FAST METHODS FOR COMPUTING THE p-RADIUS OF MATRICES[*]

RAPHAËL M. JUNGERS[†] AND VLADIMIR Y. PROTASOV[‡]

Abstract. The p-radius characterizes the average rate of growth of norms of matrices in a multiplicative semigroup. This quantity has found several applications in recent years. We raise the question of its computability. We prove that the complexity of its approximation increases exponentially with p. We then describe a series of approximations that converge to the p-radius with a priori computable accuracy. For nonnegative matrices, this gives efficient approximation schemes for the p-radius computation.

Key words. joint spectral radius, p-radius, approximation algorithm, conic programming, smoothness in L_p spaces

AMS subject classifications. 15-04, 15A60, 90C22, 26A15, 26A18, 39A99

DOI. 10.1137/090777906

1. Introduction. For a given set $\mathcal{M} = \{A_1, \ldots, A_m\}$ of $d \times d$ matrices and for a parameter $p \in [1, +\infty]$ we consider the value (where $\mathcal{M}^k$ denotes the set of all m^k products of k matrices from $\mathcal{M}$)

$$
(1.1) \qquad \rho_p =
\begin{cases}
\lim_{k \to \infty} \left[m^{-k} \sum_{B \in \mathcal{M}^k} \|B\|^p \right]^{1/(pk)}, & p < \infty, \\
\lim_{k \to \infty} \left[\max_{B \in \mathcal{M}^k} \|B\| \right]^{1/k}, & p = \infty,
\end{cases}
$$

called the L_p-*norm joint spectral radius*, or, in short, the *p-radius* of the set $\mathcal{M}$.

Thus, ρ_p is the limit of the L_p-averaged norms of all products of length k to the power $1/k$ as $k \to \infty$. This limit exists for any set $\mathcal{M}$ and does not depend on the multiplicative norm chosen. If $\mathcal{M}$ consists of one matrix A (or if all matrices of $\mathcal{M}$ coincide), then ρ_p is equal to $\rho(A)$, the spectral radius of the matrix A, which is the largest modulus of its eigenvalues.

The p-radius has found several important applications in the literature recently, but very little is known about its computability. The reason is perhaps that from a look at its definition this quantity might seem extremely difficult to compute. The goal of this paper is to survey the few results on the topic, to delineate the easy and the difficult cases, and to introduce methods for estimating the p-radius. We show that for some particular situations this leads to efficient approximation schemes.

The value ρ_∞, known as the *joint spectral radius,* was first introduced in a short paper of Rota and Strang [28] in 1960. Thirty years later Daubechies and Lagarias [10] revealed its crucial role in the theory of wavelets and functional equations. This

work attracted real interest in the joint spectral radius, which has now found many applications in various areas of mathematics, such as functional analysis, number theory, dynamical systems, and formal languages. See [15] for a monograph on this topic. Much research effort has been devoted to the computation of the joint spectral radius. This problem is known to be extremely hard for general sets of matrices; see [30] and the references therein for NP-hardness and undecidability results. Nevertheless, several algorithms have been proposed, which are efficient in many practical cases (see, for instance, [1, 12, 26]).

For finite values of the parameter p the history of the p-radius is more recent: it was introduced by Wang [31] for $p = 1$ and independently by Jia [14]. It characterizes the regularity of wavelets in L_p spaces (see [18, 22, 33] and the references therein) and the convergence of subdivision algorithms in approximation theory [14, 20], and it also rules fractal measures [24, 25, 29] and the solvability criterion for self-similarity functional equations and refinement equations with a compact mask [7, 23, 31].

In spite of the growing interest in this subject very little is known about the properties of the p-radius. It is a nondecreasing function in the parameter p and is concave in $\frac{1}{p}$ [33].

It is straightforward that for any $k \geq 1$ we have the following upper bound:

$$(1.2) \qquad \rho_p \leq U_k = \left[m^{-k} \sum_{B \in \mathcal{M}^k} \|B\|^p \right]^{1/(pk)}.$$

This is a simple consequence of submultiplicativity of the matrix norms. Since $U_k \to \rho_p$ as $k \to \infty$, this provides a theoretical opportunity to compute ρ_p. This bound U_k, however, usually becomes sharp enough only for very large values of k, which makes it impossible to compute ρ_p with some reasonable accuracy. Moreover, as no theoretical guarantee exists on the speed of convergence of this quantity, even if for some k the quantity U_k is very close to ρ_p, one could not conclude that the estimate has converged. Indeed, the approximation U_k could in theory stagnate during a long period and start decreasing again at a larger value of k. Already for $d = 2$ simple numerical examples demonstrate a very slow convergence of U_k to ρ_p.

Our main approach for deriving effective bounds for ρ_p is to find a suitable norm for which U_k rapidly converges. This idea is related to the concept of extremal norms. For any norm $\|\cdot\|$ in $\mathbb{R}^d$ and for any set $\mathcal{M}$ of m matrices, we denote by $F_p(\|\cdot\|, x)$ (abbreviated $F_p(x)$) the following function on $\mathbb{R}^d$:

$$F_p(\|\cdot\|, x) = \begin{cases} \left(\frac{1}{m} \sum_{A \in \mathcal{M}} \|Ax\|^p \right)^{1/p} & \text{if} \quad p \in [1, \infty), \\ \max_{A \in \mathcal{M}} \|Ax\| & \text{if} \quad p = \infty. \end{cases}$$

A norm $\|\cdot\|$ is called *extremal* for a given set $\mathcal{M}$ and a parameter p if there is $\lambda \geq 0$ such that $F_p(\|\cdot\|, x) = \lambda \|x\|$ for any $x \in \mathbb{R}^d$. The concept of extremal norms originated in [3] for $p = \infty$ and in [24] for all finite p. The following theorem was proved in [7] (the existence of ε-extremal norms) and in [25] (the existence of extremal norms for *irreducible* families, i.e., for matrices without common nontrivial invariant subspaces).

THEOREM 1.1. *A norm $\|\cdot\|$ is extremal if and only if for this norm $F_p(x) = \rho_p \|x\|$. If the set $\mathcal{M}$ is irreducible, then it possesses an extremal norm.*

For every $\varepsilon > 0$, every set possesses an ε-extremal norm, that is, a norm $\|\cdot\|_\varepsilon$, for which $F_p(\|\cdot\|_\varepsilon, x) \leq (\rho_p + \varepsilon)\|x\|_\varepsilon$ for any $x \in \mathbb{R}^d$.

The above theorem justifies the term "extremal" for such norms: they are the norms for which the value $\sup_{|x|=1} F_p(\|\cdot\|, x)$ is minimal (indeed this quantity is always an upper bound for ρ_p (Proposition 3.2)). Thus, knowing an extremal norm allows us to compute the value ρ_p. However, that norm is usually difficult to find. Since the initial norm in $\mathbb{R}^d$, say, the Euclidean norm, may be too far from the extremal one, estimations of type (1.2) are not sharp and the corresponding algorithms may have slow convergence. To cope with this problem, one could approximate the extremal norm, i.e., find a norm $\|\cdot\|$ for which $F_p(x)$ is close to $\lambda\|x\|$ for all x. In the case $p = \infty$ this idea was put to good use in several works, where the extremal norm was approximated by polyhedral functions, ellipsoidal norms, sums of squares polynomials, etc. [1, 4, 5, 12, 21, 23]. Recently, conic norms have been used for this purpose [26]. These methods have proved very efficient, even for large dimensions d.

We finally observe that most methods of computation that are quite efficient for the joint spectral radius ρ_∞ are inapplicable for finite p. For example, in the case $p = \infty$, one always has an easy lower bound for ρ_∞: for any k the value

$$\rho_\infty(k) = \left[\max_{B \in \mathcal{M}^k} \rho(B)\right]^{1/k},$$

where $\rho(B)$ denotes the (usual) spectral radius of the operator B, which is the largest modulus of its eigenvalues, does not exceed $\rho_\infty(\mathcal{M})$. Moreover, this lower bound converges to ρ_∞ as $k \to \infty$. In most practical cases this convergence is pretty fast, which makes $\rho_\infty(k)$ an efficient lower bound for ρ_∞.

This fact has been put to good use in most approximation algorithms for the joint spectral radius. In the case of finite p it seems natural to consider an analogous value

$$\rho_p(k) = \left[m^{-k} \sum_{B \in \mathcal{M}^k} \left(\rho(B)\right)^p\right]^{1/kp}.$$

However, simple examples show that this is not a lower bound for the p-radius. For instance, consider two 2×2 diagonal matrices $A_1 = \mathrm{diag}\,(1, \alpha)$, $A_2 = \mathrm{diag}\,(\alpha, 1)$, where $\alpha \in (0, 1)$. It is easily shown that $\rho_1(A_1, A_2) = \frac{1+\alpha}{2}$, while for $k = 1$ we have $\rho_1(1) = 1$, which is bigger than $\frac{1+\alpha}{2}$.

