On Almost Everywhere Strong Convergence of Multidimensional Continued Fraction Algorithms

D.M. Hardcastle[1], K. Khanin[1-4]
Basic Research Institute in Mathematical Sciences
HP Laboratories Bristol
HPL-BRIMS-2000-12
May, 2000

E-mail: D.M.Hardcastle@ma.hw.ac.uk
K.Khanin@newton.cam.ac.uk

simultaneous diophantine approximations, Jacobi-Perron algorithm, multidimensional continued fractions, strong convergence, Lyapunov exponents

We describe a strategy which allows one to produce computer assisted proofs of almost everywhere strong convergence of Jacobi-Perron type algorithms in arbitrary dimension. Numerical work is carried out in dimension three to illustrate our method. To the best of our knowledge this is the first result on almost everywhere strong convergence in dimension greater than two.

[1] Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, UK
[2] Isaac Newton Institute for Mathematical Sciences, 20 Clarkson Road, Cambridge CB3 0EH, UK.
[3] BRIMS, Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, UK.
[4] Landau Institute for Theoretical Physics, Kosygina Str., 2, Moscow 117332, Russia

On almost everywhere strong convergence of multidimensional continued fraction algorithms

D. M. Hardcastle[1] and K. Khanin[1−4]

[1] Department of Mathematics, Heriot-Watt University,
Edinburgh EH14 4AS, UK. E-mail: D.M.Hardcastle@ma.hw.ac.uk

[2] Isaac Newton Institute for Mathematical Sciences,
20 Clarkson Road, Cambridge CB3 0EH, UK.
E-mail: K.Khanin@newton.cam.ac.uk

[3] BRIMS, Hewlett-Packard Laboratories,
Stoke Gifford, Bristol BS12 6QZ, UK.

[4] Landau Institute for Theoretical Physics,
Kosygina Str.,2, Moscow 117332, Russia.

Abstract

We describe a strategy which allows one to produce computer assisted proofs of almost everywhere strong convergence of Jacobi-Perron type algorithms in arbitrary dimension. Numerical work is carried out in dimension three to illustrate our method. To the best of our knowledge this is the first result on almost everywhere strong convergence in dimension greater than two.

1 Introduction

The construction of simultaneous rational approximations to an irrational vector is a classical problem which has been studied for over a hundred years [15, 24, 22, 21, 7, 5, 25, 28]. By "rational approximation" we mean a sequence of vectors of the form $(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)})$ which converge to a vector $(\omega_1, \dots, \omega_d) \in I^d = [0,1]^d$. This convergence can be understood in two different ways. The sequence converges in a weak sense if

$$\lim_{n\to\infty} \left\| \left(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)} \right) - (\omega_1, \dots, \omega_d) \right\| = 0. \tag{1}$$

While (1) is a tempting definition of convergence, it is not what is really required for applications. More useful is the following definition:

$$\lim_{n\to\infty} \|q(n)(\omega_1, \dots, \omega_d) - (p_1(n), \dots, p_d(n))\| = 0, \tag{2}$$

which we call strong convergence. Dirichlet's Theorem ([8]) states that any $(\omega_1, \ldots, \omega_d) \in I^d \backslash \mathbb{Q}^d$ has infinitely many approximations of the form $(\frac{p_1}{q}, \ldots, \frac{p_d}{q})$ such that, for $1 \leq j \leq d$,

$$\left|\omega_j - \frac{p_j}{q}\right| \leq \frac{1}{q^{1+1/d}}.$$

This obviously implies the existence of a strongly convergent sequence. Moreover,

$$\|q(\omega_1, \ldots, \omega_d) - (p_1, \ldots, p_d)\| \leq \sqrt{d} q^{-1/d}.$$

However, it is very difficult to find these approximations which are guaranteed to exist by Dirichlet's Theorem. Thus it is important to have an algorithm which allows one to construct strongly convergent approximations. By "algorithm" we mean a procedure which allows one to calculate ϵ-close strong approximations, i.e. approximations which satisfy

$$\|q(\omega_1, \ldots, \omega_d) - (p_1, \ldots, p_d)\| < \epsilon,$$

using not more than $c|\log \epsilon|$ binary operations. Such an algorithm is certainly known in dimension 1, namely the famous continued fraction algorithm which has many nice properties. Different generalizations of continued fractions to the multidimensional case have been known for more than a hundred years. These include the Jacobi-Perron algorithm (JPA) and other connected algorithms (see [15, 22, 7, 23, 24]). However the properties of these generalizations are much worse than the properties of continued fractions. For example, continued fractions give strongly convergent approximations for every $\omega \in [0, 1] \backslash \mathbb{Q}$ but all known multidimensional algorithms (see [7, 1, 2, 3]) do not have this property. This means that there always exist "bad" irrational vectors for which there is no strong convergence [6, 18]. Thus the most one can hope to prove is almost everywhere strong convergence, i.e. that strong convergence holds for points in a subset of I^d of full Lebesgue measure.

Numerical studies indicate (see [17, 3]) that the Jacobi-Perron algorithm and many other algorithms have the property of almost everywhere strong convergence in any dimension. The only rigorous results concerning strong convergence are for two dimensional algorithms. Almost everywhere strong convergence for the two dimensional JPA follows from an old paper of Paley and Ursell [21]. This fact was pointed out by Khanin [16] (see also Schweiger [29]). A computer assisted proof of almost everywhere strong convergence for a two dimensional modification of the JPA was independently provided by Fujita, Ito, Keane and Ohtsuki [14, 11]. Their approach can also be carried out in dimension greater than two, and this is closely related to our approach. In this paper we consider the same modification but in higher dimensions. We call this modification the ordered Jacobi-Perron algorithm (OJPA). For the two dimensional OJPA, as for the two dimensional JPA, one can prove almost everywhere strong convergence without using a computer (see Meester [20]). However, at present there are no theoretical results on strong convergence in dimension greater than two. In fact there is a principle difference between two dimensions and higher dimensions. In dimension two one can effectively describe the set of bad vectors, i.e. the vectors for which strong convergence does not hold. For example, in the case of the JPA these are vectors for which the integer entries produced by the algorithm grow very fast. At the same time it is known that for high enough dimension ($d \geq 5$)

even noble vectors (i.e. vectors which are fixed points of the algorithm) may be bad. The difference between two dimensions and higher dimensions can also be described on a more technical level. It turns out that in the case $d \geq 3$ strong convergence is due to the cancellation of positive and negative terms in certain matrix products. The absolute values of both positive and negative contributions grow exponentially, but their difference is exponentially small, while in dimension two both positive and negative contributions are exponentially small (see [12]).

