FAST APPROXIMATION OF THE P-RADIUS, MATRIX PRESSURE OR GENERALISED LYAPUNOV EXPONENT FOR POSITIVE AND DOMINATED MATRICES

IAN D. MORRIS

Abstract. If $A_1, \ldots, A_N$ are real $d \times d$ matrices then the *p-radius*, *generalised Lyapunov exponent* or *matrix pressure* is defined to be the asymptotic exponential growth rate of the sum $\sum_{i_1,\ldots,i_n=1}^{N} \|A_{i_n} \cdots A_{i_1}\|^p$, where p is a real parameter. Under its various names this quantity has been investigated for its applications to topics including wavelet regularity and refinement equations, fractal geometry and the large deviations theory of random matrix products. In this article we present a new algorithm for computing the p-radius under the hypothesis that the matrices are all positive (or more generally under the hypothesis that they satisfy a weaker condition called *domination*) and of very low dimension. This algorithm is based on interpreting the p-radius as the leading eigenvalue of a trace-class operator on a Hilbert space and estimating that eigenvalue via approximations to the Fredholm determinant of the operator. In this respect our method is closely related to the work of Z.-Q. Bai and M. Pollicott on computing the top Lyapunov exponent of a random matrix product. For pairs of positive matrices of dimension two our method yields substantial improvements over existing methods.

Key words. p-radius, generalized Lyapunov exponent, matrix pressure, joint spectral radius

AMS subject classifications. 00X00

1. Introduction. If $(A_1, \ldots, A_N)$ is a tuple of real $d \times d$ matrices and $p \in \mathbb{R}$ a real parameter, the limit

$$\varrho_p(A_1, \ldots, A_N) := \lim_{n\to\infty} \left(\sum_{i_1,\ldots,i_n=1}^{N} \|A_{i_n} \cdots A_{i_1}\|^p \right)^{\frac{1}{n}} \tag{1.1}$$

exists by applying Fekete's subadditivity lemma to the sequence

$$a_n(p) := \log \left(\sum_{i_1,\ldots,i_n=1}^{N} \|A_{i_n} \cdots A_{i_1}\|^p \right)$$

if $p \geq 0$, or to the sequence $-a_n(p)$ if $p < 0$. The quantity (1.1), modulo some trivial variations in its definition, has been studied independently in at least three different contexts and literatures: under the name of *generalised Lyapunov exponent* the quantity $\log(N^{-1}\varrho_p(A_1, \ldots, A_N))$ has been studied for $p \in \mathbb{R}$ in [10, 41] where its investigation is motivated by the large deviations theory of random matrix products in statistical mechanics; under the name of *matrix pressure*, the quantity $\varrho_p(A_1, \ldots, A_N)$ has been investigated for $p \geq 0$ in the fractal geometry literature in view of its applications to the dimension of self-similar and self-affine limit sets ([12, 13, 14, 27, 34]); and in the joint spectral radius literature, the quantity $N^{-1/p}\varrho_p(A_1, \ldots, A_N)^{1/p}$ has been investigated for $p \geq 1$ in connection with its applications to wavelet regularity [9, 26, 43] and the control theory of discrete linear inclusions [22, 30]. Across all three literatures there has arisen the problem of computing or estimating the quantity $\varrho_p(A_1, \ldots, A_N)$ – as may be seen for example in [23, 27, 31, 34, 36, 39, 41] – and it is with this that the present article is concerned. The principal result of this article is a new algorithm for the computation of $\varrho_p(A_1, \ldots, A_N)$ in the case where the matrices $A_1, \ldots, A_N$ are positive and p is an arbitrary real number. More generally, our method extends to the case where the matrices $A_1, \ldots, A_N$ strictly preserve a cone or *multicone*.

2. Statement of main result. In order to state our result let us present the definition of a multicone. Let us say that a *cone* in $\mathbb{R}^d$ is a set $\mathcal{K} \subset \mathbb{R}^d$ which is closed, convex, has nonempty interior, satisfies $\lambda\mathcal{K} = \lambda\mathcal{K}$ for all real $\lambda > 0$ and satisfies $\mathcal{K} \cap -\mathcal{K} = \{0\}$. A *multicone* will be a tuple $(\mathcal{K}_1, \ldots, \mathcal{K}_m)$ of cones in $\mathbb{R}^d$ such that for some nonzero vector $w \in \mathbb{R}^d$ we have $\langle u, w\rangle > 0$ for all nonzero $u \in \bigcup_{j=1}^m \mathcal{K}_j$, and such that $\mathcal{K}_i \cap \mathcal{K}_j = \{0\}$ for distinct $i, j \in \{1, \ldots, m\}$. We call such a vector w a *transverse-defining vector* for the multicone, and when discussing specific multicones we will find it convenient to fix a choice of transverse-defining vector w. We say that a matrix $A \in M_d(\mathbb{R})$ strictly preserves a cone $\mathcal{K}$ if $A(\mathcal{K} \setminus \{0\}) \subseteq \operatorname{Int}\mathcal{K}$, and we say that A strictly preserves a multicone $(\mathcal{K}_1, \ldots, \mathcal{K}_m)$ if for every $i = 1, \ldots, m$ we have $A(\mathcal{K}_i \setminus \{0\}) \subseteq \operatorname{Int}\mathcal{K}_j \cup (-\operatorname{Int}\mathcal{K}_j)$ for some $j \in \{1, \ldots, m\}$ depending on i. If A strictly preserves a multicone then a simple pigeonhole argument demonstrates that some power of A strictly preserves a cone, which implies that A has a simple leading eigenvalue (which might be either positive or negative). We say that $(A_1, \ldots, A_N) \in M_d(\mathbb{R})^N$ strictly preserves a multicone $(\mathcal{K}_1, \ldots, \mathcal{K}_m)$ if every A_i strictly preserves that multicone. We say that $(A_1, \ldots, A_N)$ is *multipositive* if there exists a multicone which is strictly preserved by $(A_1, \ldots, A_N)$. The property of multipositivity admits characterisations which do not overtly refer to cones or multicones: for example, if $(A_1, \ldots, A_N) \in M_d(\mathbb{R})^N$ is a tuple of invertible matrices then the multipositivity of $(A_1, \ldots, A_N)$ is equivalent to the existence of constants $C > 0$, $\theta \in (0, 1)$ such that

$$\max\left\{\frac{\sigma_2(A_{i_n}\cdots A_{i_1})}{\sigma_1(A_{i_n}\cdots A_{i_1})} \colon 1 \leq i_1, \ldots, i_n \leq N\right\} \leq C\theta^n \tag{2.1}$$

for all $n \geq 1$, where $\sigma_k(A)$ denotes the k^{th} singular value of the matrix A, see for example [4, 6, 29]. The condition (2.1) above is sometimes called 1-*domination* or simply *domination* and has been explored in some detail in the dynamical systems literature [1, 6]; its applications to certain numerical invariants of sets of matrices have been investigated in such works as [7, 8]. We remark that the existence of a strictly preserved multicone, and the condition (2.1), are both preserved by a linear change of co-ordinates: if every A_i is conjugated by some matrix X then the constant θ in (2.1) is unchanged and the constant C is multiplied by the square of the condition number of X.

For each $N \geq 1$ we let Σ_N^* denote the set of all finite sequences $\mathtt{i} = (i_1, \ldots, i_n)$ such that $i_1, \ldots, i_n$ are integers between 1 and N. If a tuple of matrices $(A_1, \ldots, A_N) \in M_d(\mathbb{R})^N$ is understood, given $\mathtt{i} = (i_1, \ldots, i_n) \in \Sigma_N^*$ we define $A_\mathtt{i} := A_{i_n} \cdots A_{i_1}$. If $\mathtt{i} = (i_1, \ldots, i_n) \in \Sigma_N^*$ then we write $|\mathtt{i}| := n$ and call this the length of $\mathtt{i}$. Finally we let $\rho(A)$ denote the spectral radius of the matrix A, and we let $\lambda_1(A), \ldots, \lambda_d(A)$ denote the eigenvalues of A listed in decreasing order of absolute value. Since our matrices A will always strictly preserve a multicone the largest eigenvalue of A will always be unique and the definition of $\lambda_1(A)$ unambiguous.

We may now state the principal result of this article, which is the following:

THEOREM 2.1. *Let $(A_1, \ldots, A_N) \in M_d(\mathbb{R})^N$ be multipositive, where $N, d \geq 2$, and let $p \in \mathbb{R}$. For every $n \geq 1$ define*

$$t_n := \sum_{|\mathtt{i}|=n} \rho(A_\mathtt{i})^p \prod_{j=2}^{d} \left(1 - \frac{\lambda_j(A_\mathtt{i})}{\lambda_1(A_\mathtt{i})}\right)^{-1} = \sum_{|\mathtt{i}|=n} \frac{\lambda_1(A_\mathtt{i})^{d-1}\rho(A_\mathtt{i})^p}{p'_{A_\mathtt{i}}(\lambda_1(A_\mathtt{i}))}$$

where $p_B(x) := \det(xI - B)$ denotes the characteristic polynomial of the matrix B

and $p'_B(x_0)$ *its first derivative evaluated at* x_0*. Define* $a_0 := 1$ *and*

$$a_n := \frac{(-1)^n}{n!} \det \begin{pmatrix} t_1 & n-1 & 0 & \cdots & 0 & 0 \\ t_2 & t_1 & n-2 & \cdots & 0 & 0 \\ t_3 & t_2 & t_1 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & \cdots & t_1 & 1 \\ t_n & t_{n-1} & t_{n-2} & \cdots & t_2 & t_1 \end{pmatrix}$$

$$= \sum_{k=1}^{n} \frac{(-1)^k}{k!} \sum_{\substack{n_1,\ldots,n_k \geq 1 \\ n_1+\cdots+n_k=n}} \prod_{\ell=1}^{k} \frac{t_{n_\ell}}{n_\ell}$$

for every $n \geq 1$*. Then for all sufficiently large* n *there exists a smallest positive real root* $r_n > 0$ *of the polynomial* $\sum_{k=0}^{n} a_k x^k$*, and there exist constants* $K, \gamma > 0$ *such that for all large enough integers* n

$$\left| \varrho_p(A_1, \ldots, A_N) - \frac{1}{r_n} \right| \leq K \exp\left(-\gamma n^{\frac{d}{d-1}} \right). \tag{2.2}$$

The method of Theorem 2.1 is inspired by earlier results by M. Pollicott and coauthors on the estimation of various dynamical quantities via transfer operators such as in the articles [32, 33, 35]. It relies on a transfer operator theorem proved by the author in [29] which extends a technical result of the article [32] (which is itself in turn a rigorous version of an earlier argument of Z.-Q. Bai [2]).