Another powerful tool for computing ρ_∞ is to find a set S_k of *dominating products* in $\mathcal{M}^k$ which have large spectral radius. In many cases the largest spectral radius equals $\rho_\infty(\mathcal{M})^k$ for k not too large, which provides the exact value of the joint spectral radius. Even if it does not, it often allows one to compute the dominating products of larger length by using the elements of S_k as building blocks. For finite p this idea is not applicable either, because sometimes all products might be important. In the above example there are two dominating products A_1^k, A_2^k. Both have spectral radius equal to 1 for all k. Thus, their spectral radii do not converge to ρ_p.

In this paper we are going to see that there are, nevertheless, two approaches that can be extended from the case $p = \infty$ to all finite p. These are the concepts of extremal norm and the conic programming approach. We apply these ideas to derive efficient algorithms for the p-radius computation.

In this paper, we make the first step in analyzing the computability of the p-radius for general p. We show that unless $P = NP$, there is no algorithm that approximates the p-radius of an arbitrary set of nonnegative matrices in time polynomial in the accuracy, in p, and in the size of the set of matrices. Then we propose a method for deriving bounds on the p-radius for arbitrary sets of matrices within an arbitrary

accuracy. These bounds can be computed as the solution of an optimization problem. For nonnegative matrices the optimization problem appears to be convex and unconstrained and can then be solved efficiently.

2. Computability of the p-radius. In this section we analyze the computability of the p-radius. The difficulty of the task strongly depends on p as well as on properties of the matrices. We delineate the different cases and prove that the complexity of approximating the p-radius increases exponentially as $p \to \infty$, even for nonnegative matrices (i.e., matrices with nonnegative entries). Since for such matrices and fixed $p \in \mathbb{N}$, there exist polynomial time algorithms that compute the p-radius, this settles the complexity of the problem for the case where $p \in \mathbb{N}$ and the matrices are nonnegative.

The fact that ρ_p is computable in polynomial time for nonnegative matrices and fixed $p \in \mathbb{N}$ is due to the following proposition. Particular cases of this result are proved in [22, 32].

PROPOSITION 2.1. *For any set of matrices $\mathcal{M}$ and any $p, k \in \mathbb{N}$, the p-radius of $\mathcal{M}$ satisfies*

$$\rho_p = \left[\rho_{\frac{p}{k}}(\mathcal{M}^{\otimes k}) \right]^{1/k},$$

where

$$\mathcal{M}^{\otimes p} = \{A^{\otimes p} : A \in \mathcal{M}\}$$

and $A^{\otimes p}$ denotes the pth Kronecker power[1] of A.

Proof. It is well known that Kronecker powers have the following two fundamental properties (where $\|\cdot\|$ is the Euclidean norm):

$$\|A^{\otimes k}\| = \|A\|^k,$$

$$(AB)^{\otimes k} = (A)^{\otimes k}(B)^{\otimes k}.$$

Thus,

$$(2.1) \quad \lim_{t \to \infty} m^{-t} \left[\sum_{B \in (\mathcal{M}^{\otimes k})^t} \|B\|^{p/k} \right]^{k/(kpt)} = \lim_{t \to \infty} m^{-t} \left[\sum_{B \in \mathcal{M}^t} \|B^{\otimes k}\|^{(p/k)} \right]^{1/(pt)}$$

$$= \lim_{t \to \infty} \left[m^{-t} \sum_{B \in \mathcal{M}^t} \|B\|^p \right]^{1/(pt)}. \qquad \square$$

(Note that the size of the matrices in $\mathcal{M}^{\otimes p}$ is exponential in p.) As a corollary, for any $p \in \mathbb{N}$, the p-radius can be expressed as the 1-radius of an auxiliary set of matrices.

COROLLARY 2.2 (see [22, 32]). *For any set of matrices $\mathcal{M}$ and any $p \in \mathbb{N}$, the p-radius of $\mathcal{M}$ satisfies*

$$\rho_p = \rho_1(\mathcal{M}^{\otimes p})^{1/p}.$$

---

[1]The kth Kronecker power of an $n \times n$ matrix is an $n^k \times n^k$ matrix, defined inductively as $A^{\otimes k} = A \otimes A^{\otimes k-1}$, where $A \otimes B$ is the *Kronecker product* of A and B: $A \otimes B = \begin{pmatrix} A_{(1,1)}B & \cdots & A_{(1,n)}B \\ \vdots & \vdots & \vdots \\ A_{(n,1)}B & \cdots & A_{(n,n)}B \end{pmatrix}$.

In some cases the value ρ_1 can be expressed as the usual spectral radius of a certain matrix, which we now explain. Let K be a convex closed pointed cone in $\mathbb{R}^d$ with a nonempty interior (these assumptions on K are crucial in the remainder of the paper, and we will make them from now on). We say that a set of matrices $\mathcal{M}$ leaves K invariant if $AK \subset K$ for all $A \in \mathcal{M}$.

PROPOSITION 2.3 (see [4, 22]). *If a set of m matrices $\mathcal{M}$ leaves a cone K invariant, then*

$$\rho_1(\mathcal{M}) = \rho \left(\frac{1}{m} \sum_{A \in \mathcal{M}} A \right).$$

Putting Propositions 2.1 and 2.3 together, one gets the following theorem.

THEOREM 2.4 (see [22, 32]). *Let $\mathcal{M}$ be a set of matrices and let $p \in \mathbb{N}$. If either*

- *the set $\mathcal{M}$ leaves a cone K invariant, or*
- *p is even*

holds, then

$$\rho_p(\mathcal{M}) = \rho \left(\frac{1}{m} \sum_{A^{\otimes p} \in \mathcal{M}} A \right)^{1/p}.$$

Proof. For the first case, just put Propositions 2.1 and 2.3 together and use the fact that if $\mathcal{M}$ leaves K invariant, then $\mathcal{M}^{\otimes k}$ leaves $\mathrm{conv}\,\{K^{\otimes k}\}$ invariant.

For the second case, note that $\mathcal{M}^{\otimes 2p} = (\mathcal{M}^{\otimes p})^{\otimes 2}$ and that, given a set of matrices $\mathcal{M}'$ acting in $\mathbb{R}^d$, the set $\mathcal{M}'^{\otimes 2}$ leaves invariant the cone $K = \mathrm{conv}\,\{x \otimes x : x \in \mathbb{R}^d\}$. See [4, 22, 32] for details. □

So, if $p \in \mathbb{N}$ and the matrices are nonnegative (or, more generally, for sets of matrices with an invariant cone), there is an algorithm that computes the p-radius in finite (but exponential) time. We now show that this algorithm is optimal, since the p-radius computation is NP-hard. Actually, we show that it is even NP-hard to approximate it. In the following, we refer to the size of a rational number as its "bit size," that is, if $\epsilon = p/q$ for two natural numbers $p, q \in \mathbb{N}$, its size is $\log(pq)$.

THEOREM 2.5. *Unless $P = NP$, there is no algorithm that takes a set of matrices $\mathcal{M}$, an accuracy ϵ, and an integer p, and that returns an approximation ρ_p^* such that $|\rho_p - \rho_p^*|/\rho_p < \epsilon$ in a time which is polynomial in the size of the instance. This is true even if $\mathcal{M}$ consists of two binary matrices.*

Proof. We reduce the SAT problem, whose NP-completeness is well known, to the p-radius approximation. The construction of the reduction is the same as in [30, Theorem 1]. In that paper, the authors show that, given an instance of SAT with m clauses and n unknown variables, it is possible to construct in polynomial time a set of two matrices $\mathcal{M}$ whose joint spectral radius is equal to $m^{1/(n+2)}$ if and only if the instance is satisfiable. In the opposite case, $\rho(\Sigma) \leq (m-1)^{1/(n+2)}$ (we use $\rho(\Sigma)$ for $\rho_\infty(\Sigma)$ as is customary in the literature). Thus, if the instance is satisfiable, it implies (see [15]) that there is a constant K_1 such that, for all t, there is a product of length t with a norm greater than $K_1 m^{1/(n+2)}$. Thus, we have

$$\rho_p \geq \lim_{t \to \infty} \{K_1 m^{tp/(n+2)}/2^t\}^{1/(pt)} = m^{1/(n+2)}/2^{1/p}.$$

If $\rho(\Sigma) \leq (m-1)^{1/(n+2)}$, there is a constant K_2 such that, for all t, all products of length t have a norm which is smaller than $K_2(m-1)^{tp/(n+2)}t^{K_2}$. Thus we have

$$\rho_p \leq \lim_{t \to \infty} \{K_2(m-1)^{tp/(n+2)}t^{K_2}\}^{1/(pt)} = (m-1)^{1/(n+2)}.$$

Then, taking $p = m(n+2) > \frac{n+2}{\log(m/(m-1))}$, we have

$$\rho_p \geq m^{1/(n+2)}/2^{1/p}$$

if the instance is satisfiable, and

$$\rho_p \leq (m-1)^{1/(n+2)} < m^{1/(n+2)}/2^{1/p}$$

in the opposite case (logarithms are base 2). Now, if there existed a polynomial time algorithm to approximate the p-radius, choosing ϵ small enough (for instance $\epsilon < ((m^{1/(n+2)}/2^{1/p})/(m-1)^{1/(n+2)}) - 1)$, one could check whether $\rho_p = m^{1/(n+2)}/2^{1/p}$ and then check whether the instance is satisfiable. ◻

OPEN QUESTION 1. *Does there exist an algorithm that decides, given a set of matrices, whether its p-radius is less than one?*

OPEN QUESTION 2. *Does there exist an algorithm that decides, given a set of matrices, whether its p-radius is less than or equal to one?*

For matrices with an invariant cone and a finite nonnegative integer p, or for even p, the answer is affirmative, since the p-radius can be expressed as the spectral radius of a special matrix, as we have seen above. Also, Open Question 2 is known to be undecidable for $p = \infty$, even if the matrices are nonnegative. As far as we know, the other cases are open. See [8, 16] for a study of the first question with $p = \infty$.