The main purpose of this paper is to present a scheme which in principle allows one to give a rigorous proof of almost everywhere strong convergence for any Jacobi-Perron type algorithm. We will illustrate our strategy by discussing the OJPA which will be described in the next section. In our opinion, this version of the JPA is the best generalization of continued fractions. It shares many properties with the one dimensional case which makes the study of it easier than the study of other generalizations. For example, it has an extremely nice natural extension (see [13]). Our scheme consists of two parts. The first part is the reduction of the original problem to the calculation of certain concrete integrals which give estimates for the Lyapunov exponents. The second part involves carrying out a finite number of calculations numerically. In this sense the proof of the final result is computer assisted. In the present paper we deal with the first part of the scheme, namely the reduction part. The rigorous estimation of the Lyapunov exponents is carried out in a forthcoming paper [12].

Our aim here was to explain a simple approach which leads to a proof. The method is simple mathematically but not from the point of view of numerical studies. It is actually possible to use a more advanced scheme for a computer assisted proof but this will be explained in a forthcoming publication [12].

Acknowledgements. We are very grateful to Carlangelo Liverani for a suggestion which improved our original approach substantially. This suggestion was made at a conference which was part of the European Science Foundation's PRODYN (Probabilistic methods in non-hyperbolic dynamics) programme. We thank the European Science Foundation for the opportunity to participate in this programme. We also wish to thank Viviane Baladi and Arnaldo Nogueira for useful discussions. The first author is supported by the Engineering and Physical Sciences Research Council of the UK.

2 The ordered Jacobi-Perron algorithm

The ordered Jacobi-Perron algorithm (OJPA) was first introduced by Podsypanin [23] in two dimensions. It was then considered by Schweiger in arbitrary dimension [26]. We will define a slight variation of this which is, however, equivalent to Schweiger's version. In the papers by Podsypanin, Schweiger and others ([23, 26, 14, 11, 20]) this algorithm was called the modified JPA. However we think that the name ordered JPA describes its nature more accurately.

Consider the d-dimensional simplex

$$\Delta^d = \{\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d) \in I^d : 1 \geq \omega_1 \geq \omega_2 \geq \cdots \geq \omega_d \geq 0\}.$$

Define a transformation $T : \Delta^d \to \Delta^d$ by the following two steps. Firstly, given $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d) \in \Delta^d$ form the numbers $\frac{\omega_2}{\omega_1}, \ldots, \frac{\omega_d}{\omega_1}$ and $\{\frac{1}{\omega_1}\} = \frac{1}{\omega_1} - [\frac{1}{\omega_1}]$ (where $\{x\}$ and $[x]$

denote the fractional and integer parts of x respectively, so that $\{x\} = x - [x]$). Secondly, put $\{\frac{1}{\omega_1}\}$ in the position in the sequence $\frac{\omega_2}{\omega_1}, \dots, \frac{\omega_d}{\omega_1}$ prescribed by the ordering in the simplex. More precisely,

$$T(\omega_1, \dots, \omega_d) = \begin{cases} (\{\frac{1}{\omega_1}\}, \frac{\omega_2}{\omega_1}, \dots, \frac{\omega_d}{\omega_1}) & \text{if } \{\frac{1}{\omega_1}\} > \frac{\omega_2}{\omega_1}; \\ (\frac{\omega_2}{\omega_1}, \dots, \frac{\omega_j}{\omega_1}, \{\frac{1}{\omega_1}\}, \frac{\omega_{j+1}}{\omega_1}, \dots, \frac{\omega_d}{\omega_1}) & \text{if } \frac{\omega_j}{\omega_1} > \{\frac{1}{\omega_1}\} > \frac{\omega_{j+1}}{\omega_1}; \\ (\frac{\omega_2}{\omega_1}, \dots, \frac{\omega_d}{\omega_1}, \{\frac{1}{\omega_1}\}) & \text{if } \frac{\omega_d}{\omega_1} > \{\frac{1}{\omega_1}\}. \end{cases} \tag{3}$$

Definition 1. The transformation T is called the *ordered Jacobi-Perron transformation* in dimension d.

It can be shown that T is a locally expanding mapping and consequently has nice ergodic properties. Specifically, T, like all other transformations producing Jacobi-Perron type algorithms, has a unique absolutely continuous invariant probability measure which is ergodic [26]. One of the main advantages of the ordered Jacobi-Perron algorithm is that an explicit expression for the density of the invariant measure is known. Indeed, one can check that the probability measure

$$\mu(d\boldsymbol{\omega}) = \frac{1}{K}\rho(\boldsymbol{\omega})\, d\boldsymbol{\omega}, \tag{4}$$

$$\rho(\boldsymbol{\omega}) = \sum_{\pi \in S_d} \frac{1}{1+\omega_{\pi(1)}} \frac{1}{1+\omega_{\pi(1)}+\omega_{\pi(2)}} \cdots \frac{1}{1+\omega_{\pi(1)}+\omega_{\pi(2)}+\cdots+\omega_{\pi(d)}}$$

is invariant under T. Here S_d is the group of permutations of d symbols and $K = \int_{\Delta^d} \rho(\boldsymbol{\omega})\, d\boldsymbol{\omega}$. The fact that we have control over the invariant density is very important for our method.

We now explain how T can be used to construct rational approximations. Note that the main reason it can be used for this purpose is that it is expanding and all branches of T^{-1} preserve the set of rational vectors $(\frac{p_1}{q}, \dots, \frac{p_d}{q})$.