Theorem 2.1 applies in particular if the matrices A_i are all positive matrices, or if the matrices A_i all strictly preserve a single cone $\mathcal{K}$. However, multipositive matrix tuples with neither of these two properties also exist: see [1]. We remark that since $\varrho_p(A_1, \ldots, A_N) = \varrho_p(X^{-1}A_1X, \ldots, X^{-1}A_NX)$ for every invertible matrix X, a sufficient condition for the application of Theorem 2.1 is that the matrices A_i be *simultaneously conjugate* to positive matrices. The problem of estimating the implied constants K and γ in (2.2) is not attempted in this article, but is discussed briefly in §7 below. However, we believe this problem to be difficult, and at the present time this limits Theorem 2.1 to the production of empirical estimates of $\varrho_p(A_1, \ldots, A_N)$ rather than estimates whose accuracy can be rigorously guaranteed. In general the constants K and γ should be expected to diverge to infinity and zero respectively as the positivity hypothesis is weakened, that is, when the ratio of the largest entry of the matrices A_i to the smallest entry tends to infinity. In §7 we also briefly discuss the complexity of computing the quantities a_n, which we estimate to be of order roughly d^4N^n.

We believe that it would be possible in principle to adapt Theorem 2.1 so as to work with restricted products of matrices such that the products $A_{i_1} \cdots A_{i_n}$ are subject to a Markovian constraint, i.e. such that certain pairs of successive values $(i_k, i_{k+1}) \in \{1, \ldots, N\}^2$ are disallowed. This would yield an analoguous approximation formula for the exponential growth rate ϱ_p of the expressions $\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p$ for which the sum is restricted to words allowed by the Markovian constraint. The definition of t_n in this case must be correspondingly modified so as to sum only over words $\mathtt{i}$ of length n with the property that $\mathtt{ii}$ is an allowed word, similarly to formulas for the approximation of the top Lyapunov exponent given in the thesis [42]. The multipositivity hypothesis in this situation may potentially be relaxable to something resembling that of [5].

The reader will notice that the order of convergence in Theorem 2.1 is strongest when the dimension of the matrix is 2 and becomes weaker as the dimension is increased, although it is in all cases super-exponential in n. On the other hand as noted above the number of computational steps involved in computing the coefficients a_n admits a factor N^n. The method introduced in this article is therefore most useful when both the dimension d and the number of matrices N are equal to 2 and it should not be expected to compare favourably with other methods when N exceeds 4 or when d exceeds 3. We also emphasise that the positivity (or domination) requirement makes it unsuitable for use with sparse matrices, unlike the alternative methods which are reviewed in §3 below.

In the situation where both the number of matrices and the dimension are equal to 2, Theorem 2.1 substantially outperforms other methods. For example, in the previous work [23], R. Jungers and V. Yu. Protasov investigated the problem of computing what in our notation corresponds to the quantity

$$1 - \log_2 \left(2^{-\frac{1}{p}} \varrho_p(A_1, A_2)^{\frac{1}{p}} \right) = \frac{p+1}{p} - \frac{1}{p} \log_2 \varrho_p(A_1, A_2)$$

for the pair of matrices

$$A_1 := \begin{pmatrix} \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{3}{5} \end{pmatrix}, \qquad A_2 := \begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ 0 & \frac{1}{5} \end{pmatrix}$$

with $p := 3.5$, obtaining an estimate of $1.95 \leq \frac{p+1}{p} - \frac{1}{p} \log_2 \varrho_p(A_1, A_2) \leq 1.973$. Since the pair (A_1, A_2) is simultaneously conjugate to a pair of positive matrices Theorem 2.1 may be applied. Taking $n = 20$ yields the estimate

$$1.95382129317932586675038991473149255113828006412699 7\ldots \tag{2.3}$$

for the same quantity, which we have found to be empirically accurate to all significant figures shown in the sense that these digits remain stable as n is further increased.

We remark that in the literature on the generalised Lyapunov exponent, it is common to consider the quantity

$$\lim_{n\to\infty} \left(\sum_{i_1,\ldots,i_n=1}^{N} p_{i_n} \cdots p_{i_1} \|A_{i_n} \cdots A_{i_1}\|^p \right)^{\frac{1}{n}} \tag{2.4}$$

in place of the quantity $\varrho_p(A_1, \ldots, A_N)$ as defined in (1.1), where $(p_1, \ldots, p_N)$ is a probability vector. The quantity (2.4) can easily be included within the scope of (1.1) and Theorem 2.1 by replacing each instance of a matrix A_i with the corresponding matrix $p_i^{1/p} A_i$. For the remainder of the article we therefore ignore the issue of giving a probability weighting to each A_i and concentrate on the calculation of the p-radius as defined in (1.1).

It is possible to show that the quantities t_n defined in Theorem 2.1 satisfy $\lim_{n\to\infty} t_n \varrho_p(A_1, \ldots, A_N)^{-n} = 1$ and therefore increase (or decrease) exponentially with n. The efficiency of the estimate in Theorem 2.1 on the other hand relies on the quantities a_n decreasing as $O(\exp(-\gamma n^{\frac{d}{d-1}}))$. The small size of the quantities a_n thus arises from additive cancellation among the relatively large terms in the sum defining each a_n. In practical applications it is therefore important to compute the quantities t_n to a precision exceeding that desired for the approximation to $\varrho_p(A_1, \ldots, A_N)$.

The remainder of this article is structured as follows. In §3 below we review the fundamental properties of ϱ_p and describe some existing techniques for its estimation. In §4 we describe in outline the techniques underlying the proof of Theorem 2.1 and in §5 the proof itself is presented. In §6 we present some examples of the computation of ϱ_p using the algorithms described herein.

3. Methods for estimating the p-radius.

3.1. Fundamental estimates. If $(A_1, \ldots, A_N) \in M_d(\mathbb{R})^N$ and $p \in \mathbb{R}$ then by elementary estimates it follows that $\varrho_p(A_1, \ldots, A_N) = 0$ if and only if the joint spectral radius

$$\lim_{n\to\infty} \max_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^{\frac{1}{n}}$$

is zero. It is well known that the joint spectral radius is zero if and only if all of the products $A_{i_d} \cdots A_{i_1}$ of length d are zero, if and only if there exists a basis in which all of the matrices $A_1, \ldots, A_N$ are simultaneously upper triangular with all diagonal entries equal to zero (for details see [21, §2.3.1]). Since the theory of the p-radius is trivial in this situation we will for the remainder of this paper deal only with matrices for which the p-radius is assumed to be nonzero. We remark that in the multipositive case considered in Theorem 2.1 every product $A_{\mathtt{i}}$ has a simple leading eigenvalue and in particular is not the zero matrix, so in this case $\varrho_p(A_1, \ldots, A_N)$ is guaranteed to be nonzero.

When $p > 0$ the p-radius admits an elementary description as the limit of a convergent sequence of upper bounds,

$$\varrho_p(A_1, \ldots, A_N) = \lim_{n\to\infty} \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p \right)^{\frac{1}{n}} = \inf_{n\geq 1} \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p \right)^{\frac{1}{n}}, \tag{3.1}$$

as a consequence of the submultiplicativity relation

$$\sum_{|\mathtt{i}|=m+n} \|A_{\mathtt{i}}\|^p \leq \left(\sum_{|\mathtt{i}|=m} \|A_{\mathtt{i}}\|^p \right) \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p \right).$$

Less trivially, when $p > 0$ it may also be expressed as the limit of a convergent sequence of lower bounds:

$$\begin{aligned} \varrho_p(A_1, \ldots, A_N) &= \lim_{n\to\infty} \left(\frac{\sum_{|\mathtt{i}|=nd} \|A_{\mathtt{i}}\|^p}{K(p,d) \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p \right)^{d-1}} \right)^{\frac{1}{n}} \\ &= \sup_{n\geq 1} \left(\frac{\sum_{|\mathtt{i}|=nd} \|A_{\mathtt{i}}\|^p}{K(p,d) \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p \right)^{d-1}} \right)^{\frac{1}{n}} \end{aligned} \tag{3.2}$$

where $K(p,d) := d^{2+(d+1)p} \max\{d^{1-p}, 1\}$, see [27, Theorem 1.2]. In particular the p-radius can in principle be approximated to within any prescribed error ε by systematically computing the upper and lower bounds until they eventually agree to within the prescribed amount. However, since the computational effort involved increases exponentially with n and the relative error may reasonably be presumed to be

at least of the order of $K(p,d)^{1/n}$, and since the constant $K(p,d)$ is relatively large even in the case $d=2$, this procedure seems unlikely to have any value for practical computations. An illustration of this is presented in §6 below. We remark that an additional theoretical consequence of the above expressions is that the p-radius varies continuously both in p and in the matrix entries when p is positive, since it is then equal to both an upper and a lower pointwise limit of sequences of continuous functions, hence continuous. When $p<0$ the computability and continuity of the p-radius do not seem to have been as thoroughly investigated, but based on the related works [7, 28, 40] it seems likely that continuity should not hold and that systematic upper and lower estimation might be infeasible, at least when the matrices are not assumed to be positive or invertible.