3. Approximating the p-radius. In this section we describe four quantities that approximate the p-radius within an approximation error of at most $1/d$, where d is the dimension of the matrices. Then we show how to iterate this in order to get more and more precise approximations. We start with two auxiliary results. The first is a consequence of Theorem 1.1.

LEMMA 3.1 (see [25]). *Let $\mathcal{M}$ be a set of matrices. For any $x \in \mathbb{R}^d$ there is a constant C_1 depending only on $\mathcal{M}$ and x such that, for all $k \geq d$,*

$$\left(m^{-k} \sum_{B \in \mathcal{M}^k} \|Bx\|^p \right)^{1/p} \leq C_1 k^{d-1}(\rho_p)^k.$$

Moreover, if x does not belong to a common invariant subspace of the matrices of the set $\mathcal{M}$, then there is a constant C_2 depending only on $\mathcal{M}$ and x such that

$$C_2(\rho_p)^k \leq \left(m^{-k} \sum_{B \in \mathcal{M}^k} \|Bx\|^p \right)^{1/p}.$$

Let $V_p(\mathcal{M}) = \sup_{\|x\|=1} F_p(x)$. This value depends on p, $\mathcal{M}$, and the norm $\|\cdot\|$.

PROPOSITION 3.2. *For any norm one has $V_p \geq \rho_p$. For extremal norms this inequality becomes an equality.*

Proof. We claim that for any x, $\|x\| = 1$, and for any k,

$$(V_p)^k \geq \left(m^{-k} \sum_{B \in \mathcal{M}^k} \|Bx\|^p \right)^{1/p}.$$

We prove it by induction on k. For $k = 1$ there is nothing to prove, by definition of V_p. Now,

$$
\begin{aligned}
(3.1) \quad m^{-k} \sum_{B \in \mathcal{M}^k} \|Bx\|^p &= m^{-1} \sum_{A \in \mathcal{M}} m^{-(k-1)} \sum_{B \in \mathcal{M}^{k-1}} \|BAx\|^p, \\
&= m^{-1} \sum_{A \in \mathcal{M}} m^{-(k-1)} \sum_{B \in \mathcal{M}^{k-1}} \|By\|^p \|Ax\|^p, \\
&= m^{-1} \sum_{A \in \mathcal{M}} \|Ax\|^p \left[m^{-(k-1)} \sum_{B \in \mathcal{M}^{k-1}} \|By\|^p \right], \\
&\leq (V_p V_p^{k-1})^p.
\end{aligned}
$$

In the above equations the vector y refers to $(Ax)/\|Ax\|$. Now, we can suppose without loss of generality that there is a vector that does not belong to a common invariant subspace. Thus, by Lemma 3.1 there is $x \in \mathbb{R}^d, \|x\| = 1$, such that

$$
(V_p)^k \geq \left(m^{-k} \sum_{B \in \mathcal{M}^k} \|Bx\|^p \right)^{1/p} \geq C(\rho_p)^k, \qquad k \in \mathbb{N}.
$$

As $k \to \infty$, this inequality implies that $V_p \geq \rho_p$. □

The following analysis will be restricted to families $\mathcal{M}$ with a common invariant cone.

As usual, any cone K defines a corresponding order in the following way: for any $x, y \in \mathbb{R}^d$, $x \geq_K y$ if $x - y \in K$, and $x >_K y$ if $x - y \in \mathrm{int} K$. Thus, $x \geq_K 0$ if $x \in K$. For a matrix A we write $A \geq_K 0$ if $AK \subset K$. We deal with norms $\| \cdot \|$ defined for $x \in K$. Any such a norm can easily be extended onto the whole space $\mathbb{R}^d$ in a standard way by defining its unit ball as $\mathbf{co}\{B_K, -B_K\}$, where $\mathbf{co}$ is the convex hull and $B_k = \{x \in K : \|x\| \leq 1\}$. We spot two special families of norms in the cone K. The *primal conic norms* $r_x(\cdot)$ are defined by unit balls $B_K = K \cap (x - K)$, where $x \in \mathrm{int} K$ is a given point. Thus, $r_x(y) = \inf \{ \lambda > 0 \mid y \leq_K \lambda x \}$. The *dual conic norms* are defined by unit balls $B_K^* = \{ y \in K \mid (v, y) \leq 1 \}$, where $v \in \mathrm{int} K^*$ is an arbitrary functional ($K^* = \{ v \in \mathbb{R}^d \mid \inf_{x \in K} (v, x) \geq 0 \}$ denotes as usual the cone dual to K). For self-dual cones K (for, instance, if $K = \mathbb{R}_+^d$) these norms are dual to each other (with $x = v$).

In what follows, we need the following geometrical characteristic of a cone.

DEFINITION 3.3. *For a given cone $K \subset \mathbb{R}^d$ the value $h(K)$ is the supremum of the numbers h possessing the following property: for any norm $\| \cdot \|$ in $\mathbb{R}^d$ there exist points $x \in \mathrm{int} K, v \in \mathrm{int} K^*$ such that $\|x\| = 1$ and for any $y \in K$,*

$$
(3.2) \qquad\qquad hr_x(y) \leq (v, y) \leq \|y\|.
$$

When the choice of the cone K is clear from the context, we simply write h for $h(k)$.

The constant h is related with other geometrical characteristics of cones such as the parameter of self-concordant barriers [19]. It is related to the constant $\alpha(K)$ introduced in [26] (see this reference for more information on this constant). The key property that we need for h is the following, largely inspired from [26, Theorem 2.7].

PROPOSITION 3.4. *For any cone $K \subset \mathbb{R}^d$, one has $h(K) \geq \frac{1}{d}$.*

Proof. Let B be the unit ball of a given norm. We define v and x as follows. Among all hyperplanes that do not intersect $\mathrm{int} B$ there exists a hyperplane $H = \{ y \in \mathbb{R}^d \mid (v, y) = 1 \}$ that cuts from K a convex compact set of the smallest possible volume. Then the center of gravity $x = \mathrm{gr}(G)$ of the $(d-1)$-dimensional cross-section $G = K \cap H$ belongs to B [2, p. 229].

Clearly, $x \in \operatorname{int} K$, and $v \in \operatorname{int} K^*$ since G is bounded. Also, the second inequality in Definition 3.3 (equation (3.2)) is satisfied by construction.

We now prove the first inequality: $r_x(y) \le d(v, y)$ for any $y \in K$. This means that $d(v, y) x \ge_K y$ for any $y \in K$, or, writing $z = y/(v, y)$, we obtain $d x \ge_K z$ for any $z \in G$. Now we apply the Minkowski–Radon theorem [27], which states that, for any segment $[z_1, z_2]$ passing through the center of gravity x of a $(d-1)$-dimensional convex set and having its ends on the boundary of G, one has $|z_1 - x| \le (d-1)|z_2 - x|$. This implies that for any $z \in G$ one has $x + \frac{1}{d-1}(x - z) \in G$. Therefore, $\frac{1}{d-1}(d x - z) \in K$, and so $d x \ge_K z$, which completes the proof. $\qquad\square$

Remark 1. It is possible to show (see [26, Theorem 2.7]) that for any polyhedral cone in dimension d, as well as for the $\frac{d^2+d}{2}$-dimensional cone $\mathcal{K}_d$ of symmetric positive semidefinite $d \times d$ matrices, the constant h is equal to $\frac{1}{d}$. For any Lorentz (or spherical) cone $h = \frac{1}{2}$.

3.1. The primal conic radius α_p. We consider a set $\mathcal{M}$ with an invariant cone K. For any point $x >_K 0$ we first focus on the primal conic norm $r_x(\cdot)$ on K. The corresponding induced matrix norm appears to be particularly easy to evaluate: indeed, $y \le_K x$ implies that $Ay \le_K Ax$ for any matrix $A \ge_K 0$, and hence $r_x(Ay) \le r_x(Ax)$. Thus the corresponding induced matrix norm is

$$\|A\|_{r_x} = \sup_{r_x(y) \le 1, y \ge_K 0} r_x(Ay) = r_x(Ax).$$

As we have seen above (Proposition 3.2), any norm provides an upper bound on ρ_p, and thus

$$(3.3) \qquad \alpha_p(x) = \sup_{y \in K, r_x(y) \le 1} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left[r_x(Ay) \right]^p \right]^{1/p}$$

$$(3.4) \qquad = \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left[r_x(Ax) \right]^p \right]^{1/p}$$

$$(3.5) \qquad \ge \rho_p.$$

Then, minimizing over x in order to get the best upper bound, one obtains

$$(3.6) \qquad \alpha_p = \inf_{x >_K 0} \alpha_p(x) \ge \rho_p.$$

We call α_p the primal conic radius of the set $\mathcal{M}$. The ratio between α_p and ρ_p can be estimated in terms of the constant h defined above (Definition 3.3).