First define a map $m : \Delta^d \to \mathbb{N}$ by $m(\boldsymbol{\omega}) = [1/\omega_1]$. Also, let $j(\boldsymbol{\omega}) = i$ where the i^{th} coordinate of $T(\boldsymbol{\omega})$ is $\{1/\omega_1\}$.

Consider the trajectory of a point $\boldsymbol{\omega}_1$ under T, i.e.

$$\boldsymbol{\omega}_1 \xrightarrow{T} \boldsymbol{\omega}_2 \xrightarrow{T} \cdots \xrightarrow{T} \boldsymbol{\omega}_n \xrightarrow{T} \cdots. \tag{5}$$

If the sequence $(m_1, j_1), \dots, (m_n, j_n)$ is known, where

$$m_i = m(T^{i-1}\boldsymbol{\omega}_1), j_i = j(T^{i-1}\boldsymbol{\omega}_1),$$

then for $1 \le i \le n$ it is possible to find the point $\boldsymbol{\omega}_{i-1}$ for which $T(\boldsymbol{\omega}_{i-1}) = \boldsymbol{\omega}_i$. Let $S_{(m_i,j_i)}$ denote the inverse of T on the branch specified by (m_i, j_i). Take the vector $(\frac{0}{1}, \dots, \frac{0}{1})$ as an approximation to $\boldsymbol{\omega}_n$, and consider the image of this vector under $S_{(m_1,j_1)} \circ S_{(m_2,j_2)} \circ \cdots \circ S_{(m_n,j_n)}$. Then we obtain a vector $(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)})$ which is the n^{th} approximation to $\boldsymbol{\omega}_1$.

This procedure leads to the following formal description. For each $m \in \mathbb{N}$, $j = 1, \dots, d$, let us define a matrix $\tilde{A}^{(m,j)} \in GL(d+1, \mathbb{Z})$. The first row of $\tilde{A}^{(m,j)}$ has only two nonzero entries:

$$a_{1,1} = m, a_{1,j+1} = 1.$$

All other rows have only one nonzero entry, which is equal to 1. In fact, $a_{i,i-1} = 1$ for $i = 2, \ldots, j+1$ and $a_{i,i} = 1$ for $i = j+2, \ldots, d+1$. In short,

$$\tilde{A}^{(m,j)} = \begin{pmatrix} m & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}. \tag{6}$$

Now define $A^{(m,j)} = (\tilde{A}^{(m,j)})^t$ where A^t denotes the transpose of A. Define

$$C_n = A^{(m_n,j_n)} \cdots A^{(m_2,j_2)} A^{(m_1,j_1)}, \tag{7}$$

where $m_i = m(T^{i-1}\boldsymbol{\omega}_1)$, $j_i = j(T^{i-1}\boldsymbol{\omega}_1)$. Then the first row of C_n gives the numerators and the denominator of the approximation of the n^{th} step described above. More precisely, if the first row of C_n is $(c_{1,1}^{(n)}, \ldots, c_{1,d}^{(n)}, c_{1,d+1}^{(n)})$ then

$$\left(\frac{p_1(n)}{q(n)}, \ldots, \frac{p_d(n)}{q(n)}\right) = \left(\frac{c_{1,2}^{(n)}}{c_{1,1}^{(n)}}, \ldots, \frac{c_{1,d+1}^{(n)}}{c_{1,1}^{(n)}}\right).$$

The other rows of C_n correspond to approximations which were obtained previously, i.e., approximations $(\frac{p_1(n')}{q(n')}, \ldots, \frac{p_d(n')}{q(n')})$ where $n' < n$. The $d+1$ rows of C_n give $d+1$ points in Δ^d which form a simplex $\Delta(n)$ which contains $\boldsymbol{\omega}_1$. In some sense the whole simplex $\Delta(n)$ is the rational approximation to $\boldsymbol{\omega}_1$.

We introduce a matrix valued function $A : \Delta^d \to GL(d+1, \mathbb{Z})$ defined by $A(\boldsymbol{\omega}) = A^{(m(\boldsymbol{\omega}),j(\boldsymbol{\omega}))}$. Then

$$C_n = A(T^{n-1}\boldsymbol{\omega}_1) \cdots A(T\boldsymbol{\omega}_1) A(\boldsymbol{\omega}_1).$$

Since the approximations are constructed through the product of matrices along the trajectory of an ergodic dynamical system, it is clear that the typical quality of approximations will depend crucially on the corresponding Lyapunov exponents. This will be explained in more detail in section 3.

We finish this section with two definitions which are used in the rest of the paper. For $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we set $\|\boldsymbol{x}\| = \sum_{j=1}^d |x_j|$.

Definition 2. A sequence of rational vectors $\boldsymbol{x}_n = (\frac{p_1(n)}{q(n)}, \ldots, \frac{p_d(n)}{q(n)})$ is said to be *convergent to $\boldsymbol{\omega}$ in the weak sense* if

$$\|\boldsymbol{\omega} - \boldsymbol{x}_n\| \to 0 \quad \text{as} \quad n \to \infty. \tag{8}$$

The sequence is *exponentially convergent to $\boldsymbol{\omega}$ in the weak sense* if there exist constants $K > 0, \alpha > 0$ such that

$$\|\boldsymbol{\omega} - \boldsymbol{x}_n\| \le K q(n)^{-\alpha}. \tag{9}$$

Definition 3. A sequence of rational vectors $\boldsymbol{x}_n = (\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)})$ is said to be *strongly convergent to* $\boldsymbol{\omega}$ if

$$\|q(n)\boldsymbol{\omega} - (p_1(n), \dots, p_d(n))\| \to 0 \quad \text{as} \quad n \to \infty. \tag{10}$$

The sequence is *exponentially strongly convergent to* $\boldsymbol{\omega}$ if there exist constants $K > 0, \alpha > 0$ such that

$$\|q(n)\boldsymbol{\omega} - (p_1(n), \dots, p_d(n))\| \leq K q(n)^{-\alpha}. \tag{11}$$

As we will see later, the denominators $q(n)$ grow exponentially with n for almost all $\boldsymbol{\omega}$, and this explains the word exponential in the above definitions.