When p is a positive even integer, or when p is a positive integer and the matrices $A_1,\ldots,A_N$ preserve a cone, the identity

$$\varrho_p(A_1,\ldots,A_N)=\rho\left(\sum_{i=1}^N A_i^{\otimes p}\right) \tag{3.3}$$

has been discovered independently on several occasions [11, 36, 44]. (Here $A^{\otimes p}$ denotes the p^{th} Kronecker power of the matrix A, see for example [18, §4.2].) When p is a positive integer and $A_1,\ldots,A_N$ are not necessarily positive, the inequality

$$\varrho_p(A_1,\ldots,A_N)\leq\rho\left(\sum_{i=1}^N A_i^{\otimes p}\right)$$

may be obtained by the same means. Whilst in principle (3.3) represents an easy method for computing the p-radius of positive matrices, the size of the auxiliary matrix $\sum_{i=1}^N A_i^{\otimes p}$ increases exponentially with p which prevents the use of the formula when p is sufficiently large. For non-integer p these results may nonetheless be exploited so as to yield upper bounds as follows. We observe that if p_1 and p_2 are real numbers such that $0<p_1<p_2$, and $\lambda\in(0,1)$, then for each $n\geq 1$

$$\sum_{|\mathtt{i}|=n}\|A_{\mathtt{i}}\|^{\lambda p_1+(1-\lambda)p_2}\leq\left(\sum_{|\mathtt{i}|=n}\|A_{\mathtt{i}}\|^{p_1}\right)^{\lambda}\left(\sum_{|\mathtt{i}|=n}\|A_{\mathtt{i}}\|^{p_2}\right)^{1-\lambda}$$

using Hölder's inequality with $p:=\frac{1}{\lambda}$ and $q:=\frac{1}{1-\lambda}$. It follows easily that

$$\log\varrho_{\lambda p_1+(1-\lambda)p_2}(A_1,\ldots,A_N)\leq\lambda\log\varrho_{p_1}(A_1,\ldots,A_N)+(1-\lambda)\log\varrho_{p_2}(A_1,\ldots,A_N)$$

and hence the function $p\mapsto\log\varrho_p(A_1,\ldots,A_N)$ is convex. This yields the upper bound

$$\begin{aligned}\varrho_p(A_1,\ldots,A_N)&\leq\varrho_{\lfloor p\rfloor}(A_1,\ldots,A_N)^{p-\lfloor p\rfloor}\varrho_{1+\lfloor p\rfloor}(A_1,\ldots,A_N)^{1+\lfloor p\rfloor-p}\\&\leq\rho\left(\sum_{i=1}^N A_i^{\otimes\lfloor p\rfloor}\right)^{p-\lfloor p\rfloor}\rho\left(\sum_{i=1}^N A_i^{\otimes(1+\lfloor p\rfloor)}\right)^{1+\lfloor p\rfloor-p}\end{aligned} \tag{3.4}$$

valid for all $p>0$ and $A_1,\ldots,A_N\in M_d(\mathbb{R})$, which does not seem to have been previously noted in the literature. We will see in §6 below that despite its crudity this estimate does not automatically provide a bad approximation and should not be discounted out of hand.

3.2. Resampled Monte Carlo methods. In [41], J. Vanneste introduced a method based on the interpretation of the p-radius as an asymptotic moment of a random matrix product: given $A_1, \ldots, A_N \in M_d(\mathbb{R})$, $n \geq 1$ and $p \in \mathbb{R}$ we may view the sum $\frac{1}{N^n} \sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p$ as the expectation of the random variable $\mathtt{i} \mapsto \|A_{\mathtt{i}}\|^p$ where each word $\mathtt{i}$ of length n is chosen with probability $1/N^n$. This suggests the possibility of approximating $\frac{1}{N^n} \sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p$ for large n by Monte Carlo estimation: if we choose M words $\mathtt{i}_1, \ldots, \mathtt{i}_M$ independently then by the law of large numbers, the average $\frac{1}{M} \sum_{k=1}^{M} \|A_{\mathtt{i}_k}\|^p$ should for large enough M give a reasonable approximation to the value $\frac{1}{N^n} \sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p$ which is that random variable's expectation and hence a good approximation to $\varrho_p(A_1, \ldots, A_N)$ as long as n is reasonably large. However, except which p is small, the variance of this random variable will be prohibitively large – indeed exponentially large in n – which makes convergence in the strong law of large numbers unreasonably slow. To compensate for this Vanneste introduced a "go-with-the-winners" resampling scheme along the lines of [16], which successively modifies the distribution of the random variable $\mathtt{i} \mapsto \|A_{\mathtt{i}}\|^p$ so as to retain the same mean while reducing the variance; see discussion in [41, §III] for details. The particular strength of this method is that it has very limited dependence on the number of matrices and their dimension; on the other hand, the accuracy of the results is relatively low in practice. See §6 below for further discussion.

3.3. The convex optimisation bounds of Jungers and Protasov. The article [23] introduced new systematic upper and lower bounds for the p-radius in the case $p \geq 1$. If $(A_1, \ldots, A_N)$ are non-negative matrices, Jungers and Protasov showed that the quantities

$$\mathsf{a}_p(n) = \inf_{(u_1,\ldots,u_d)\in\mathbb{R}^d} \sum_{|\mathtt{i}|=n} \left(\max_{1\leq i\leq d} \sum_{j=1}^{d} (A_{\mathtt{i}})_{ij} e^{u_j - u_i} \right)^p,$$

$$\mathsf{b}_p(n) = \inf_{(v_1,\ldots,v_d)\in\mathbb{R}^d} \max_{1\leq j\leq d} \sum_{|\mathtt{i}|=n} \left(\sum_{i=1}^{d} (A_{\mathtt{i}})_{ij} e^{v_i - v_j} \right)^p,$$

where $(B)_{ij}$ denotes the (i,j) entry of the matrix $B \in M_d(\mathbb{R})$, satisfy

$$\max\left\{ d^{-\frac{p}{n}} \mathsf{a}_p(n)^{\frac{1}{n}}, d^{\frac{1-p}{n}} \mathsf{b}_p(n)^{\frac{1}{n}} \right\} \leq \varrho_p(A_1, \ldots, A_N) \leq \mathsf{b}_p(n)^{\frac{1}{n}}$$

for every $n \geq 1$. (Here we have modified the statement of their results in concordance with our definition of ϱ_p.) The quantities $\mathsf{a}_p(n)$ and $\mathsf{b}_p(n)$ are solutions to convex optimisation problems and as such may be efficiently approximated. In the case where $(A_1, \ldots, A_N)$ preserves a more general cone $\mathcal{K}$ (in the weak sense that $A_i\mathcal{K} \subseteq \mathcal{K}$ for each $i = 1, \ldots, N$) analogous upper and lower bounds are given, but these are not in general the solutions to convex optimisation problems and as such are more difficult to efficiently or rigorously estimate. Since the matrices $A_1^{\otimes 2}, \ldots, A_N^{\otimes 2}$ always preserve a cone irrespective of the structure of the original matrices $A_1, \ldots, A_N$, and since $\varrho_p(A_1, \ldots, A_N) = \varrho_{p/2}(A_1^{\otimes 2}, \ldots, A_N^{\otimes 2})$ for all $p \in \mathbb{R}$, this more general version of their method permits the estimation of $\varrho_p(A_1, \ldots, A_N)$ for arbitrary $A_1, \ldots, A_N \in M_d(\mathbb{R})$ and $p \geq 2$.

As with the upper and lower bounds (3.1) and (3.2) this system of estimation requires the computation of N^n matrix products in order to obtain the n^{th} approximation and as such is best suited to cases in which N is small.

3.4. Eigenvalue methods. As has been previously observed by J. Vanneste [41, §II.B], the quantity $\varrho_p(A_1,\ldots,A_N)$ can be represented as the leading eigenvalue of a linear operator on an infinite-dimensional function space in the following manner. Suppose that $A_1,\ldots,A_N \in M_d(\mathbb{R})$ are invertible matrices and let $p \in \mathbb{R}$. Let $\mathbb{RP}^{d-1}$ denote the space of lines through the origin in $\mathbb{R}^d$, with the distance between two lines defined to be the angle at which they intersect. For each nonzero $u \in \mathbb{R}^d$ let $\overline{u} \in \mathbb{RP}^{d-1}$ denote the line spanned by u. Define an operator on the space $C^\alpha(\mathbb{RP}^{d-1})$ of α-Hölder continuous functions $f\colon \mathbb{RP}^{d-1} \to \mathbb{R}$ by

$$(\mathcal{L}_p f)(\overline{u}) := \sum_{i=1}^N \left(\frac{\|A_i u\|}{\|u\|}\right)^p f(\overline{A_i u})$$

and observe that by a simple calculation

$$\left(\mathcal{L}_p^n f\right)(\overline{u}) = \sum_{|\mathtt{i}|=n} \left(\frac{\|A_{\mathtt{i}} u\|}{\|u\|}\right)^p f(\overline{A_{\mathtt{i}} u})$$

for every $n \geq 1$, $f \in C^\alpha(\mathbb{RP}^{d-1})$ and $\overline{u} \in \mathbb{RP}^{d-1}$. With only a little more work one may show that in fact

$$\lim_{n\to\infty} \left\|\mathcal{L}_p^n\right\|^{\frac{1}{n}} = \lim_{n\to\infty} \left(\sum_{|\mathtt{i}|=n} \|A_{\mathtt{i}}\|^p\right)^{\frac{1}{n}},$$

and under mild algebraic non-degeneracy conditions on the matrices A_i, a rather longer argument shows that $\varrho_p(A_1,\ldots,A_N)$ is the largest eigenvalue of $\mathcal{L}_p$ acting on $C^\alpha(\mathbb{RP}^{d-1})$ if $\alpha > 0$ is chosen sufficiently small (see for example [17, Théorème 8.8]). This suggests the idea of calculating $\varrho_p(A_1,\ldots,A_N)$ by approximating the operator $\mathcal{L}_p$ with a large matrix representing the action of the matrices A_i on a discretised version of $\mathbb{RP}^{d-1}$. This approach was previously described in [41, §IV.A] but does not seem to have been investigated in detail. A version of this method was also suggested in [29, §8] for the purpose of estimating the Hausdorff dimensions of some self-affine limit sets.