THEOREM 3.5. *For any set $\mathcal{M}$ with an invariant cone K we have $h\alpha_p \le \rho_p \le \alpha_p$.*

Proof. We need to prove that $h\alpha_p \le \rho_p$. For some small $\varepsilon > 0$ consider an ε-extremal norm $\|\cdot\|_\varepsilon$ for the set $\mathcal{M}$ (Theorem 1.1). Take a point x corresponding to this norm, for which $\|x\|_\varepsilon = 1$ and $(h - \varepsilon)r_x(y) \le \|y\|_\varepsilon$ for all $y \in K$ (Definition 3.3). We have

$$\rho_p + \varepsilon \ge \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \|Ax\|_\varepsilon^p \right]^{1/p} \ge \left[\frac{1}{m} \sum_{A \in \mathcal{M}} (h - \varepsilon)^p \left(r_x(Ax) \right)^p \right]^{1/p} \ge (h - \varepsilon) \alpha_p.$$

Taking now $\varepsilon \to 0$, we complete the proof. $\qquad\square$

Invoking now Proposition 3.4, we obtain the following corollary.

COROLLARY 3.6. *For any set $\mathcal{M} \subset \mathbb{R}^{d \otimes d}$ with an invariant cone K we have*

$$\frac{1}{d} \, \alpha_p \leq \rho_p \leq \alpha_p.$$

We end this subsection by commenting on the computability of α_p. As we have seen, given a vector x, the optimization problem (3.3) is trivial (its solution is x), and so it is very easy to obtain the corresponding upper bound $\alpha_p(x)$ on ρ_p. However, in order to compute α_p (to have a guaranteed accuracy), one then has to solve the optimization problem (3.6) on x. In general, $\alpha_p(x)$ is not even quasi-convex, and so the computation of the real value α_p for general cones might be difficult. We will see below particular cones for which the situation is much more appealing.

3.2. The dual conic radius β_p. We now do a similar analysis for the dual estimate β_p. To an arbitrary $v \in \operatorname{int} K^*$ we associate the dual conic norm (v, x) on K. Let us introduce

$$(3.7) \qquad \beta_p(v) = \sup_{x >_K 0, (v,x) \leq 1} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} (v, Ax)^p \right]^{1/p}$$

$$(3.8) \qquad = \sup_{x >_K 0} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, Ax)^p}{(v, x)^p} \right]^{1/p}$$

and

$$\beta_p = \inf_{v >_{K^*} 0} \beta_p(v).$$

Again, by Proposition 3.2, $\beta_p(v) \geq \rho_p$ for any v. And again, it appears that β_p is within the same approximation factor h as α_p. To show this we start with more auxiliary results.

We call a norm $\| \cdot \|$ *consistent* with a cone K if it is monotone in the cone, i.e., for any $x, y \in K$ the relation $x \geq_K y$ implies $\|x\| \geq \|y\|$. Geometrically this means that for any point $x \in K$ from the unit ball the set $(x - K) \cap K$ is contained in the unit ball.

PROPOSITION 3.7. *For any set $\mathcal{M}$ with an invariant cone K and for any $\varepsilon > 0$ there exists a consistent ε-extremal norm.*

Proof. For an arbitrary $v \in \operatorname{int} K^*$ let

$$f_\lambda(x) = \left(\sum_{k=1}^{\infty} (\lambda m)^{-k} \sum_{B \in \mathcal{M}^k} (v, Bx)^p \right)^{1/p}.$$

By Lemma 3.1 this series converges for any $\lambda > \rho_p^p$. The function $f_\lambda(x)$ is convex in x, because it is the l_p-norm of the two-indexed sequence

$$\{(\lambda m)^k (v, Bx)\}, \quad B \in \mathcal{M}^k, \ k \in \mathbb{N}$$

(see [6, Example 3.14]). This function is also positively homogeneous and monotone on the cone K; moreover,

$$\frac{1}{m} \sum_{A \in \mathcal{M}} f_\lambda^p(Ax) + \frac{1}{m} \sum_{A \in \mathcal{M}} (v, Ax)^p = \lambda [f_\lambda(x)]^p,$$

and therefore

$$\left(\frac{1}{m} \sum_{A \in \mathcal{M}} f_\lambda^p(Ax) \right)^{1/p} \leq \lambda^{1/p} f_\lambda(x).$$

Taking now $\lambda = (\rho_p + \varepsilon)^p$, we obtain an ε-extremal consistent norm $f_\lambda(\cdot)$ defined on the cone K. We then extend it in a standard way to the whole space $\mathbb{R}^d$. $\square$

PROPOSITION 3.8. *For any cone $K \subset \mathbb{R}^d$, for any consistent norm $\| \cdot \|$ in $\mathbb{R}^d$, and for any $\varepsilon > 0$, there is a vector $v \in \mathrm{int} K^*$ such that $(h - \varepsilon)\|y\| \leq (v, y) \leq \|y\|$ for all $y \in K$.*

Proof. By Definition 3.3, there are points $x \in \mathrm{int} K, v \in \mathrm{int} K^*$ such that $\|x\| = 1$ and $(h - \varepsilon) r_x(y) \leq (v, y) \leq \|y\|$ for all $y \in K$. If the norm is consistent, then $r_x(y) \geq \|y\|$, which concludes the proof. $\square$

THEOREM 3.9. *For any set $\mathcal{M}$ with an invariant cone K we have $h\beta_p \leq \rho_p \leq \beta_p$.*

Proof. Fix $\varepsilon > 0$ and consider a consistent ε-extremal norm $\| \cdot \|_\varepsilon$ of the set $\mathcal{M}$, which exists by Proposition 3.7. Take the corresponding $v \in \mathrm{int} K^*$ from Proposition 3.8. For any $y \in \mathrm{int} K$ we have $\|Ay\|_\varepsilon \geq (v, Ay)$ and $(v, y) \geq (h - \varepsilon)\|y\|_\varepsilon$. Combining these two inequalities, we get $\frac{\|Ay\|_\varepsilon}{\|y\|_\varepsilon} \geq (h - \varepsilon)\frac{(v, Ay)}{(v, y)}$. It now follows that for all $y \in \mathrm{int} K$,

$$\rho_p + \varepsilon \geq \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{\|Ay\|_\varepsilon^p}{\|y\|_\varepsilon^p} \right]^{1/p} \geq \left[\frac{1}{m} \sum_{A \in \mathcal{M}} (h - \varepsilon)^p \frac{(v, Ay)^p}{(v, y)^p} \right]^{1/p}.$$

The supremum of the right-hand side over all $y >_K 0$ equals $(h - \varepsilon) \beta_p(v)$. Thus, for any $\varepsilon > 0$

$$\rho_p + \varepsilon \geq (h - \varepsilon) \beta_p(v) \geq (h - \varepsilon) \beta_p,$$

from which the theorem follows. $\square$

COROLLARY 3.10. *For any set $\mathcal{M}$ with an invariant cone K we have $\frac{1}{d}\beta_p \leq \rho_p \leq \beta_p$.*

In the next section we compare more precisely the quality of the estimates α_p and β_p.

We end this subsection by commenting on the computational aspect for the value β_p. The function

$$\beta_p(v) = \sup_{x >_K 0} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, Ax)^p}{(v, x)^p} \right]^{1/p}$$

is quasi-convex. To see this, consider the level sets of the function $f(v) = [\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, Ax)^p}{(v, x)^p}]^{1/p}$. The inequality $f(v) \leq c$ is equivalent to

$$\left[\frac{1}{m} \sum_{A \in \mathcal{M}} (v, Ax)^p \right]^{1/p} \leq c(v, x).$$

The left-hand side is convex in v, and the right-hand side is linear, so the set of solutions v of this inequality is convex. Therefore, the function $\beta_p(v)$ is quasi-convex, being a supremum of quasi-convex functions. Hence the value $\beta_p = \inf_{v >_K 0} \beta_p(v)$

can be effectively computed, provided that we can compute $\beta_p(v)$ for any v. This problem, however, may be hard for general cones. Indeed, in order to compute $\beta_p(v)$ we need to find the maximum of a quasi-convex function (see (3.7)) on the cone K, for which there are no efficient algorithms. Already for $p = 4$, $K = \mathcal{K}_d$ (i.e., K is the cone of positive semidefinite $d \times d$ matrices), this problem amounts to maximizing a homogeneous polynomial of degree 4 on a linear subset of $\mathcal{K}_d$, which is known to be NP-hard.