In section 4 we prove that for almost all $\boldsymbol{\omega} \in \Delta^d$ the OJPA provides approximations which are exponentially convergent to $\boldsymbol{\omega}$ in the weak sense. The geometrical meaning of convergence in the weak sense and in the strong sense is the following. Weak convergence means that $(q(n), p_1(n), \dots, p_d(n))$ converges to $(1, \omega_1, \dots, \omega_d)$ in the directional sense, i.e. the angle between the two vectors converges to 0 as $n \to \infty$. Strong convergence means that the vector $(q(n), p_1(n), \dots, p_d(n))$ itself converges to the ray

$$\Gamma_{\boldsymbol{\omega}} = \{\lambda(1, \omega_1, \dots, \omega_d) : \lambda > 0\}$$

defined by the vector $(1, \omega_1, \dots, \omega_d)$ (see [6, 18]).

3 Lyapunov exponents and strong convergence

In this section we prove a general theorem which is similar to Theorem 4.1 of [18]. However, our aim is to establish conditions for exponential strong convergence while Lagarias studies exponents which characterise the quality of approximations. We will comment on the connections with Lagarias' Theorem in more detail in the conclusions.

The result which we prove in this section is quite general. It applies not just to the OJPA, but to general approximation schemes based on the product of matrices $A(\boldsymbol{\omega}) \in GL(d+1, \mathbb{Z})$ along the trajectory of a transformation T of $\Omega \subset [0,1]^d$. As in the previous section, we write $C_n(\boldsymbol{\omega}) = A(T^{n-1}\boldsymbol{\omega}) \cdots A(T\boldsymbol{\omega})A(\boldsymbol{\omega})$ and set

$$\boldsymbol{x}_j^{(n)} = \left(\frac{c_{j,2}^{(n)}}{c_{j,1}^{(n)}}, \dots, \frac{c_{j,d+1}^{(n)}}{c_{j,1}^{(n)}} \right).$$

We consider $\boldsymbol{x}_1^{(n)}$ as the n^{th} approximation to $\boldsymbol{\omega}$ and sometimes denote it by

$$\boldsymbol{x}_n = \left(\frac{p_1(n)}{q(n)}, \dots, \frac{p_d(n)}{q(n)} \right)$$

where $q(n) = c_{1,1}^{(n)}$ and $p_i(n) = c_{1,i+1}^{(n)}$.

We will suppose that T and the matrix valued function $A(\boldsymbol{\omega})$ satisfy the following conditions:

(P1) T is ergodic with respect to an invariant probability measure $\mu(d\boldsymbol{\omega})$.

(P2) μ is equivalent to Lebesgue measure.

(P3) The matrices $A(\boldsymbol{\omega})$ are nondegenerate and have non-negative entries for μ-almost all $\boldsymbol{\omega}$.

(P4) $\int_\Omega \log(\max(\|A(\boldsymbol{\omega})\|, 1))\, \mu(d\boldsymbol{\omega}) < \infty$.

(P5) For μ-almost all $\boldsymbol{\omega} \in \Omega$, $a_{11}(\boldsymbol{\omega}) > 0$ and for all $2 \leq i \leq d+1$ there exists $1 \leq k \leq d+1$ such that

$$a_{ij}(\boldsymbol{\omega}) = \begin{cases} 0 & \text{if } j \neq k; \\ 1 & \text{if } j = k. \end{cases}$$

(P6) There exists a matrix $M \in GL(d+1, \mathbb{Z})$ with strictly positive elements such that

$$l(\boldsymbol{\omega} \in \Omega : \exists k > 0 \text{ such that } C_k(\boldsymbol{\omega}) = M) > 0 \tag{12}$$

where l denotes Lebesgue measure on Ω.

(P7) There exists a constant $c_0 > 1$ such that for almost all $\boldsymbol{\omega}$ there exists $n_0(\boldsymbol{\omega})$ such that for all $n \geq n_0(\boldsymbol{\omega})$

$$\max_{1 \leq j \leq d+1} \|\boldsymbol{\omega} - \boldsymbol{x}_j^{(n)}\| \leq c_0^{-n}. \tag{13}$$

Condition (P5) is technical and it holds for most algorithms which are considered in this area. It implies that the denominator $q(n) = c_{1,1}^{(n)}$ of $\boldsymbol{x}_n$ is the largest element of the first column of C_n. As we will see (P5) and (P6) imply that for almost all $\boldsymbol{\omega}$ for large enough n all rows of C_n correspond to some approximation $\boldsymbol{x}_m = \boldsymbol{x}_1^{(m)}$, where $m < n$.

Denote

$$\Omega_{M,k} = \{\boldsymbol{\omega} \in \Omega : C_k(\boldsymbol{\omega}) = M\}.$$

Define $n_{M,k}(\boldsymbol{\omega})$ to be the number of visits to the set $\Omega_{M,k}$ by the trajectory $T^i(\boldsymbol{\omega})$, $0 \leq i \leq n-1$, i.e.

$$n_{M,k}(\boldsymbol{\omega}) = \sum_{i=0}^{n-1} \chi_{M,k}(T^i \boldsymbol{\omega}), \tag{14}$$

where $\chi_{M,k}(\boldsymbol{\omega})$ is the indicator function of the set $\Omega_{M,k}$:

$$\chi_{M,k}(\boldsymbol{\omega}) = \begin{cases} 1 & \text{if } \boldsymbol{\omega} \in \Omega_{M,k}; \\ 0 & \text{if } \boldsymbol{\omega} \notin \Omega_{M,k}. \end{cases}$$

Lemma 1. *Suppose (P1), (P2) and (P6) hold. Then there is $k > 0$ such that for almost all $\boldsymbol{\omega} \in \Omega$*

$$\lim_{n \to \infty} \frac{n_{M,k}(\boldsymbol{\omega})}{n} = \mu(\Omega_{M,k}) > 0. \tag{15}$$

Proof. It follows from (12) that there exists $k > 0$ such that $l(\Omega_{M,k}) > 0$. Hence, by (P2), $\mu(\Omega_{M,k}) > 0$. By Birkhoff's Ergodic Theorem, for almost all $\boldsymbol{\omega} \in \Omega$

$$\lim_{n \to \infty} \frac{n_{M,k}(\boldsymbol{\omega})}{n} = \mu(\Omega_{M,k}) > 0.$$

□

Denote

$$k(\boldsymbol{\omega}) = \min\{k : A(T^{k-1}\boldsymbol{\omega}) \cdots A(T\boldsymbol{\omega})A(\boldsymbol{\omega}) \text{ is strictly positive}\}$$

and set $k(\boldsymbol{\omega}) = +\infty$ if no such k exists.