To give a concrete example, in the case $d = 2$ write $u(\theta) := (\cos\theta, \sin\theta)$ for each $\theta \in [0,\pi)$ and for $\overline{u}, \overline{v}$ let $[\overline{u}, \overline{v})$ denote the shorter of the two arcs in $\mathbb{RP}^1$ from $\overline{u}$ to $\overline{v}$, including the former endpoint but not the latter. Fix an integer $n \geq 1$. For each $i = 1,\ldots,N$ define an $n \times n$ matrix $B_i = [b_{jk}^{(i)}]_{j,k=0}^{n-1}$ by $b_{jk}^{(i)} := \|A_i u(j\pi/n)\|^p$ if $\overline{Au(j\pi/n)} \in [\overline{u(k\pi/n)}, \overline{u((k+1)\pi/n)})$ and $b_{jk}^{(i)} := 0$ otherwise. Define now the matrix $B := \sum_{i=1}^N B_i$. Since B corresponds to a version of $\mathcal{L}_p$ acting on functions defined on a discretisation of $\mathbb{RP}^1$ into n evenly-spaced points, we expect that for large n the spectral radius of B should give a reasonable approximation to $\rho(\mathcal{L}_p) = \varrho_p(A_1,\ldots,A_N)$. In principle it may be possible to demonstrate this rigorously using the methods of [25], but this does not seem to have so far been attempted in the literature and is certainly a problem beyond the scope of this article.

For two-dimensional matrices this method appears to yield approximations accurate to several decimal places in a tolerable amount of time (see §6 below) and it is apparent from the definition that the effect of increasing the number of matrices N has at worst a polynomial effect on the running time of the algorithm. However the size of the matrix required in order to discretise $\mathbb{RP}^{d-1}$ into a mesh of prescribed size

ε rises exponentially with the dimension d, suggesting that this method is unlikely to be very useful for matrices which are not of low dimension. The question also arises of whether better estimates may be obtained by adapting the mesh locally so as to include more mesh points in regions where the derivative of one of the maps $\overline{u} \mapsto \overline{A_i u}$ is large and fewer mesh points where it is small. Since the principal purpose of this article is to introduce the new algorithm given by Theorem 2.1, we leave these questions to other investigators.

4. Overview of the proof of Theorem 2.1. In the previous subsection we observed that $\varrho_p(A_1, \ldots, A_N)$ admits an interpretation as the leading eigenvalue of a linear operator on an infinite-dimensional function space and considered the possibility of approximating such an operator directly by operators on finite-dimensional spaces. This is however not the only mechanism by which the leading eigenvalue of an operator may be calculated. In order to describe our chosen alternative we will briefly and informally review some concepts from the theory of trace-class linear operators; thorough formal treatments of this topic may be found in e.g. [15, 38].

If an operator $\mathscr{L}$ on an infinite-dimensional Hilbert space has the property that the sequence of approximation numbers

$$\mathfrak{s}_n(\mathscr{L}) := \inf\{\|\mathscr{L} - \mathscr{F}\| \colon \operatorname{rank} \mathscr{F} < n\}$$

is summable then it is called *trace-class*. If this is the case then $\mathscr{L}$ is a compact operator (since it is a limit in the norm topology of a sequence of finite-rank operators) and therefore its spectrum consists of 0 together with a finite or infinite set of eigenvalues, each of finite algebraic multiplicity, which has no nonzero accumulation points. It is not difficult to see that $\mathfrak{s}_n(\mathscr{L}^k) \leq \|\mathscr{L}^{k-1}\|\mathfrak{s}_n(\mathscr{L})$ for every $k, n \geq 1$ by direct manipulation of the definition and consequently every power of a trace-class operator is also trace-class. If $\mathscr{L}$ is a trace-class operator on $\mathscr{H}$ with finite or infinite sequence of nonzero eigenvalues $(\lambda_n)_{n=1}^M$, it is classical that the series $\sum_{n=1}^M \lambda_n$ converges absolutely to a quantity which is called the *trace* of $\mathscr{L}$ and denoted $\operatorname{tr} \mathscr{L}$. Moreover the quantity

$$\det(I - z\mathscr{L}) := \prod_{n=1}^{M} (1 - z\lambda_k),$$

called the *Fredholm determinant* of $\mathscr{L}$, defines an entire holomorphic function in the variable z with power series $\sum_{n=0}^{\infty} a_n z^n$, say. It is also classical that in this case the zeros of $z \mapsto \det(I - z\mathscr{L})$ are precisely the reciprocals of the nonzero eigenvalues of $\mathscr{L}$ and that additionally

$$a_n = \sum_{i_1 < i_2 < \cdots < i_n} \lambda_{i_1} \cdots \lambda_{i_n} = \frac{1}{n!} \sum_{n_1 + \cdots + n_k = n} \prod_{i=1}^{k} \left(-\frac{\operatorname{tr} \mathscr{L}^{n_i}}{n_i} \right) \tag{4.1}$$

for every $n \geq 1$, where $a_0 := 1$ and where λ_k is interpreted as zero if $k > M$. It follows that if the traces $\operatorname{tr} \mathscr{L}^k$ can be easily calculated for $k = 1, \ldots, n$, say, then an approximation $\sum_{k=0}^{n} a_k z^k$ to the Fredholm determinant can be constructed using (4.1) and it might be hoped that the smallest positive real root of the polynomial $\sum_{k=0}^{n} a_k z^k$ would provide a good estimate for the reciprocal of the leading eigenvalue $\rho(\mathscr{L})$ of $\mathscr{L}$ as long as the remainder $\sum_{k=n+1}^{\infty} a_k z^k$ is extremely small. In view of the equation (4.1) it follows that if the sequence $(\lambda_n)_{n=1}^M$ can be shown to decay stretched-exponentially then this remainder will in fact be super-exponentially small, and this is indeed the approach which we will take in estimating $\varrho_p(A_1, \ldots, A_N)$.

This general approach to estimating dynamical quantities via operator eigenvalues has been previously exploited in a number of prior articles of which we note [2, 19, 20, 29, 32, 33, 35].

The proof of Theorem 2.1 therefore proceeds via the introduction of a trace-class operator $\mathscr{L}$ on a Hilbert space $\mathscr{H}$ with the properties required by the argument sketched above: a stretched-exponential estimate on the singular values $\mathfrak{s}_n(\mathscr{L})$ (which implies a stretched-exponential estimate on the eigenvalues via Weyl's inequality), the property $\rho(\mathscr{L}) = \varrho_p(A_1, \ldots, A_N)$, and a simple, computationally-feasible formula for the sequence of traces $\operatorname{tr} \mathscr{L}^n$. The following result from [29] saves us the necessity of constructing such an operator from first principles:

THEOREM 4.1. *Let $d, N \geq 2$, let $A_1, \ldots, A_N$ be real $d \times d$ matrices and suppose that $(\mathcal{K}_1, \ldots, \mathcal{K}_m)$ is a multicone for $(A_1, \ldots, A_N)$ with transverse-defining vector $w \in \mathbb{R}^d$. Then there exists a nonempty bounded open subset Ω of the complex hyperplane $\{z \in \mathbb{C}^d \colon \langle z, w\rangle = 1\}$ such that the following properties hold. Let $\mathcal{A}^2(\Omega)$ denote the separable complex Hilbert space of holomorphic functions $\Omega \to \mathbb{C}$ for which the integral $\int_\Omega |f(z)|^2 dV(z)$ is finite, where V denotes $2(d-1)$-dimensional Lebesgue measure on Ω. For each $p \in \mathbb{C}$ define an operator $\mathscr{L}_p \colon \mathcal{A}^2(\Omega) \to \mathcal{A}^2(\Omega)$ by*

$$(\mathscr{L}_p f)(z) := \sum_{i=1}^{N} \left(\frac{\langle A_i z, w\rangle}{\operatorname{sign} \Re(\langle A_i z, w\rangle)} \right)^p f\left(\langle A_i z, w\rangle^{-1} A_i z \right).$$

Then the operators $\mathscr{L}_p$ are well-defined bounded linear operators on $\mathcal{A}^2(\Omega)$ and:

(i) There exist $C, \kappa, \gamma > 0$ such that for all $p \in \mathbb{C}$ and $n \geq 1$ we have

$$\mathfrak{s}_n(\mathscr{L}_p) \leq C \exp\left(\kappa |p| - \gamma n^{\frac{1}{d-1}} \right).$$

In particular each $\mathscr{L}_p$ is trace-class.