Nevertheless, for polyhedral cones this problem is effectively solvable. In this case the cone K has a finite number of extreme directions; i.e., K is the conic hull of several points $\{x_j\}_{j=1}^N$. Whence the supremum of a quasi-convex homogeneous function is attained at one of these points:

$$(3.9) \qquad \beta_p(v) = \max_{j=1,\ldots,N} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, Ax_j)^p}{(v, x_j)^p} \right]^{1/p}.$$

Thus, this supremum over all $x >_K 0$ can be replaced with the maximum over a finite number of points x_i that can be found by exhaustion for any fixed v. We show this in detail for the case $K = \mathbb{R}_+^d$ is section 4.

3.3. The transposed quantities α_p^*, β_p^* and the algorithm. In this subsection we first establish relations between the primal and dual conic radii. For this we introduce some more notation. Let A^* denote the adjoint matrix of A, and let $\mathcal{M}^*$ be the set of adjoint matrices of the matrices in $\mathcal{M}$, and let $h^* = h(K^*)$. Clearly, $\rho_p(\mathcal{M}) = \rho_p(\mathcal{M}^*)$, and if the set $\mathcal{M}$ possesses an invariant cone K, then K^* is an invariant cone for $\mathcal{M}^*$. Finally denote $\alpha_p^* = \alpha_p(\mathcal{M}^*, K^*)$ and similarly with β_p^*. Applying Theorem 3.5 and taking into account that $\rho_p(\mathcal{M}) = \rho_p(\mathcal{M}^*)$, we obtain the following corollary.

COROLLARY 3.11. *For any set $\mathcal{M}$ with an invariant cone K we have*

$$h^* \beta_p^* \leq \rho_p \leq \beta_p^*, \qquad h^* \alpha_p^* \leq \rho_p \leq \alpha_p^*.$$

It appears that the upper bounds β_p and β_p^* are always tighter than α_p^* and α_p, respectively.

PROPOSITION 3.12. *For any set $\mathcal{M}$ with an invariant cone K we have $\beta_p(v) \leq \alpha_p^*(v)$ for all $v \in \mathrm{int}K^*$ and $\beta_p^*(x) \leq \alpha_p(x)$ for all $x \in \mathrm{int}K$.*

Proof. Let us establish the second inequality; the proof of the first is the same. Since $z \in K$ if and only if $\inf_{v \in K^*}(v, z) \geq 0$, we see that $y \leq_K \lambda x$ for $x, y \in K$ if and only if $(v, y) \leq \lambda(v, x)$ for all $v \in K^*$. Therefore, $r_x(y) = \sup_{v >_{K^*} 0} \frac{(v, y)}{(v, x)}$. Consequently,

$$(3.10) \qquad \alpha_p(x) = \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(r_x(Ax) \right)^p \right]^{1/p}$$

$$(3.11) \qquad = \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\sup_{v >_{K^*} 0} \frac{(v, Ax)}{(v, x)} \right)^p \right]^{1/p}$$

$$(3.12) \qquad = \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\sup_{v >_{K^*} 0} \frac{(A^*v, x)}{(v, x)} \right)^p \right]^{1/p}$$

$$(3.13) \qquad \geq \sup_{v >_{K^*} 0} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(A^*v, x)^p}{(v, x)^p} \right]^{1/p} = \beta_p^*(x). \qquad \square$$

COROLLARY 3.13. *For any set $\mathcal{M}$ with an invariant cone K we have*

$$\beta_p \leq \alpha_p^*, \qquad \beta_p^* \leq \alpha_p.$$

Remark 2. For $p = \infty$ we have $\beta_p = \alpha_p^*$ and $\beta_p^* = \alpha_p$. This is because in this case inequality (3.13) above becomes an equality. Therefore, for $p = \infty$ the primal conic radius is equal to the dual conic radius of the adjoint matrices.

Corollary 3.13 yields that the values β_p, β_p^* provide better upper bounds for ρ_p than the values α_p, α_p^*, but the latter may provide better lower bounds. We obtain the following theorem.

THEOREM 3.14. *For any set $\mathcal{M}$ with an invariant cone K we have*

$$(3.14) \qquad\qquad a_p \leq \rho_p \leq b_p,$$

where

$$(3.15) \qquad b_p = \min\{\beta_p, \beta_p^*\}, \qquad a_p = \max\{h\alpha_p, h\beta_p, h^*\alpha_p^*, h^*\beta_p^*\}.$$

Observe that $a_p \geq \max\{h, h^*\}b_p$. Hence the ratio between the lower and upper bounds is greater than or equal to $\max\{h, h^*\}$, which is always at least $\frac{1}{d}$.

Remark 3. In the case $p = 1$ the upper bound in (3.14) is always sharp. The reader will have no difficulty in showing that $\beta_1 = \beta_1^* = \rho_1$.

Theorem 3.14 allows us to compute the p-radius of matrices that possess an invariant cone with any prescribed relative accuracy $\varepsilon > 0$. To do this we apply inequality (3.14) to the set $\mathcal{M}^k$ that consists of all m^k products of length k of matrices from $\mathcal{M}$. Taking into account that $\rho_p(\mathcal{M}^k) = [\rho_p(\mathcal{M})]^k$, we obtain the following bounds for $\rho_p(\mathcal{M})$:

$$(3.16) \qquad\qquad \left[a_p(\mathcal{M}^k)\right]^{1/k} \leq \rho_p(\mathcal{M}) \leq \left[b_p(\mathcal{M}^k)\right]^{1/k}.$$

Since the families $\mathcal{M}^k$ and $(\mathcal{M}^*)^k$ still have the same invariant cones K and K^*, respectively, we can use the same constants $h(K)$ and $h(K^*)$ and obtain

$$\left(\max\{h, h^*\}\right)^{1/k}\left[b_p(\mathcal{M}^k)\right]^{1/k} \leq \left[a_p(\mathcal{M}^k)\right]^{1/k}.$$

Thus, the ratio between the lower and upper bounds is greater than or equal to $(\max\{h, h^*\})^{1/k}$, which is always at least $d^{-1/k}$. Therefore, the estimate (3.16) of the p-radius has a relative error bounded by

$$(3.17) \qquad\qquad 1 - \left(\max\{h, h^*\}\right)^{1/k} \leq -\frac{1}{k}\ln\max\{h, h^*\},$$

which is smaller than or equal to $\frac{\ln d}{k}$. As we shall see in section 5, in practice this estimate can even be much better.

4. Nonnegative matrices. In this section we focus on the case $K = \mathbb{R}_+^d$ and show that in this case both α_p and β_p can be found as solutions of convex minimization problems. In this case $h = h^* = \frac{1}{d}$ (Remark 1). Hence formula (3.15) for the bounds a_p and b_p becomes

$$(4.1) \qquad\qquad b_p = \min\{\beta_p, \beta_p^*\}, \qquad a_p = \frac{1}{d}\max\{\alpha_p, \alpha_p^*\}.$$

Let us start with the primal conic radius α_p. We have $r_x(y) = \max_{i=1,\ldots,d} \frac{y_i}{x_i}$, and then

$$\alpha_p = \inf_{x >_{\mathbb{R}^d_+} 0} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(r_x(Ax) \right)^p \right]^{1/p}$$

$$= \inf_{x >_{\mathbb{R}^d_+} 0} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\max_{i=1,\ldots,d} \frac{(Ax)_i}{x_i} \right)^p \right]^{1/p}.$$

Each function $\frac{(Ax)_i}{x_i}$ is not convex in x, but it is quasi-convex. Therefore, so is the function $\max_{i=1,\ldots,d} \frac{(Ax)_i}{x_i}$ as a maximum of quasi-convex functions. However, $\alpha_p(x)$ as an L_p-average of quasi-convex functions is a priori not quasi-convex. This difficulty can be overcome by the exponential change of variables $x_i = e^{u_i}$, $i = 1, \ldots, d$. We have

$$(4.2) \qquad \alpha_p = \inf_{u \in \mathbb{R}^d} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\max_{i=1,\ldots,d} (a_i, e^u) e^{-u_i} \right)^p \right]^{1/p}.$$

Note that all the functions $(a_i, e^u) e^{-u_i} = \sum_{j=1}^d a_{ij} e^{u_j - u_i}$ are convex in u, and hence their maximum $\max_{i=1,\ldots,d} (a_i, e^u) e^{-u_i}$ is also convex, and so is the function $\alpha_p(x)$ as an L_p-average of nonnegative convex functions. Thus, we have the following theorem.