Corollary 1. *If (P1)-(P3) and (P6) hold then $k(\boldsymbol{\omega})$ is finite for almost all $\boldsymbol{\omega}$.*

Proof. Notice that if $T^i\boldsymbol{\omega} \in \Omega_{M,k}$ then all entries of the matrix

$$C_{k+i}(\boldsymbol{\omega}) = C_k(T^i\boldsymbol{\omega})C_i(\boldsymbol{\omega})$$

are positive. □

If (P1) and (P4) hold then one can define Lyapunov exponents $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d+1}$ in the usual way (see [30]). The following theorem is connected to Theorem 4.1 of [18].

Theorem 1. *Suppose (P1)-(P7) hold. Then:*

(i) the largest Lyapunov exponent λ_1 is greater than 0 and is simple, i.e., it has multiplicity 1;

(ii) for almost all $\boldsymbol{\omega} \in \Omega$,

$$\lim_{n\to\infty} \frac{1}{n} \log q(n) = \lambda_1;$$

(iii) for Lebesgue almost all $\boldsymbol{\omega}$ the sequence of approximations $\boldsymbol{x}_n$ is exponentially strongly convergent to $\boldsymbol{\omega}$ if and only if $\lambda_2 < 0$.

Proof. The proof of (i) is the same as in the proof of Theorem 4.1 of [18]. It also follows from this proof that for almost all $\boldsymbol{\omega}$ the vector $\boldsymbol{e}_1 = (1, 0, \ldots, 0) \in \mathbb{R}^{d+1}$ does not belong to the space $E_2(\boldsymbol{\omega})$ corresponding to the Lyapunov exponents $\lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_{d+1}$. Since the first column of C_n,

$$\begin{pmatrix} c_{1,1}^{(n)} \\ c_{2,1}^{(n)} \\ \vdots \\ c_{d+1,1}^{(n)} \end{pmatrix} = C_n \boldsymbol{e}_1$$

we get

$$\left\| \begin{pmatrix} c_{1,1}^{(n)} \\ c_{2,1}^{(n)} \\ \vdots \\ c_{d+1,1}^{(n)} \end{pmatrix} \right\| = \exp(\lambda_1 n(1 + o(1))). \tag{16}$$

As was mentioned above, property (P5) implies that $q(n) = c_{1,1}^{(n)}$ is the largest element of the first column. Hence, for Lebesgue almost all $\boldsymbol{\omega}$,

$$q(n) = \exp(\lambda_1 n(1 + o(1))). \tag{17}$$

Statement (ii) follows immediately from (17). The following formula was also proved in [18]:

$$\sum_{i=1}^{d+1}\sum_{j=1}^{d} |c_{i,1}^{(n)}\omega_j - c_{i,j+1}^{(n)}| = \exp((\lambda_2 + o(1))n). \tag{18}$$

It follows from (18) that for all i

$$\sum_{j=1}^{d} |c_{i,1}^{(n)}\omega_j - c_{i,j+1}^{(n)}| \le \exp((\lambda_2 + o(1))n). \tag{19}$$

This, together with (17), immediately implies

$$\|q(n)\boldsymbol{\omega} - (p_1(n), \ldots, p_d(n))\| \le q(n)^{\frac{\lambda_2+o(1)}{\lambda_1(1+o(1))}} = q(n)^{\frac{\lambda_2+o(1)}{\lambda_1}}. \tag{20}$$

This implies statement (iii) in the case $\lambda_2 < 0$. Moreover, the constant α in Definition 3 can be taken as $-\frac{\lambda_2}{\lambda_1} - \epsilon$ for any $\epsilon > 0$.

Finally, if $\lambda_2 \ge 0$ then from (18), for at least one i,

$$\sum_{j=1}^{d} |c_{i,1}^{(n)}\omega_j - c_{i,j+1}^{(n)}| = \exp((\lambda_2 + o(1))n). \tag{21}$$

Conditions (P5) and (P6) imply that for n large enough every row of C_n corresponds to some approximation $\boldsymbol{x}_{m(n)}$, $m(n) \le n$. Corollary 1 implies that $m(n) \to \infty$ as $n \to \infty$. Moreover one can show that with probability 1 for n large enough $m(n) \ge \beta n$, $1 > \beta > 0$. Indeed it follows from Lemma 1 that

$$n_{M,k} \ge \frac{3}{4}\mu(\Omega_{M,k})n$$

for n large enough. Denote

$$i_{n,M,k} = \max\{i : 0 \le i \le n-k, T^i\boldsymbol{\omega} \in \Omega_{M,k}\}.$$

It is easy to see that $m(n) > i_{n,M,k}$. Notice that $i_{n,M,k} \ge n_{M,k} - k$. Hence for n large enough

$$m(n) > n_{M,k} - k > \beta n,$$

where $\beta = \mu(\Omega_{M,k})/2$. It follows from (21) that

$$\|q(m(n))(\boldsymbol{x}_{m(n)} - \boldsymbol{\omega})\| = \exp((\lambda_2 + o(1))n).$$

Since $q(m(n)) = \exp(\lambda_1(1 + o(1))m(n))$ one has

$$\|q(m(n))(\boldsymbol{x}_{m(n)} - \boldsymbol{\omega})\| = q(m(n))^{\frac{(\lambda_2+o(1))n}{\lambda_1(1+o(1))m(n)}}. \tag{22}$$

It follows from (22) that with probability 1 there is no exponential strong convergence. Indeed if $\lambda_2 > 0$ then the exponent

$$\delta = \frac{(\lambda_2 + o(1))n}{\lambda_1(1 + o(1))m(n)}$$

is also positive for n large enough. If $\lambda_2 = 0$ then $\delta = o(1)$ since $1 \le \frac{n}{m(n)} \le \frac{1}{\beta}$ for n large enough. □