(ii) For every $p \in \mathbb{C}$ and $n \geq 1$ we have

$$\operatorname{tr} \mathscr{L}_p^n = \sum_{|\mathtt{i}|=n} \rho(A_\mathtt{i})^p \prod_{j=2}^{d} \left(1 - \frac{\lambda_j(A_\mathtt{i})}{\lambda_1(A_\mathtt{i})} \right)^{-1} = \sum_{|\mathtt{i}|=n} \frac{\lambda_1(A_\mathtt{i})^{d-1} \rho(A_\mathtt{i})^p}{p'_{A_\mathtt{i}}(\lambda_1(A_\mathtt{i}))}$$

(iii) For every $p \in \mathbb{R}$ the spectral radius of $\mathscr{L}_p$ is equal to

$$\lim_{n\to\infty} \left(\sum_{|\mathtt{i}|=n} \|A_\mathtt{i}\|^p \right)^{\frac{1}{n}}.$$

(iv) For all $p \in \mathbb{R}$ the spectral radius of $\mathscr{L}_p$ is a simple eigenvalue of $\mathscr{L}_p$ and there are no other eigenvalues of the same modulus.

Theorem 4.1 above is precisely the special case $\ell = 1$ of [29, Theorem 11].

Theorem 2.1 can thus be seen as a version of the eigenvalue-problem approach discussed in the previous section, but one which takes advantage of the special additional structure of trace-class operators. Note that since trace-class operators are compact operators they are very far from being invertible, and indeed an important feature of the hypotheses of Theorem 4.1 is that the transformations A_i map a (not necessarily connected) patch of $\mathbb{RP}^{d-1}$ strictly inside itself – which results in a non-invertible action on the associated function space – as opposed to acting transitively on $\mathbb{RP}^{d-1}$.

This feature is precisely the content of the multicone hypothesis, and indeed the non-invertibility of the action on $\mathbb{RP}^{d-1}$ is critical in constructing a space on which the operators $\mathscr{L}_p$ can act in a trace-class manner. As such any extension of the method of Theorem 2.1 to families of matrices with non-real eigenvalues is therefore likely to be impossible since such matrices would tend to act transitively on the phase space $\mathbb{RP}^{d-1}$, preventing the construction of a suitable domain for a trace-class operator to act upon.

The operator considered in Theorem 4.1 is closely related to an operator constructed in earlier work of M. Pollicott [32], but with two differences. In the first instance, Pollicott's construction deals only with the simpler case where the parameter p is equal to zero – in which case several factors in Theorem 4.1 evaluate to 1 and disappear, simplifying many steps of the argument. Secondly, Pollicott's article treats only the case $d = 2$ in detail and omits many details of the construction of the domain Ω in other cases. The additional complexities inherent in this construction in higher dimensions are responsible for much of the length of the article [29] in which Theorem 4.1 is proved.

5. Proof of Theorem 2.1. The following result summarises the classical results on traces and determinants of trace-class operators on Hilbert spaces which will be required in our proof. It is a combination of several results from [38, §3], with the exception of the determinant formula for a_n which may be found instead in, for example, [37, Theorem 6.8] or [15, Theorem IV.5.2].

THEOREM 5.1. *Let $\mathscr{H}$ be a complex separable Hilbert space, let $\mathscr{L}$ be a trace-class operator acting on $\mathscr{H}$, and define $a_0 := 1$ and*

$$\begin{aligned} a_n :=& \frac{(-1)^n}{n!} \det \begin{pmatrix} \operatorname{tr}\mathscr{L} & n-1 & 0 & \cdots & 0 & 0 \\ \operatorname{tr}\mathscr{L}^2 & \operatorname{tr}\mathscr{L} & n-2 & \cdots & 0 & 0 \\ \operatorname{tr}\mathscr{L}^3 & \operatorname{tr}\mathscr{L}^2 & \operatorname{tr}\mathscr{L} & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \operatorname{tr}\mathscr{L}^{n-1} & \operatorname{tr}\mathscr{L}^{n-2} & \operatorname{tr}\mathscr{L}^{n-3} & \cdots & \operatorname{tr}\mathscr{L} & 1 \\ \operatorname{tr}\mathscr{L}^{n} & \operatorname{tr}\mathscr{L}^{n-1} & \operatorname{tr}\mathscr{L}^{n-2} & \cdots & \operatorname{tr}\mathscr{L}^2 & \operatorname{tr}\mathscr{L} \end{pmatrix} \\ =& \sum_{k=1}^{n} \frac{(-1)^k}{k!} \sum_{\substack{n_1,\ldots,n_k \geq 1 \\ n_1+\cdots+n_k=n}} \prod_{\ell=1}^{k} \frac{\operatorname{tr}\mathscr{L}^{n_\ell}}{n_\ell} \end{aligned}$$

for every $n \geq 1$. Then the power series $\mathscr{D}(z) := \sum_{n=0}^{\infty} a_n z^n$ converges for all $z \in \mathbb{C}$. The function $\mathscr{D}\colon \mathbb{C} \to \mathbb{C}$ is holomorphic, the zeros of $\mathscr{D}$ are precisely the reciprocals of the nonzero eigenvalues of $\mathscr{L}$, and the degree of each zero of $\mathscr{D}$ is equal to the algebraic multiplicity of the corresponding eigenvalue of $\mathscr{L}$. Moreover the coefficients a_n satisfy the estimate

$$|a_n| \leq \sum_{i_1<i_2<\cdots<i_n} \mathfrak{s}_{i_1}(\mathscr{L}) \cdots \mathfrak{s}_{i_n}(\mathscr{L}) \tag{5.1}$$

for every $n \geq 1$.

Proof. In the notation of all three of [15, 37, 38] the function $\mathscr{D}$ corresponds to the function $z \mapsto \det(I - z\mathscr{L})$. The fact that this function satisfies $\mathscr{D}(z) = \sum_{n=0}^{\infty} a_n z^n$ with each a_n given by the determinant formula above, and with the series being absolutely convergent for all $z \in \mathbb{C}$, follows from [15, Theorem IV.5.2]. Obviously

$\mathscr{D}(0) = 1 \neq 0$, so the claim regarding the zeros of $\mathscr{D}$ pertains to its values at nonzero numbers $z \in \mathbb{C}$ only. If $\lambda \in \mathbb{C} \setminus \{0\}$ is *not* the reciprocal of an eigenvalue of $\mathscr{L}$ then $I - \lambda\mathscr{L}$ is invertible, since the nonzero points of the spectrum of a compact operator are precisely its eigenvalues; thus $\mathscr{D}(\lambda) = \det(I - \lambda\mathscr{L}) \neq 0$ by [37, Theorem 3.9]. On the other hand if $\lambda \in \mathbb{C} \setminus \{0\}$ is the reciprocal of an eigenvalue of $\mathscr{L}$ of algebraic multiplicity $k > 0$ then λ is a zero of $\mathscr{D}$ of degree k by [37, Theorem 3.10]. As a consequence of [38, Lemma 3.3] and in particular formula (3.5) in that lemma, the absolute value of the coefficient a_n for $n \geq 1$ is precisely the absolute value of the trace of the exterior power $\wedge^n \mathscr{L}$. By [38, Theorem 3.1] this trace is bounded by the trace norm $\| \wedge^n \mathscr{L}\|_1$ and by equation (3.8) in the proof of [38, Lemma 3.3] this in turn is bounded by the expression on the right-hand side of (5.1). This yields (5.1). By equation (5.12) in the proof of [38, Theorem 5.4] we have

$$\mathscr{D}(z) = \det(I - z\mathscr{L}) = \exp\left(-\sum_{m=1}^{\infty} \frac{z^m \cdot \operatorname{tr} \mathscr{L}^m}{m}\right)$$

for all z in a sufficiently small complex neighbourhood of 0, with the series being absolutely convergent. Comparing the coefficients of z^n on either side of the resulting equation

$$\sum_{n=0}^{\infty} a_n z^n = \sum_{k=0}^{\infty} \frac{1}{k!} \left(-\sum_{m=1}^{\infty} \frac{z^m \cdot \operatorname{tr} \mathscr{L}^m}{m}\right)^k$$

yields the second expression for the coefficients a_n, and the proof is complete. ∎

We also require the following elementary lemma:

LEMMA 5.2. *For each $\gamma, \alpha > 0$ there exists a constant $K = K(\alpha, \gamma) > 0$ such that*

$$\sum_{n=m}^{\infty} \exp\left(-\gamma n^{\alpha}\right) \leq K \exp\left(-\frac{\gamma}{2^{1+\alpha}} m^{\alpha}\right)$$

for all $m \geq 1$.