THEOREM 4.1. *For any $p \in [1, +\infty]$ and for any set of nonnegative matrices $\mathcal{M}$, the value $\alpha_p(\mathbb{R}^d_+, \mathcal{M})$ is a solution of the following unconstrained convex minimization problem:*

$$(4.3) \qquad \alpha_p = \inf_{u \in \mathbb{R}^d} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\max_{i=1,\ldots,d} \sum_{j=1}^d a_{ij} e^{u_j - u_i} \right)^p \right]^{1/p}.$$

Now consider the dual conic radius β_p. In the case $K = \mathbb{R}^d_+$, formula (3.9) becomes

$$\beta_p(v) = \max_{j=1,\ldots,d} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, a^j)^p}{v_j^p} \right]^{1/p},$$

where a^j denotes the jth column of the matrix A. Thus, the value β_p is the solution of a quasi-convex minimization problem:

$$(4.4) \qquad \beta_p = \inf_{v >_{\mathbb{R}^d_+} 0} \max_{j=1,\ldots,d} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \frac{(v, a^j)^p}{v_j^p} \right]^{1/p}.$$

Again, we can apply the exponential change of variables $v_j = e^{z_j}$, after which $\frac{(v, a^j)^p}{v_j^p} = (a^j, e^z)^p e^{-p z_j} = (\sum_{i=1}^d a_{ij} e^{z_i - z_j})^p$. We thus obtain the following theorem.

THEOREM 4.2. *For any $p \in [1, +\infty]$ and for any set of nonnegative matrices $\mathcal{M}$ the value $\beta_p(\mathbb{R}^d_+, \mathcal{M})$ is a solution of the following unconstrained convex minimization problem:*

$$(4.5) \qquad \beta_p = \inf_{z \in \mathbb{R}^d} \max_{j=1,\ldots,d} \left[\frac{1}{m} \sum_{A \in \mathcal{M}} \left(\sum_{i=1}^d a_{ij} e^{z_i - z_j} \right)^p \right]^{1/p}.$$

We now make a last improvement on the estimation (4.1). It appears that in the case $K = \mathbb{R}^d_+$, the lower bound of Corollary 3.10 can be sharpened.

THEOREM 4.3. *For any set of nonnegative matrices one has*

$$d^{\frac{1}{p}-1}\beta_p \leq \rho_p \leq \beta_p.$$

Thus, for any p we have

$$(4.6) \qquad \max\{d^{\frac{1}{p}-1}\beta_p, d^{\frac{1}{p}-1}\beta_p^*, d^{-1}\alpha_p, d^{-1}\alpha_p^*\} \leq \rho_p \leq \min\{\beta_p, \beta_p^*\}.$$

For small values of p this gives a significant improvement. The ratio between the upper and lower bounds does not exceed $d^{1-\frac{1}{p}}$. We briefly sketch the proof of the above theorem. In the following we call a set $\mathcal{M}$ *reducible* if all of its matrices get a block upper-triangular form after some permutation of coordinates.

Proof. The proof is based on the following successive arguments, which are of independent interest:

- *We can assume without loss of generality that our set $\mathcal{M}$ is irreducible.* Indeed it is known that if there is a permutation that puts all matrices in $\mathcal{M}$ in the same upper-triangular form, then

$$\rho_p(\mathcal{M}) = \max_{i=1,\dots,s} \rho_p(\mathcal{M}_i),$$

  where $\{\mathcal{M}_i\}$ is the ith set of diagonal blocks [22]. Moreover the permutation can be found in polynomial time [17].
- *If the set $\mathcal{M}$ is irreducible, then the infimum β_p in (4.4) is attained at some point $\bar{v} >_{\mathbb{R}^d_+} 0$, and, moreover,*

$$\left(\frac{1}{m}\sum_{A \in \mathcal{M}}(\bar{v}, a^j)^p\right)^{1/p} = \beta_p \bar{v}_j$$

  for each $j = 1, \dots, d$.
  It is not difficult to see that if $\mathcal{M}$ is irreducible, $\beta_p(v)$ is finite only for $v >_{\mathbb{R}^d_+} 0$, and that if for some j

$$(4.7) \qquad \left(\frac{1}{m}\sum_{A \in \mathcal{M}}(\bar{v}, a^j)^p\right)^{1/p} < \beta_p \bar{v}_j,$$

  then it is possible to decrease $\beta_p(\bar{v})$ by decreasing $\bar{v}_j$.
- *The above point implies that $d^{\frac{1}{p}-1}\beta_p \leq \rho_p$.*
  We have the following inequalities:

$$(4.8) \qquad \rho_p \geq \inf_{y \geq_{\mathbb{R}^d_+} 0, (\tilde{v},y)=1} \left(\frac{1}{m}\sum_{A \in \mathcal{M}}(\tilde{v}, Ay)^p\right)^{1/p}$$

$$= \left(\frac{1}{m}\sum_{A \in \mathcal{M}}(\tilde{v}, Ax)^p\right)^{1/p} \qquad \text{for some } x \geq_{\mathbb{R}^d_+} 0$$

$$(4.9) \qquad \geq d^{\frac{1}{p}-1} \sum_{j=1}^{d} \left(\frac{1}{m} \sum_{A \in \mathcal{M}} \big(\tilde{v}, A(x_j e_j) \big)^p \right)^{1/p}$$

$$= \sum_{j=1}^{d} x_j \left(\frac{1}{m} \sum_{A \in \mathcal{M}} (\tilde{v}, a^j)^p \right)^{1/p}$$

$$= d^{\frac{1}{p}-1} \sum_{j=1}^{d} \beta_p(\tilde{v}) x_j v_j = d^{\frac{1}{p}-1} \beta_p(\tilde{v}).$$

Inequality (4.8) is well known; it is proved in the same way as Proposition 3.2. In order to obtain inequality (4.9), we write $\big(\tilde{v}, Ax \big)$ as $\sum_{j=1}^{d} \big(\tilde{v}, A(x_j e_j) \big)$, and use the well-known inequality $\big\| \sum_{j=1}^{d} b_j \big\|_p \geq d^{\frac{1}{p}-1} \sum_{j=1}^{d} \| b_j \|_p$, valid for nonnegative vectors b_j, with b_j the vector in $\mathbb{R}^m$ with entries

$$m^{-1/p} \big(\tilde{v}, A(x_j e_j) \big), \quad A \in \mathcal{M}. \qquad \square$$

We now comment on the algorithmic complexity of our method. As we have seen ((3.17) and the equation below), it is possible to obtain an estimate with relative accuracy ϵ by applying our methods to $\mathcal{M}^k$, provided that $k \geq \frac{\ln d}{\epsilon}$. The cost of such an estimate is mainly to compute the set $\mathcal{M}^k$ and then to perform computations on each of these matrices to compute the function $\alpha_p(u)$ (these computations take polynomial time). Then, since $\alpha_p(u)$ is convex, only a polynomial (in $1/\epsilon$) number of calls to this function are necessary to approximate its minimum with a relative error of ϵ, say with a quasi-Newton method. Thus, the algorithm provides an accuracy of ϵ in a computation time which is $O(m^{\log d/\epsilon} q(\epsilon, m, d)) = O(d^{\log m/\epsilon} q(\epsilon, m, d))$, where q is a polynomial. This exponential behavior cannot be avoided because of Theorem 2.5 (unless $P = NP$), and our results show that for nonnegative matrices, the p-radius can be computed up to any required accuracy in an a priori computable amount of time, which to the best of our knowledge was not known before.

5. Applications and examples. We apply our technique to problems of convergence of subdivision schemes in the L_p-spaces, smoothness of refinable functions, and smoothness of compactly supported wavelets. We end this section by reporting computations on randomly generated nonnegative matrices.

5.1. Chaikin's subdivision scheme. An important application of the p-radius is that it expresses the rate of convergence of subdivision schemes and the exponent of regularity of refinable functions in L_p (see [7] for details and many references). For univariate functions the set of matrices consists of two block Toeplitz $N \times N$ matrices A_1, A_2 defined for a sequence $c_0, \dots, c_N$ as follows:

$$(5.1) \qquad (A_1)_{ij} = c_{2i-j-1}, \qquad (A_2)_{ij} = c_{2i-j}, \qquad i, j \in \{1, \dots, N\}.$$

We focus on the simplest case of Chaikin's subdivision scheme and the corresponding refinable functions, De Rham curves. In this case $N = 2$ and $c_0 = \omega$, $c_1 = 1 - 2\omega$, $c_2 = \omega$ for a real parameter $\omega \in (0, 1/2)$. Thus, we have

$$A_1 = \begin{pmatrix} \omega & 0 \\ \omega & 1 - 2\omega \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 - 2\omega & \omega \\ 0 & \omega \end{pmatrix}.$$
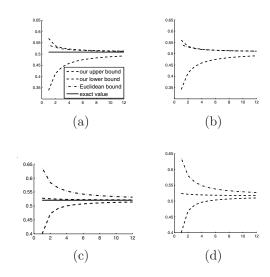
FIG. 5.1. *Evolution (with the length of the computed products) of the different bounds on ρ_p for the Chaikin interpolation with $\omega = 1/3$ ((a) and (b)) and $\omega = 1/5$ ((c) and (d)). For $p = 4$, ((a) and (c)), the exact value is available thanks to Proposition 2.1. For $p = 3.5$, ((b) and (d)) the exact value is unknown.*

The matrices A_1, A_2 are nonnegative, so we can find the values α_p, β_p effectively and squeeze the p-radius between an upper and a lower bound. We show here the results for $w = 1/3$ (where the set $\mathcal{M}$ is proportional to a set of binary matrices) and for $w = 1/5$. We compute bounds on $\mathcal{M}^k$ for larger and larger k. The variable on the x-axis is the length k of the products.