We finish this section by checking that conditions (P1)-(P7) are satisfied for the OJPA. (P1) follows from [26] (see also [19]). (P2) follows immediately from the formula (4) for the invariant measure μ. (P3) is obvious since $|\det A(\boldsymbol{\omega})| = 1$. In order to check (P4) notice that $\|A(\boldsymbol{\omega})\| = [\frac{1}{\omega_1}] + 1$ where

$$\|A\| = \max_{1\leq j\leq d+1} \sum_{i=1}^{d+1} a_{ij}.$$

Using the explicit formula for μ we can easily check that

$$\int_{\Delta^d} \log\left(\frac{1}{\omega_1}\right) \mu(d\boldsymbol{\omega}) < \infty.$$

(P5) is obvious. Property (P6) is considered in the following lemma.

Lemma 2. *Property (P6) holds for the JPA and the OJPA.*

Proof. Consider the "golden" vector $\boldsymbol{\omega}_G$ which is defined by the properties $T(\boldsymbol{\omega}_G) = \boldsymbol{\omega}_G$, $m(\boldsymbol{\omega}_G) = 1$ and $j(\boldsymbol{\omega}_G) = d$. It is easy to show that $\boldsymbol{\omega}_G = (\omega, \omega^2, \ldots, \omega^d)$ where $0 < \omega < 1$ is the only positive root of the equation $\omega^{d+1} + \omega - 1 = 0$. Notice that in a neighbourhood of $\boldsymbol{\omega}_G$ the transformations corresponding to the JPA and the OJPA coincide. Using the fact that $C_n(\boldsymbol{\omega}_G) = (A^{(1,d)})^n$, one can check that all the elements of $C_{2d}(\boldsymbol{\omega}_G)$ are positive. Take $M = C_{2d}(\boldsymbol{\omega}_G)$. Since for all $\boldsymbol{\omega}$ in a small enough open neighbourhood of $\boldsymbol{\omega}_G$, $C_{2d}(\boldsymbol{\omega}) = C_{2d}(\boldsymbol{\omega}_G) = M$ we have

$$l(\boldsymbol{\omega} \in \Omega : C_{2d}(\boldsymbol{\omega}) = M) > 0.$$

□

In the next section we will prove that property (P7) holds for the OJPA.

4 Exponential convergence in the weak sense

Since $q(n)$ grows exponentially, property (P7) is equivalent to the exponential convergence in the weak sense of $\boldsymbol{x}_j^{(n)}$, $1 \leq j \leq d+1$, to the vector $\boldsymbol{\omega}$. We will show in this section that convergence in the weak sense is a Perron-Frobenius type property which basically follows from the positivity of the matrices

$$C_k(\boldsymbol{\omega}) = A(T^{k-1}\boldsymbol{\omega}) \cdots A(T\boldsymbol{\omega})A(\boldsymbol{\omega})$$

for large k. However, there is a small problem connected to the fact that the matrices $A(\boldsymbol{\omega})$ are only non-negative. This difficulty can be overcome by the use of condition (P6).

Consider n large enough so that the first column of C_n is positive, i.e.,

$$\min_{1\leq j\leq d+1} c_{j,1}^{(n)} > 0.$$

Then the approximations

$$\boldsymbol{x}_i^{(n)} = \frac{1}{c_{i,1}^{(n)}}(c_{i,2}^{(n)}, c_{i,3}^{(n)}, \dots, c_{i,d+1}^{(n)}), \quad 1 \le i \le d+1,$$

are defined. The points $\boldsymbol{x}_i^{(n)} \in I^d$, $1 \le i \le d+1$, are the vertices of the simplex

$$\Delta_n = \{\boldsymbol{y} \in \mathbb{R}^d : \boldsymbol{y} = \sum_{i=1}^{d+1} \alpha_i \boldsymbol{x}_i^{(n)}, \alpha_i \ge 0, \sum_{i=1}^{d+1} \alpha_i = 1\}. \tag{23}$$

Denote

$$\operatorname{diam} \Delta_n = \max_{1 \le i,j \le d+1} \|\boldsymbol{x}_i^{(n)} - \boldsymbol{x}_j^{(n)}\|.$$

Theorem 2. *Suppose that (P1)-(P6) hold. Then there exists a constant* $0 < \gamma < 1$ *such that, for almost all* $\boldsymbol{\omega}$,

$$\operatorname{diam} \Delta_n \le \gamma^n \tag{24}$$

for all n *large enough.*

Proof. The simplex Δ_n is formed by the rows of the matrix

$$C_n(\boldsymbol{\omega}) = A(T^{n-1}\boldsymbol{\omega}) \cdots A(T\boldsymbol{\omega})A(\boldsymbol{\omega}). \tag{25}$$

We write $A_i = A(T^i \boldsymbol{\omega})$, $0 \le i \le n-1$. It follows from Lemma 1 that for almost all $\boldsymbol{\omega}$ and for n large enough $n_{M,k}(\boldsymbol{\omega}) > \beta n$ where $\beta = \mu(\Omega_{M,k})/2$. Define

$$I_n = \{i : 0 \le i \le n-1, T^i(\boldsymbol{\omega}) \in \Omega_{M,k}\}.$$

Then for n large enough, the set I_n contains at least βn elements. Denote $t_1 = \min\{i : i \in I_n\}$ and $t_{l+1} = \min\{i : i \in I_n, i \ge t_l + k\}$ for $l \ge 1$. Suppose there are $n' + 1$ elements in the sequence $\{t_l\}$, $n' \ge [\frac{\beta n}{k}] - 1$. For each t_l, except for the last one, we replace the matrices $A_{t_l+k-1}, \dots, A_{t_l}$ in the product (25) by the single matrix M. Then

$$C_n = B_{m-1} \cdots B_0, \quad m = n - n'(k-1), \tag{26}$$

where each B_s, $0 \le s \le m-1$, is equal either to some matrix A_j, $0 \le j \le n-1$, or to M. Clearly, there are n' matrices M in the product (26). Denote the matrix elements of B_s by $b_s(i,j)$, $1 \le i,j \le d+1$. For each $0 \le s \le m$ consider the index j as a spin variable j_s taking values $1, 2, \dots, d+1$. Define a probability distribution

$$P(j_m, j_{m-1}, \dots, j_0) = \frac{1}{Z} \prod_{s=0}^{m-1} b_s(j_{s+1}, j_s), \tag{27}$$

where

$$Z = \sum_{1 \le j_0, j_1, \dots, j_m \le d+1} \prod_{s=0}^{m-1} b_s(j_{s+1}, j_s). \tag{28}$$