Proof. Fix γ and α. By adjusting the constant K if necessary we may without loss of generality assume $m \geq 2$. Define

$$C := \sup\left\{u^{\frac{1}{\alpha}-1} \exp\left(-\frac{\gamma}{2}u\right) : u \geq 1\right\} > 0.$$

Since clearly $e^{-\gamma n^{\alpha}} \leq \int_{n-1}^{n} e^{-\gamma t^{\alpha}} dt$ for every integer n we have

$$\begin{aligned}
\sum_{n=m}^{\infty} \exp\left(-\gamma n^{\alpha}\right) &\leq \int_{m-1}^{\infty} \exp\left(-\gamma t^{\alpha}\right) dt \\
&= \frac{1}{\alpha} \int_{(m-1)^{\alpha}}^{\infty} u^{\frac{1}{\alpha}-1} \exp\left(-\gamma u\right) du \\
&\leq \frac{C}{\alpha} \int_{(m-1)^{\alpha}}^{\infty} \exp\left(-\frac{\gamma}{2}u\right) du \\
&= \frac{C}{\alpha} \exp\left(-\frac{\gamma}{2}(m-1)^{\alpha}\right) \leq \frac{C}{\alpha} \exp\left(-\frac{\gamma}{2^{1+\alpha}} m^{\alpha}\right)
\end{aligned}$$

for every $m \geq 2$ and the result follows. ∎

We may now begin the proof of Theorem 2.1. Fix $A_1, \ldots, A_N$ and $p \in \mathbb{R}$ as in Theorem 2.1. By Theorem 4.1 there exist a complex separable Hilbert space $\mathscr{H}$ and a trace-class linear operator $\mathscr{L}_p \colon \mathscr{H} \to \mathscr{H}$ such that $\varrho_p(A_1, \ldots, A_N)$ is a simple isolated eigenvalue of $\mathscr{L}_p$, such that all other eigenvalues have absolute value strictly smaller than $\varrho_p(A_1, \ldots, A_N)$, such that

$$\operatorname{tr} \mathscr{L}_p^n = \sum_{|\mathtt{i}|=n} \rho(A_\mathtt{i})^p \prod_{j=2}^{d} \left(1 - \frac{\lambda_j(A_\mathtt{i})}{\lambda_1(A_\mathtt{i})}\right) = \sum_{|\mathtt{i}|=n} \frac{\lambda_1(A_\mathtt{i})^{d-1} \rho(A_\mathtt{i})^p}{p'_{A_\mathtt{i}}(\lambda_1(A_\mathtt{i}))}$$

for every $n \geq 1$ and such that there exist constants $C_1, \gamma_1 > 0$ such that $\mathfrak{s}_n(\mathscr{L}_p) \leq C_1 \exp(-\gamma_1 n^{\frac{1}{d-1}})$ for every $n \geq 1$. Define the sequence (t_n) in accordance with Theorem 2.1 and note that we have $t_n = \operatorname{tr} \mathscr{L}_p^n$ for every $n \geq 1$. For each $n \geq 0$ let a_n be as defined in Theorem 5.1 and note that this coincides with the definition of the sequence a_n in Theorem 2.1. We claim that there exist $C_2, \gamma_2 > 0$ such that

$$|a_n| \leq C_2 \exp\left(-\gamma_2 n^{\frac{d}{d-1}}\right) \tag{5.2}$$

for every $n \geq 1$. To see this let $n \geq 1$ and observe that by Theorem 5.1

$$\begin{aligned}
|a_n| \leq \sum_{i_1 < \cdots < i_n} \mathfrak{s}_{i_1}(\mathscr{L}_p) \cdots \mathfrak{s}_{i_n}(\mathscr{L}_p) &\leq \sum_{i_1 < \cdots < i_n} \prod_{k=1}^{n} C_1 \exp\left(-\gamma_1 i_k^{\frac{1}{d-1}}\right) \\
&= C_1^n \sum_{i_1 < \cdots < i_n} \exp\left(-\gamma_1 \sum_{k=1}^{n} i_k^{\frac{1}{d-1}}\right) \\
&\leq C_1^n \sum_{i_1=1}^{\infty} \cdots \sum_{i_n=n}^{\infty} \exp\left(-\gamma_1 \sum_{k=1}^{n} i_k^{\frac{1}{d-1}}\right) \\
&= C_1^n \prod_{k=1}^{n} \sum_{i_k=k}^{\infty} \exp\left(-\gamma_1 i_k^{\frac{1}{d-1}}\right) \\
&\leq C_1^n K^n \prod_{k=1}^{n} \exp\left(-\frac{\gamma_1}{2^{\frac{d}{d-1}}} k^{\frac{1}{d-1}}\right) \\
&= C_1^n K^n \exp\left(-\frac{\gamma_1}{2^{\frac{d}{d-1}}} \sum_{k=1}^{n} k^{\frac{1}{d-1}}\right) \\
&\leq C_1^n K^n \exp\left(-\frac{(d-1)\gamma_1}{d 2^{\frac{d}{d-1}}} n^{\frac{d}{d-1}}\right)
\end{aligned}$$

where we have used Lemma 5.2 with $\alpha = \frac{1}{d-1}$ and have also used the elementary inequality

$$\sum_{k=1}^{n} k^{\frac{1}{d-1}} \geq \int_0^n t^{\frac{1}{d-1}} \, dt = \frac{d-1}{d} n^{\frac{d}{d-1}}$$

which is valid since the series is an upper Riemann sum of the integral. The claim follows easily.

Now define a function $\mathscr{D} \colon \mathbb{C} \to \mathbb{C}$ by $\mathscr{D}(z) := \sum_{n=0}^{\infty} a_n z^n$. It is clear from the estimate (5.2) that this power series has infinite radius of convergence and therefore $\mathscr{D}$ is a well-defined holomorphic function on $\mathbb{C}$. By Theorem 5.1 we have $\mathscr{D}(z) = \det(I - z\mathscr{L}_p)$ for all $z \in \mathbb{C}$ and the zeros of $\mathscr{D}$ are precisely the reciprocals of the

nonzero eigenvalues of $\mathscr{L}_p$ with the degree of each zero being equal to the algebraic multiplicity of the corresponding eigenvalue. By Theorem 4.1, $\varrho_p(A_1,\dots,A_N)$ is the largest eigenvalue of $\mathscr{L}_p$ in absolute value and is a simple eigenvalue. It follows that we may choose a circular contour Γ in $\mathbb{C}$ which is centred somewhere on the real line, passes through 0, encloses $1/\varrho_p(A_1,\dots,A_N)$ and does not enclose or intersect the reciprocal of any eigenvalue of $\mathscr{L}_p$ other than $\varrho_p(A_1,\dots,A_N)$. Let $c\in\mathbb{R}$ and $R>0$ denote the centre point and radius of Γ respectively. Since Γ does not intersect the reciprocal of any eigenvalue of $\mathscr{L}_p$ the function $\mathscr{D}$ does not have any zeros on Γ, so by compactness

$$\inf_{|z-c|=R}|\mathscr{D}(z)|>0.$$

For each $n\geq 1$ define a function $\mathscr{D}_n\colon\mathbb{C}\to\mathbb{C}$ by $\mathscr{D}_n(z):=\sum_{k=0}^{n}a_kz^k$. Obviously each $\mathscr{D}_n$ is a polynomial and is therefore holomorphic on $\mathbb{C}$. Via Lemma 5.2 the estimate (5.2) implies

$$\begin{aligned}\sup_{|z-c|\leq R}\sum_{k=n}^{\infty}|a_kz^k| &\leq \sum_{k=n}^{\infty}C_2(2R)^k\exp\left(-\gamma_2k^{\frac{d}{d-1}}\right)\\ &\leq \sum_{k=n}^{\infty}C_3\exp\left(-\gamma_3k^{\frac{d}{d-1}}\right)\leq C_4\exp\left(-\gamma_4n^{\frac{d}{d-1}}\right)\end{aligned}\tag{5.3}$$

for all $n\geq 1$ and some suitable constants $C_3,C_4,\gamma_3,\gamma_4>0$. In particular

$$\lim_{n\to\infty}\sup\{|\mathscr{D}(z)-\mathscr{D}_n(z)|\colon |z-c|\leq R\}=0\tag{5.4}$$

and therefore there exists $n_0\geq 1$ such that for all $n\geq n_0$

$$\sup_{|z-c|=R}|\mathscr{D}(z)-\mathscr{D}_n(z)|<\inf_{|z-c|=R}|\mathscr{D}(z)|.$$

Applying Rouché's theorem on the circular contour Γ we deduce that for all $n\geq n_0$ the functions $\mathscr{D}$ and $\mathscr{D}_n$ have the same number of zeros inside the contour Γ, and the total degree of the zeros inside Γ is the same for the function $\mathscr{D}$ as it is for the function $\mathscr{D}_n$. Since $\mathscr{D}$ has a unique zero inside Γ and that zero is simple this means that $\mathscr{D}_n$ has a unique zero inside Γ for all large enough n, and this zero is simple. Call this zero r_n. Since $\mathscr{D}_n$ is a polynomial with real coefficients its zeros are symmetrically located with respect to reflection in the real axis. Since the contour Γ is circular with real centre, a zero of $\mathscr{D}_n$ is enclosed by Γ if and only if the complex conjugate of that zero is also so enclosed. It follows that the complex conjugate of r_n is also enclosed by the contour Γ and is therefore also a zero of $\mathscr{D}_n$. But $\mathscr{D}_n$ has a unique zero inside Γ. These statements can only be compatible if r_n is equal to its own complex conjugate, and we conclude that r_n is real. Since r_n is enclosed by Γ and is real it necessarily lies on the interval $(0,2R)$ and is the unique zero of $\mathscr{D}_n$ on that interval. In particular it is the smallest positive zero of the polynomial $\mathscr{D}_n$.

Define $r_\infty:=1/\varrho_p(A_1,\dots,A_N)\in(0,2R)$. To complete the proof of the theorem we will show that

$$\left|\frac{1}{r_\infty}-\frac{1}{r_n}\right|=O\left(\exp\left(-\gamma_4n^{\frac{d}{d-1}}\right)\right).$$

We first require a lower bound for the derivative $\mathscr{D}'(z)$ for z close to r_∞. Since $r_\infty=1/\varrho_p(A_1,\dots,A_N)$ is a simple zero of $\mathscr{D}$ we have $\mathscr{D}'(r_\infty)\neq 0$, and since it is also necessarily an isolated zero we may choose $\delta>0$ such that $|\mathscr{D}'(z)|\neq 0$ for all $z\in\mathbb{C}$

with $|z - r_\infty| \le \delta$, such that $\mathscr{D}(z) \neq 0$ for all $z \in \mathbb{C}$ with $0 < |z - r_\infty| \le \delta$, and such that the closed disc of radius δ and centre r_∞ is enclosed by the contour Γ. Since by compactness

$$\inf_{|z-r_\infty|=\delta} |\mathscr{D}(z)| > 0$$

it follows via (5.4) in the same manner as before that there exists $n_1 \ge n_0$ such that for all $n \ge n_1$

$$\sup_{|z-r_\infty|=\delta} |\mathscr{D}(z) - \mathscr{D}_n(z)| < \inf_{|z-r_\infty|=\delta} |\mathscr{D}(z)|.$$

Applying Rouché's theorem again, this time to the circular contour with centre r_∞ and radius δ, we see that for each $n \ge n_1$ there is a unique zero of $\mathscr{D}_n$ within distance δ of r_∞. Since the disc of radius δ and centre r_∞ is enclosed by Γ, and Γ encloses a unique zero of $\mathscr{D}_n$, we conclude that this zero must be r_n and therefore $|r_n - r_\infty| < \delta$ for all $n \ge n_1$.