We choose $p = 4$ and $p = 3.5$. In the first case we are able to compute the exact value of the p-radius thanks to Proposition 2.1, so that one can see in Figure 5.1 the rate of convergence to the real value. We compare our results with the upper bound obtained by applying (1.1) with the Euclidean norm. One can see not only that the latter quantity provides only an upper bound on the p-radius, but, moreover, that it converges far more slowly than the bound obtained with the best conic norm in the case $\omega = 1/5$. In the case $\omega = 1/3$, both the best conic norm and the Euclidean norm seem to converge very quickly to the exact value.

For $p = 3.5$, $\omega = 1/3$ we obtain $0.49 \leq \rho_p \leq 0.512$. Thus, the Hölder exponent of the corresponding Chaikin–De Rham refinable function φ in the space L_p, which is equal to $1 - \log_2 \rho_p$ [18], is between 1.96 and 2.02. In the case $\omega = 1/5$ it is between 1.95 and 1.973.

5.2. Daubechies wavelets. The regularity of Daubechies wavelets in various functional spaces has been thoroughly studied in the literature. Our results make it possible to estimate their exponents of regularity in the spaces L_p for all $p \in [1, +\infty]$. It is known [23] that the L_p-regularity of the nth Daubechies wavelet function ψ_n is equal to $n - \log_2 \rho_p(\mathcal{M})$, where $\mathcal{M} = \{A_1, A_2\}$ and the matrices A_1, A_2 are constructed by formulas (5.1) for $N = n - 1$ with a special sequence $c_0, \ldots, c_{n-1}$ (see [9] for details). For $n = 1$ we obtain the Haar wavelets, for which everything is known, and for $n = 2$ all exponents of regularity are precisely computed for all p [23, sect. 12]. Therefore, we focus on the case $n \geq 3$. For every n the $(n-1) \times (n-1)$ matrices A_1, A_2 have positive and negative entries, which forbids the use of Theorems 4.1 and 4.2 in order to compute α_p and β_p exactly. However, one can apply Proposition 2.1 and compute
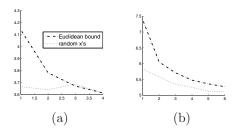
FIG. 5.2. *Evolution of the different bounds on ρ_p ($p = 8$) for the third* (a) *and fourth* (b) *Daubechies wavelets.*

instead the $\frac{p}{2}$-radius of $\mathcal{M}^{\otimes 2}$, which leaves the cone $K = \{x \otimes x : x \in \mathbb{R}^d\}$ invariant. One can then derive upper bounds on the p-radius by applying Theorems 3.5 and 3.9, by picking at random an $x \in \mathbb{R}^d$ to build a norm $r_{\tilde{x}}$, where $\tilde{x} = x \otimes x$. One can then iterate to build several norms and take the lowest upper bound obtained. This method appears to work well in practice.

We present in Figures 5.2(a) and 5.2(b) the result of approximation of ρ_8 for the third and fourth Daubechies wavelet functions ψ_3, ψ_4, respectively. One can see in (a) that our method can also not behave so well in some examples, because we do not have accuracy guarantees when matrices have negative entries. However, it seems that the bigger the dimension is, the better the results are, and for the fourth Daubechies wavelet our method already performs far better than the estimation with the Euclidean norm. We see that for ψ_3 we have $\rho_8 \leq 3.61$. Hence the derivative ψ_3' belongs to $L_8(\mathbb{R})$, and its exponent of regularity in this space is $2 - \log_2 \rho_8 \geq 0.14$. For ψ_4' we have $\rho_8 \leq 5.114$. Hence ψ_4' also belongs to $L_8(\mathbb{R})$, and its exponent of regularity in this space is $3 - \log_2 \rho_8 \geq 0.64$.

We then focus on the fifth Daubechies wavelet ψ_5, obtained with the following set of matrices (rounded up to the second decimal):

$$(5.2) \quad \left\{ \begin{pmatrix} 7.24 & 0 & 0 & 0 \\ 4.82 & -8.90 & 7.24 & 0 \\ 0.15 & -1.32 & 4.82 & -8.90 \\ 0 & 0 & 0.15 & -1.32 \end{pmatrix}, \begin{pmatrix} -8.90 & 7.24 & 0 & 0 \\ -1.32 & 4.82 & -8.90 & 7.24 \\ 0 & 0.15 & -1.32 & 4.82 \\ 0 & 0 & 0 & 0.15 \end{pmatrix} \right\}.$$

It is possible to show that for these matrices, $\rho_\infty = 8.174\ldots$ with existing techniques, such as those introduced in [26]. Since $5 - \log_2 \rho_\infty = 1.9689\ldots$, we see that ψ_5 belongs to $C^1(\mathbb{R})$ but not to $C^2(\mathbb{R})$, and the Hölder exponent of the derivative ψ_5' is $0.9689\ldots$, so ψ_5 is "almost" in C^2. On the other hand, $\rho_1 \geq \rho((A_1 + A_2)/2) = 4.08\ldots$, so for any $p \geq 1$ the p-radius satisfies $4.08 < \rho_p < 8.175$. In Figure 5.3 we present results of approximation of $\rho_p(\mathcal{M}_5)$ for $p = 4, 6, 6.5$, and 8. For $p = 4, 6$ the exact value is known from Proposition 2.1. Since $\rho_6 < 8$, it follows that $\psi_5'' \in L_6$. For larger even p, Proposition 2.1 is no longer applicable in practice because it involves computing the spectral radius of a matrix of dimension at least $4^8 \approx 100000$.

For $p = 6.5, 8$ the upper bound seems to converge to some value around 8, but we were not able to prove that the exact value is smaller than 8. In a separate computation with products of length 13, we obtained that for $p = 8$, the p-radius is smaller than 8.055. Thus, the exponent of regularity of ψ_5' in L_8 is at least $0.99\ldots$. It remains an open question whether ψ_5'' belongs to L_8 and what is the largest p for which $\psi_5'' \in L_p$, but we conjecture this value to be 8.
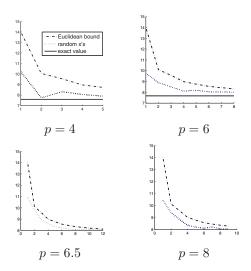
FIG. 5.3. *Evolution of the different bounds on ρ_p for the fifth Daubechies wavelet.*

5.3. The generalized four-point scheme. The generalized four-point scheme is one of the best studied interpolatory subdivision algorithms for approximation and curve design. Introduced in 1990 in [11], it has been analyzed and modified in many papers (see [13] and the references therein). The scheme depends on the so-called tension parameter $\omega \in \mathbb{R}$. For each value of ω it generates a *limit curve*, i.e., a continuous refinable function $\varphi_\omega(x)$. One of the main questions in this context is the smoothness of this function for various ω. It is known that the Hölder exponent of the derivative φ'_ω in the space L_p is equal to $-\log_2 \rho_p(A_0, A_1)$, where A_0, A_1 are two special 4×4 matrices with coefficients depending on ω. In 2009 it was shown [13] that $\varphi_\omega \in C^1(\mathbb{R})$ precisely when $0 < \omega < \omega^*$, where $\omega^* = 0.19273\ldots$ is the unique real solution of the cubic equation $32t^3 + 4t - 1 = 0$. Thus, $\varphi_\omega \notin C^1(\mathbb{R})$ for $\omega \geq \omega^*$. We computed upper bounds on $\rho_p(A_0, A_1)$ for various values of ω. Our calculations show that, for instance, for $\omega = 0.227$ the 3-radius of A_0, A_1 is smaller than 0.9978. Hence, for $\omega = 0.227$ the function φ_ω is differentiable, its derivative is in $L_3(\mathbb{R})$, and the Hölder exponent of φ'_ω in that space is bigger than $-\log_2 0.9978 > 0.003$. For $\omega = 0.22$, the 5-radius of A_0, A_1 is smaller than 0.9987. Therefore, for this omega we have $\varphi'_\omega \in L_5(\mathbb{R})$, and the Hölder exponent in L_5 is bigger than $-\log_2 0.9987 > 0.0018$. Let us observe that for this particular example, as well as for the butterfly scheme below, the Euclidean norm also provides good bounds. However, for a fixed length of the constructed products, we were always able to find a particular conic norm that outperforms the Euclidean norm. For instance, for $\omega = 0.22$, the Euclidean norm is not able to prove that $\rho_3 < 1$, as the bound obtained with this norm is 1.0021. This is not surprising, as we have shown that there are always very good norms among the conic norms (Theorem 3.5).

5.4. The butterfly scheme. The butterfly scheme, originated in [11], is an interpolatory two-variate subdivision scheme. It is a natural generalization of the univariate four-point subdivision algorithm. The scheme also depends on the tension parameter ω. It is known that only for $\omega = \frac{1}{16}$ the scheme reproduces all the polynomials of degree three, while for all other ω it reproduces only linear functions. Therefore, for $\omega = \frac{1}{16}$ the butterfly scheme has the best approximation properties.