Denote by $P_t(j_t, \dots, j_0 | j_m = j)$ the conditional distribution for the first $t+1$ spins under the condition $j_m = j$. It is easy to see that up to normalization the j^{th} row of $C_n(\boldsymbol{\omega})$ is equal to $P_0(\,\cdot\,| j_m = j)$:

$$P_0(j_0 | j_m = j) = \frac{c^{(n)}_{j,j_0}}{\sum_{i=1}^{d+1} c^{(n)}_{j,i}}, \quad 1 \le j_0 \le d+1. \tag{29}$$

Denote by $0 < d_t(j, j') < 1$ the variational distance between $P_t(\,\cdot\,| j_m = j)$ and $P_t(\,\cdot\,| j_m = j')$:

$$d_t(j, j') = \frac{1}{2} \sum_{1 \le j_t, \dots, j_0 \le d+1} |P_t(j_t, \dots, j_0 | j) - P_t(j_t, \dots, j_0 | j')|. \tag{30}$$

Statements (31) and (32) below are well known (see [9]). For fixed j, j' the sequence $d_t(j, j')$ satisfies

$$d_{t-1}(j, j') \le d_t(j, j'), \quad 1 \le t \le m-1. \tag{31}$$

Suppose now that $B_t = M$. Then there exists a constant $0 < \delta < 1$ such that

$$d_t(j, j') \le \delta d_{t+1}(j, j'). \tag{32}$$

It follows from (31) and (32) that

$$d_0(j, j') \le \delta^{n'} \le \delta^{\frac{\beta n}{k} - 2} = \frac{1}{\delta^2} \delta_1^n, \tag{33}$$

where $\delta_1 = \delta^{\frac{\beta}{k}}$. The vertices of Δ_n correspond to the vectors

$$\boldsymbol{x}^{(n)}_j = \frac{1}{c^{(n)}_{j,1}} (c^{(n)}_{j,2}, \dots, c^{(n)}_{j,d+1}), \quad 1 \le j \le d+1. \tag{34}$$

Consider the $(d+1)$-dimensional vectors

$$\boldsymbol{y}^{(n)}_j = (1, \boldsymbol{x}^{(n)}_j), \quad \boldsymbol{z}^{(n)}_j = \frac{c^{(n)}_{j,1}}{\sum_{i=1}^{d+1} c^{(n)}_{j,i}} \boldsymbol{y}^{(n)}_j. \tag{35}$$

It is easy to see that

$$\boldsymbol{z}^{(n)}_j = (P_0(1|j), P_0(2|j), \dots, P_0(d+1|j)). \tag{36}$$

Hence

$$\|\boldsymbol{z}^{(n)}_j - \boldsymbol{z}^{(n)}_{j'}\| = 2 d_0(j, j') \le \frac{2}{\delta^2} \delta_1^n, \quad 1 \le j, j' \le d+1. \tag{37}$$

Denote $c_j^{(n)} = \sum_{i=1}^{d+1} c_{j,i}^{(n)}$. Since $\boldsymbol{y}_j^{(n)} = \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \boldsymbol{z}_j^{(n)}$ we have

$$\begin{aligned} \|\boldsymbol{y}_j^{(n)} - \boldsymbol{y}_{j'}^{(n)}\| &= \left\| \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \boldsymbol{z}_j^{(n)} - \frac{c_{j'}^{(n)}}{c_{j',1}^{(n)}} \boldsymbol{z}_{j'}^{(n)} \right\| \\ &\leq \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \|\boldsymbol{z}_j^{(n)} - \boldsymbol{z}_{j'}^{(n)}\| + \left| \frac{c_j^{(n)}}{c_{j,1}^{(n)}} - \frac{c_{j'}^{(n)}}{c_{j',1}^{(n)}} \right| \\ &= \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \|\boldsymbol{z}_j^{(n)} - \boldsymbol{z}_{j'}^{(n)}\| + \left| \frac{c_{j',1}^{(n)}}{c_{j'}^{(n)}} - \frac{c_{j,1}^{(n)}}{c_j^{(n)}} \right| \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \frac{c_{j'}^{(n)}}{c_{j',1}^{(n)}} \\ &\leq \frac{c_j^{(n)}}{c_{j,1}^{(n)}} \left(\frac{c_{j'}^{(n)}}{c_{j',1}^{(n)}} + 1 \right) \|\boldsymbol{z}_j^{(n)} - \boldsymbol{z}_{j'}^{(n)}\|. \end{aligned} \tag{38}$$

In the last inequality we have used the following trivial fact:

$$\left| \frac{c_{j,1}^{(n)}}{c_j^{(n)}} - \frac{c_{j',1}^{(n)}}{c_{j'}^{(n)}} \right| \leq \|\boldsymbol{z}_j^{(n)} - \boldsymbol{z}_{j'}^{(n)}\|.$$

It follows from (P5) that $c_{1,1}^{(n)} = \exp(\lambda_1 n(1+o(1)))$. Since $c_1^{(n)} > c_{1,1}^{(n)}$ and $c_1^{(n)} \leq \exp(\lambda_1 n(1+o(1)))$ we conclude that $c_1^{(n)} = \exp(\lambda_1 n(1+o(1)))$. Hence

$$0 < \frac{c_{1,1}^{(n)}}{c_1^{(n)}} = e^{no(1)} < 1.$$

Since

$$\left| \frac{c_{1,1}^{(n)}}{c_1^{(n)}} - \frac{c_{j,1}^{(n)}}{c_j^{(n)}} \right| \leq \|\boldsymbol{z}_1^{(n)} - \boldsymbol{z}_j^{(n)}\| \leq \frac{2}{\delta^2} \delta_1^n, \tag{39}$$

we have

$$0 < \frac{c_{j,1}^{(n)}}{c_j^{(n)}} = e^{no(1)} < 1, \quad 1 \leq j \leq d+1. \tag{40}$$