Now define

$$\kappa := \inf \{|\mathscr{D}'(z)| \colon |z - r_\infty| \le \delta\} > 0.$$

Since $\mathscr{D}_n$ is a polynomial with real coefficients it takes only real values when restricted to $\mathbb{R}$ and therefore the same is true of $\mathscr{D}$ since it is the pointwise limit of $\mathscr{D}_n$ as $n \to \infty$. Let $n \ge n_1$ and suppose that $r_n \neq r_\infty$. By the Mean Value Theorem it follows that there exists a real number t in the interval from r_n to r_∞ such that

$$\frac{\mathscr{D}(r_n) - \mathscr{D}(r_\infty)}{r_n - r_\infty} = \mathscr{D}'(t).$$

Since clearly $|r_\infty - t| \le |r_\infty - r_n| \le \delta$ we have $|\mathscr{D}'(t)| \ge \kappa$ and therefore

$$|r_n - r_\infty| \le \kappa^{-1} |\mathscr{D}(r_n) - \mathscr{D}(r_\infty)|.$$

This inequality is obviously also true for integers $n \ge n_1$ such that $r_n = r_\infty$. In particular for all $n \ge n_1$ we have

$$|r_n - r_\infty| \le \kappa^{-1} |\mathscr{D}(r_n) - \mathscr{D}_n(r_n)|$$

using the fact that $\mathscr{D}_n(r_n) = 0 = \mathscr{D}(r_\infty)$. Thus

$$|r_n - r_\infty| \le \kappa^{-1} |\mathscr{D}(r_n) - \mathscr{D}_n(r_n)| = \kappa^{-1} \left| \sum_{k=n+1}^{\infty} a_k r_n^k \right| \le \kappa^{-1} C_4 \exp\left(-\gamma_4 n^{\frac{d}{d-1}}\right)$$

for all $n \ge n_1$ using (5.3). We in particular have $\lim_{n\to\infty} r_n = r_\infty$. If $n_2 \ge n_1$ is taken large enough that for all $n \ge n_2$ we have $r_n \ge \frac{1}{2} r_\infty$, then for all $n \ge n_2$ we have

$$\left| \frac{1}{r_n} - \frac{1}{r_\infty} \right| = \frac{|r_n - r_\infty|}{r_n r_\infty} \le \frac{|r_n - r_\infty|}{\frac{1}{2} r_\infty^2} \le \frac{2C_4}{\kappa r_\infty^2} \exp\left(-\gamma_4 n^{\frac{d}{d-1}}\right)$$

and this completes the proof of the theorem.

6. Example: a pair of matrices considered by Jungers and Protasov. In the article [23] the p-radius of the pair (A_1, A_2) defined by

$$A_1 := \begin{pmatrix} \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{3}{5} \end{pmatrix}, \qquad A_2 := \begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ 0 & \frac{1}{5} \end{pmatrix}$$

was investigated motivated by its connection with Chaikin's subdivision schemes and the L^p regularity of refinable functions. The reader may easily check that if we define

$$X := \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

then the matrices $X^{-1}A_1X$ and $X^{-1}A_2X$ are both positive, so the pair (A_1, A_2) strictly preserves a cone and Theorem 2.1 may be applied thereto. The results of applying the various methods of estimation to $\varrho_{3.5}(A_1, A_2)$ are tabulated in Figures 6.1–6.5 below. The reader will notice that by far the best results are those obtained by Theorem 2.1: the estimate obtained by evaluating all products $A_{\mathtt{i}}$ of length up to 12 yields the estimate $0.19773298680753190957\ldots$, where all 20 significant figures remain stable at higher values of n. Estimates of comparable complexity using the method of §3.3 give only the first two significant figures, albeit rigorously; the naïve upper and lower estimates described in §3.1 are not even sufficient to establish the first significant digit of $\varrho_{3.5}(A_1, A_2)$. The methods of §3.2 and §3.4 perform somewhat better, being able to give non-rigorous estimates accurate to several decimal places. We also observe that the upper estimate arising from logarithmic convexity,

$$\varrho_{3.5}(A_1, A_2) \simeq \sqrt{\varrho_3(A_1, A_2)\varrho_4(A_1, A_2)} = \sqrt{\rho\left(A_1^{\otimes 3} + A_2^{\otimes 3}\right)\rho\left(A_1^{\otimes 4} + A_2^{\otimes 4}\right)},$$

gives a rigorous upper bound of

$$\varrho_{3.5}(A_1, A_2) \leq 0.1986720360\ldots$$

which, remarkably, is more accurate than several of the other methods employed. Applying Theorem 2.1 with $n = 20$ gives the estimate

$$\varrho_{3.5}(A_1, A_2) \simeq 0.19773298680753190957347710334795037036402463415 67\ldots$$

where the digits displayed were likewise found empirically to be stable with respect to increase of n, and this estimate provides the value (2.3) mentioned in the introduction.

n	Naïve upper estimate	Naïve lower estimate
1	0.41014 02388	0.00003 71719
2	0.29717 45163	0.00265 32644
3	0.26212 69438	0.01107 32061
4	0.24497 10624	0.02270 50356
5	0.23489 87259	0.03497 48389
6	0.22831 70520	0.04666 81491
7	0.22369 66328	0.05735 30955
8	0.22028 14135	0.06694 70201
9	0.21765 70884	—
10	0.21557 86195	—
11	0.21389 22442	—
12	0.21249 67903	—

FIG. 6.1. *The rigorous upper and lower estimates (3.1) and (3.2) applied to the pair (A_1, A_2) with $p = 3.5$. The upper estimate requires the computation of 2^n matrix products and the lower estimate 4^n products. For $n > 8$ the lower estimate was omitted due to the large number of products to be computed and the poor quality of the estimates.*

n	Upper estimate $\mathsf{b}_p(n)^{\frac{1}{n}}$	Lower estimate $d^{-\frac{1}{n}}\mathsf{a}_p(n)^{\frac{1}{n}}$
1	0.20779 00346	0.08095 43081
2	0.20474 70800	0.14134 17665
3	0.20294 52224	0.16241 04530
4	0.20180 54158	0.17198 46647
5	0.20104 31937	0.17732 22741
6	0.20050 82647	0.18073 86055
7	0.20011 68386	0.18313 47477
8	0.19982 00191	0.18492 14944
9	0.19958 80621	0.18631 15004
10	0.19940 21599	0.18742 65582
11	0.19924 99839	0.18834 21232
12	0.19912 31811	0.18910 78446

FIG. 6.4. *Rigorous upper and lower estimates given by the algorithm of Jungers and Protasov applied to the pair* (A_1, A_2) *with* $p = 3.5$.

Mesh size	Estimate
10	0.22765 40788
100	0.19986 86395
1000	0.19785 78266
10000	0.19774 13329
100000	0.19773 40963

FIG. 6.2. *Estimates of* $\varrho_{3.5}(A_1, A_2)$ *given by the eigenvalue method described in §3.4.*

Sample length	Number of runs	Resampled Monte Carlo estimate
10	10	0.20663 64774
10	100	0.19336 14906
10	1000	0.19472 39505
100	10	0.19078 48295
100	100	0.19724 80647
100	1000	0.19706 73206
1000	10	0.19171 01011
1000	100	0.19752 20499
1000	1000	0.19768 32282
10000	10	0.19460 13140
10000	100	0.19737 86045
10000	1000	0.19766 64507

FIG. 6.3. *Some representative instances of J. Vanneste's resampled Monte Carlo scheme applied to the pair* (A_1, A_2) *with* $p = 3.5$ *over various parameter ranges.*

7. Conclusions. We have introduced a new method for estimating the p-radius of low-cardinality sets of positive or dominated matrices and investigated its effective-

n	Estimate $1/r_n$
1	0.50193 86416 68481 22831 92327
2	—
3	0.25470 11941 19890 64296 65247
4	—
5	0.19747 18486 52733 86575 36851
6	0.19773 76208 73169 67676 89071
7	0.19773 30386 40809 03204 40047
8	0.19773 29865 81371 43318 96314
9	0.19773 29868 07433 20636 81181
10	0.19773 29868 07532 62503 56803
11	0.19773 29868 07531 90980 60910
12	0.19773 29868 07531 90957 29023

FIG. 6.5. *Estimates of $\varrho_{3.5}(A_1, A_2)$ provided by Theorem 2.1. For $n = 2, 4$ the polynomial $\sum_{k=0}^{n} a_k x^k$ has no real roots and the quantity $1/r_n$ is undefined.*

ness in the case of a particular pair of matrices considered by Jungers and Protasov in connection with applications to Chaikin's subdivision scheme. We have compared its results to those of a number of other estimation methods in the case of that example and obtained results apparently accurate to within an absolute error of approximately 10^{-20}, versus approximately 10^{-2} to 10^{-6} for rival methods.