For this value of ω the Hölder smoothness of the limiting surface $\varphi(x_1, x_2)$ is equal to $-\log_2 \rho_\infty(\mathcal{M})$, where $\rho_\infty(\mathcal{M})$ is the joint spectral radius of the family $\mathcal{M}$ of four special 17×17 matrices $A_1, \ldots, A_4$ (we follow here the construction of [11]). The usual spectral radius of three of those matrices is equal to $\frac{1}{4}$, therefore $\rho_\infty(\mathcal{M}) \geq \frac{1}{4}$. Hence, $\varphi \notin C^2(\mathbb{R}^2)$. On the other hand, there is a conjecture that actually $\rho_\infty = \frac{1}{4}$, and so $\varphi \in C^{2-\varepsilon}(\mathbb{R}^2)$ for any $\varepsilon > 0$. For this reason we focus here on finite p. Using our algorithm we estimate the 3-radius of $\mathcal{M}$. It appears that $\rho_3 \leq 0.2415$, and therefore the Hölder exponent of φ in the space $L_3(\mathbb{R}^2)$ is at least $-\log_2 0.2415 = 2.0499 \ldots$.

5.5. p-radius of random matrices. To end this section we report computations on random matrices with nonnegative entries. In this situation too, our methods perform well in comparison with other classical norms. However, we observe that the Euclidean norm also performs very well for random matrices, at least if they are not too sparse. The sparser the matrices become, the more the bound derived from the Euclidean norm gives a bad estimate for the p-radius, while the quantities α_p, β_p do not seem sensitive to sparsity.

As an example, we took $p = 3.5$, and we ran computations on pairs of sparse random 7×7 matrices with entries between 0 and 10 and probability $9/10$ to be zero. For products of length 10, the upper bound for $\rho_{3.5}$ obtained with the Euclidean norm gave 7.76, while the bound with our method gave 7.46, which we think is roughly the exact value. As a matter of comparison, the ∞-norm (maximum row sum) gives an upper bound of 7.97. Of course only our method allowed us to derive a lower bound, which was equal to 6.49. As always, the upper bound seems to have converged to the exact value much faster than expected theoretically. As we have observed on all the examples, the ∞-norm gives much worse results than the two other above-mentioned bounds.

6. Conclusion. In this paper we have proposed new methods for approximating the p-radius of a set of matrices. Their interest is both theoretical and practical. Indeed, to the best of our knowledge, our methods are the first that allow one to obtain an arbitrary accurate estimate for the p-radius in the case where p is not an integer and when the matrices have nonnegative entries. The previous methods, based on the definition (1.1), only allowed us to derive closer and closer upper bounds on the desired value, without knowing the quality of the approximation. Also, in the case where the matrices have negative entries we provide a way to obtain good upper bounds on the p-radius. However, in this latter case we leave open the question of knowing whether there is an algorithm which, given a set of matrices and a prescribed accuracy ϵ, returns an estimate of the p-radius which is within an accuracy ϵ of the exact value.

Our methods also make it particularly easy in practice to obtain a good upper bound: thanks to the several parameters in the approximation algorithms (one can raise the set to a certain Kronecker power $\mathcal{M}^{\otimes k}$, one can take the set of products of length k $\mathcal{M}^k$, one can try several norms $r_x(\cdot)$ in the case where the matrices have negative entries), it is possible to tune the methods to rapidly obtain good results by trial and error. In the last section we have shown the efficiency of our technique by computing the exponents of regularity in L_p spaces for some well-studied refinable functions, subdivision schemes, and the Daubechies wavelets.

REFERENCES

[1] A. A. AHMADI, R. M. JUNGERS, P. PARRILO, AND M. ROOZBEHANI, *Analysis of the joint spectral radius via Lyapunov functions on path-complete graphs*, in Proceedings of Hybrid Systems: Computation and Control (HSCC'11), Chicago, ACM, New York, 2011.

[2] V. M. ALEXEEV, E. M. GALEEV, AND V. M. TIKHOMIROV, *Book of Problems in Optimization*, Nauka, Moscow, 1984.

[3] N. BARABANOV, *On the Lyapunov exponent of discrete inclusions*. I–III, Autom. Remote Control, 49 (1988), pp. 152–157, 283–287, 558–565.

[4] V. D. BLONDEL AND Y. NESTEROV, *Computationally efficient approximations of the joint spectral radius*, SIAM J. Matrix Anal. Appl., 27 (2005), pp. 256–272.

[5] 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.

[6] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, New York, 2004.

[7] C. A. CABRELLI, C. HEIL, AND U. M. MOLTER, *Self-similarity and multiwavelets in higher dimensions*, Mem. Amer. Math. Soc., 170 (2004), no. 807.

[8] A. CICONE, N. GUGLIELMI, S. SERRA CAPIZZANO, AND M. ZENNARO, *Finiteness property of pairs of 2×2 sign-matrices via real extremal polytope norms*, Linear Algebra Appl., 432 (2009), pp. 796–816.

[9] I. DAUBECHIES, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61, SIAM, Philadelphia, 1992.

[10] 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.

[11] N. DYN, J. GREGORY, AND D. LEVIN, *A butterfly subdivision scheme for surface interpolation with tension control*, ACM Trans. Graph., 9 (1990), pp. 160–169.

[12] N. GUGLIELMI AND M. ZENNARO, *An algorithm for finding extremal polytope norms of matrix families*, Linear Algebra Appl., 428 (2008), pp. 2265–2282.

[13] J. HECHLER, B. MOSSNER, AND U. REIF, C^1-*Continuity of the generalized four-point scheme*, Linear Algebra Appl., 430 (2009), pp. 3019–3029.

[14] R. Q. JIA, *Subdivision schemes in L_p spaces*, Adv. Comput. Math., 3 (1995), pp. 309–341.

[15] R. M. JUNGERS, *The Joint Spectral Radius. Theory and Applications*, Lecture Notes in Control and Inform. Sci. 385, Springer-Verlag, Berlin, 2009.

[16] R. M. JUNGERS AND V. D. BLONDEL, *On the finiteness property for rational matrices*, Linear Algebra Appl., 428 (2008), pp. 2283–2295.

[17] R. M. JUNGERS, V. PROTASOV, AND V. D. BLONDEL, *Efficient algorithms for deciding the type of growth of products of integer matrices*, Linear Algebra Appl., 428 (2008), pp. 2296–2311.

[18] K.-S. LAU AND J. WANG, *Characterization of L^p-solutions for the two-scale dilation equations*, SIAM J. Math. Anal., 26 (1995), pp. 1018–1046.

[19] Y. NESTEROV AND A. NEMIROVSKII, *Interior Point Polynomial Algorithms in Convex Programming*, Stud. in Appl. Math. 13, SIAM, Philadelphia, 1994.

[20] P. OSWALD, *Optimality of multilevel preconditioning for nonconforming $P1$ finite elements*, Numer. Math., 111 (2008), pp. 267–291.

[21] P. PARRILO AND A. JADBABAIE, *Approximation of the joint spectral radius using sum of squares*, Linear Algebra Appl., 428 (2008), pp. 2385–2402.

[22] V. YU. PROTASOV, *A generalized spectral radius. A geometric approach*, Izv. Math., 61 (1997), pp. 995–1030.

[23] V. YU. PROTASOV, *Fractal curves and wavelets*, Izv. Math., 70 (2006), pp. 975–1013.

[24] V. YU. PROTASOV, *Self-similarity equations and the p-radius of operators*, Russian Math. Surveys, 62 (2007), pp. 1221–1223.

[25] V. YU. PROTASOV, *Extremal L_p-norms of linear operators and self-similar functions*, Linear Algebra Appl., 428 (2008), pp. 2339–2356.

[26] V. YU. PROTASOV, R. M. JUNGERS, AND V. D. BLONDEL, *Joint spectral characteristics of matrices: A conic programming approach*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2146–2162.

[27] J. RADON, *Über eine Erweiterung des Begriffs der konvexen Funktionen, mit einer Anwendung auf die Theorie der konvexen Körper*, S.-B. Akad. Wiss. Wien, 125 (1916), pp. 241–258.

[28] G. C. ROTA AND G. STRANG, *A note on the joint spectral radius*, Indag. Math., 22 (1960), pp. 379–381.

[29] I. SHEIPAK, *On the construction and some properties of self-similar functions in the spaces $L_p[0, 1]$*, Math. Notes, 81 (2007), pp. 827–839.

[30] 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.

[31] Y. WANG, *Two-scale dilation equations and the mean spectral radius*, Random Comput. Dynam., 4 (1996), pp. 49–72.

[32] D.-X. ZHOU, *The p-norm joint spectral radius for even integers*, Methods Appl. Anal., 5 (1998), pp. 39–54.

[33] D.-X. ZHOU, *The p-norm joint spectral radius and its applications in wavelet analysis*, Wavelet Analysis and Applications (Guangzhou, China, 1999), AMS/IP Stud. Adv. Math. 25, AMS, Providence, RI, 2002, pp. 305–326.