The new method has the disadvantage that the number of matrix products whose leading eigenvalue and characteristic polynomial must be computed in order to obtain the n^{th} approximation to $\varrho_p(A_1, \ldots, A_N)$ grows at an exponential rate of N^n with respect to the integer n. Since those two properties of a matrix product are invariant with respect to cyclic permutation of the product concerned, each coefficient thus t_n requires the consideration of about N^n/n distinct matrix products, one for each equivalence class modulo cyclic permutation. Computing all d eigenvalues of each product by the QR method may be expected to take a total time of order d^3 per eigenvalue, making d^4 per product; and since the n^{th} approximation requires the computation of the first n coefficients t_k this suggests a complexity on the order of $d^4 N^n$. In particular if the number of matrices N being considered is greater than around 4, the computational burden of producing accurate results may be prohibitively large. On the other hand this disadvantage of a factor of N^n in the complexity is shared by the methods of §3.1 and §3.3. In view of this consideration, when N is large the methods of §3.2 and §3.4 may be preferable. Our method also, as presently formulated, does not provide a rigorous estimate of its own accuracy, and if rigorous bounds are sought then the method of §3.3, possibly in combination with the logarithmic-convexity bound (3.4) may be applied instead.

The constants K and γ of Theorem 2.1 both depend on the estimate in Theorem 4.1(i) for the singular values of the operator $\mathscr{L}_p$, and those estimates in turn are functions of the strength of contractiveness of the action of the matrices $A_1, \ldots, A_N$ on the associated multicone. In the case of two-dimensional positive matrices it is possible to give explicit estimates for the decay of the singular values of $\mathscr{L}_p$, and this is attempted in [24] at the parameter value $p = 0$. On the other hand the constant K is further affected by factors such as the distance separating the largest two eigenvalues of the operator $\mathscr{L}_p$, and as such it may be more difficult to obtain *a priori* estimates

for this constant. Furthermore, in dimensions higher than two even the constant γ is not obtained in a fully constructive manner since it is affected by the cardinality of the relative covers arising in the application of [3, Theorem 4.7] to the domain Ω which arises from certain slices of complex cones, and the finiteness of that cardinality is established nonconstructively. The problem of obtaining explicit estimates of γ in higher dimensions is therefore substantially more challenging.

8. Acknowledgements. This research was supported by the Leverhulme Trust (Research Project Grant number RPG-2016-194). The author thanks two referees for their helpful and thorough comments and suggestions.

REFERENCES

[1] Avila, A., Bochi, J., and Yoccoz, J.-C. Uniformly hyperbolic finite-valued $\mathrm{SL}(2,\mathbb{R})$-cocycles. *Comment. Math. Helv. 85*, 4 (2010), 813–884.

[2] Bai, Z.-Q. On the cycle expansion for the Lyapunov exponent of a product of random matrices. *J. Phys. A 40*, 29 (2007), 8315–8328.

[3] Bandtlow, O. F., and Jenkinson, O. Explicit eigenvalue estimates for transfer operators acting on spaces of holomorphic functions. *Adv. Math. 218*, 3 (2008), 902–925.

[4] Barnsley, M. F., and Vince, A. Real projective iterated function systems. *J. Geom. Anal. 22*, 4 (2012), 1137–1172.

[5] Berger, G. O., and Jungers, R. M. A converse Lyapunov theorem for p-dominant switched linear systems. In *2019 18th European Control Conference (ECC)* (2019), pp. 1263–1268.

[6] Bochi, J., and Gourmelon, N. Some characterizations of domination. *Math. Z. 263*, 1 (2009), 221–231.

[7] Bochi, J., and Morris, I. D. Continuity properties of the lower spectral radius. *Proc. Lond. Math. Soc. (3) 110*, 2 (2015), 477–509.

[8] Brundu, M., and Zennaro, M. Invariant multicones for families of matrices. *Ann. Mat. Pura Appl. (4)*. To appear.

[9] Cabrelli, C. A., Heil, C., and Molter, U. M. Self-similarity and multiwavelets in higher dimensions. *Mem. Amer. Math. Soc. 170*, 807 (2004), viii+82.

[10] Crisanti, A., Paladin, G., and Vulpiani, A. Generalized Lyapunov exponents in high-dimensional chaotic dynamics and products of large random matrices. *J. Statist. Phys. 53*, 3-4 (1988), 583–601.

[11] Crisanti, A., Paladin, G., and Vulpiani, A. *Products of random matrices in statistical physics*, vol. 104 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, 1993. With a foreword by Giorgio Parisi.

[12] Feng, D.-J., and Lau, K.-S. The pressure function for products of non-negative matrices. *Math. Res. Lett. 9*, 2-3 (2002), 363–378.

[13] Feng, D.-J., Lo, C.-H., and Shen, S. Uniformity of Lyapunov exponents for non-invertible matrices. *Ergodic Theory Dynam. Systems* (2019). To appear.

[14] Feng, D.-J., and Shmerkin, P. Non-conformal repellers and the continuity of pressure for matrix cocycles. *Geom. Funct. Anal. 24*, 4 (2014), 1101–1128.

[15] Gohberg, I., Goldberg, S., and Krupnik, N. *Traces and determinants of linear operators*, vol. 116 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2000.

[16] Grassberger, P. Go with the winners: a general Monte Carlo strategy. In *Proceedings of the Europhysics Conference on Computational Physics (CCP 2001) (Aachen)* (2002), vol. 147, pp. 64–70.

[17] Guivarc'h, Y., and Le Page, E. Simplicité de spectres de Lyapounov et propriété d'isolation spectrale pour une famille d'opérateurs de transfert sur l'espace projectif. In *Random walks and geometry*. Walter de Gruyter, Berlin, 2004, pp. 181–259.

[18] Horn, R. A., and Johnson, C. R. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[19] Jenkinson, O., and Pollicott, M. Calculating Hausdorff dimensions of Julia sets and Kleinian limit sets. *Amer. J. Math. 124*, 3 (2002), 495–545.

[20] Jenkinson, O., and Pollicott, M. Orthonormal expansions of invariant densities for expanding maps. *Adv. Math. 192*, 1 (2005), 1–34.

[21] Jungers, R. *The joint spectral radius: theory and applications*, vol. 385 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 2009.

[22] Jungers, R. M., and Protasov, V. Y. Weak stability of switching dynamical systems and

fast computation of the p-radius of matrices. In *49th IEEE Conference on Decision and Control (CDC)* (Dec 2010), pp. 7328–7333.

[23] Jungers, R. M., and Protasov, V. Y. Fast methods for computing the p-radius of matrices. *SIAM J. Sci. Comput. 33*, 3 (2011), 1246–1266.

[24] Jurga, N., and Morris, I. D. Effective estimates on the top Lyapunov exponent for random matrix products. *Nonlinearity 32* (2019), 4117–4146.

[25] Keller, G., and Liverani, C. Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28*, 1 (1999), 141–152.

[26] Lau, K.-S., and Wang, J. Characterization of L^p-solutions for the two-scale dilation equations. *SIAM J. Math. Anal. 26*, 4 (1995), 1018–1046.

[27] Morris, I. D. An inequality for the matrix pressure function and applications. *Adv. Math. 302* (2016), 280–308.

[28] Morris, I. D. On Falconer's formula for the generalized Rényi dimension of a self-affine measure. *Ann. Acad. Sci. Fenn. Math. 42*, 1 (2017), 227–238.

[29] Morris, I. D. Fast approximation of the affinity dimension for dominated affine iterated function systems. preprint arXiv:1807.09084v1, 2018.

[30] Ogura, M., and Martin, C. F. Generalized joint spectral radius and stability of switching systems. *Linear Algebra Appl. 439*, 8 (2013), 2222–2239.

[31] Ogura, M., Preciado, V. M., and Jungers, R. M. Efficient method for computing lower bounds on the p-radius of switched linear systems. *Systems Control Lett. 94* (2016), 159–164.

[32] Pollicott, M. Maximal Lyapunov exponents for random matrix products. *Invent. Math. 181*, 1 (2010), 209–226.

[33] Pollicott, M., and Jenkinson, O. Computing invariant densities and metric entropy. *Comm. Math. Phys. 211*, 3 (2000), 687–703.

[34] Pollicott, M., and Vytnova, P. Estimating singularity dimension. *Math. Proc. Cambridge Philos. Soc. 158*, 2 (2015), 223–238.

[35] Pollicott, M., and Weiss, H. How smooth is your wavelet? Wavelet regularity via thermodynamic formalism. *Comm. Math. Phys. 281*, 1 (2008), 1–21.

[36] Protasov, V. Y. The generalized joint spectral radius: a geometric approach. *Izv. Ross. Akad. Nauk Ser. Mat. 61*, 5 (1997), 99–136.

[37] Simon, B. Notes on infinite determinants of Hilbert space operators. *Advances in Math. 24*, 3 (1977), 244–273.

[38] Simon, B. *Trace ideals and their applications*, vol. 35 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge-New York, 1979.

[39] Sturman, R., and Thiffeault, J.-L. Lyapunov exponents for the random product of two shears. *Journal of Nonlinear Science* (September 2018).

[40] Tsitsiklis, J. N., and Blondel, V. D. 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*, 1 (1997), 31–40.

[41] Vanneste, J. Estimating generalized Lyapunov exponents for products of random matrices. *Phys. Rev. E 81* (Mar 2010), 036701.

[42] Wang, F. *Ergodic and algebraic properties of transfer operators for products of random matrices*. 2018. Thesis (D.Phil.) – Oxford University.

[43] Wang, Y. Two-scale dilation equations and the mean spectral radius. *Random Comput. Dynam. 4*, 1 (1996), 49–72.

[44] Zhou, D.-X. The p-norm joint spectral radius for even integers. *Methods Appl. Anal. 5*, 1 (1998), 39–54.