Using (37), (38) and (40) one has

$$\|\boldsymbol{x}_j^{(n)} - \boldsymbol{x}_{j'}^{(n)}\| = \|\boldsymbol{y}_j^{(n)} - \boldsymbol{y}_{j'}^{(n)}\| \leq e^{no(1)} e^{(\log \delta_1)n} = e^{n(\log \delta_1 + o(1))}. \tag{41}$$

This implies (24) for all $\delta_1 < \gamma < 1$. □

The estimate (24) implies (P7) if $\boldsymbol{\omega} \in \Delta_n$ for all n large enough. As it is proved below, this property holds for all Jacobi-Perron type algorithms. In order to prove it in the general setting one has to impose additional conditions on the transformation T and the matrix valued function $A(\boldsymbol{\omega})$ which will guarantee that $\boldsymbol{\omega} \in \Delta_n$ when Δ_n is defined.

We make the following assumptions about the transformation T (see [27]). Suppose

$$T(\omega_1, \ldots, \omega_d) = \left(\frac{L_1(\boldsymbol{\omega}) + b_1}{L_0(\boldsymbol{\omega}) + b_0}, \ldots, \frac{L_d(\boldsymbol{\omega}) + b_d}{L_0(\boldsymbol{\omega}) + b_0} \right) \tag{42}$$

where $b_0, b_1, \ldots, b_d$ are piecewise constant functions on Ω and $L_0, L_1, \ldots, L_d$ are linear forms on Ω with piecewise constant coefficients, i.e.,

$$L_i(\boldsymbol{\omega}) = \sum_{j=1}^{d} l_{ij}\omega_j, \quad 0 \le i \le d,$$

where l_{ij} are piecewise constant functions. Denote $T(\boldsymbol{\omega}) = (\omega_1', \ldots, \omega_d')$. Clearly, (42) can be written in the form

$$(L_0(\boldsymbol{\omega}) + b_0)\begin{pmatrix} 1 \\ \omega_1' \\ \vdots \\ \omega_d' \end{pmatrix} = D(\boldsymbol{\omega})\begin{pmatrix} 1 \\ \omega_1 \\ \vdots \\ \omega_d \end{pmatrix} \tag{43}$$

where $D(\boldsymbol{\omega}) = (d_{ij})$ is the piecewise constant matrix valued function on Ω given by $d_{ij} = \begin{cases} b_{i-1} & \text{if } j = 1; \\ l_{i-1,j-1} & \text{if } j \neq 1, \end{cases}$ $1 \le i, j \le d+1$.

The matrices $A(\boldsymbol{\omega})$ have to be connected with $D(\boldsymbol{\omega})$ in order to produce good rational approximations. We will formulate this connection as condition (P8):

(P8) T satisfies (42) and $A(\boldsymbol{\omega}) = (D^{-1}(\boldsymbol{\omega}))^t$.

This condition is clearly satisfied for the JPA and the OJPA, since the construction of A is based on the procedure explained above. In both cases $L_0(\boldsymbol{\omega}) = \omega_1$ and $b_0 = 0$. Notice that (P5), (P8) imply $L_0(\boldsymbol{\omega}) + b_0 > 0$.

Let Σ_n denote the closed cone in $\mathbb{R}^{d+1}$ generated by the rows of $C_n(\boldsymbol{\omega})$:

$$\Sigma_n = \{(z_1, \ldots, z_{d+1}) \in \mathbb{R}^{d+1} : (z_1, \ldots, z_{d+1}) = \sum_{j=1}^{d+1} \alpha_j (c_{j,1}^{(n)}, \ldots, c_{j,d+1}^{(n)}), \alpha_j \ge 0\}. \tag{44}$$

Denote $\tilde{\boldsymbol{\omega}} = (1, \boldsymbol{\omega}) = (1, \omega_1, \ldots, \omega_d) \in \mathbb{R}^{d+1}$.

Lemma 3. *Suppose that (P5) and (P8) hold. Then, for almost all $\boldsymbol{\omega} \in \Omega$, $\tilde{\boldsymbol{\omega}} \in \Sigma_n$ for all $n \ge 1$.*

Proof. Denote $\boldsymbol{\omega}^{(i)} = T^i(\boldsymbol{\omega})$ and $\tilde{\boldsymbol{\omega}}^{(i)} = (1, \boldsymbol{\omega}^{(i)})$ for $i \ge 0$. It follows from (43) that for $i \ge 1$

$$\tilde{\boldsymbol{\omega}}^{(i)} A(\boldsymbol{\omega}^{(i-1)}) = \frac{1}{L_0(\boldsymbol{\omega}^{(i-1)}) + b_0}\tilde{\boldsymbol{\omega}}^{(i-1)}.$$

Hence

$$\tilde{\boldsymbol{\omega}}^{(n)} A(\boldsymbol{\omega}^{(n-1)}) \cdots A(\boldsymbol{\omega}^{(0)}) = \prod_{i=0}^{n-1}\left(\frac{1}{L_0(\boldsymbol{\omega}^{(i)}) + b_0}\right)\tilde{\boldsymbol{\omega}}^{(0)}. \tag{45}$$

Denote $\alpha^{(n)}(\boldsymbol{\omega}) = \prod_{k=0}^{n-1}(L_0(\boldsymbol{\omega}^{(k)}) + b_0)$. Then (45) can be rewritten as

$$\alpha^{(n)}(\boldsymbol{\omega})\tilde{\boldsymbol{\omega}}^{(n)} C_n(\boldsymbol{\omega}) = \tilde{\boldsymbol{\omega}}^{(0)}. \tag{46}$$

It follows from (46) that

$$\tilde{\boldsymbol{\omega}} = \tilde{\boldsymbol{\omega}}^{(0)} = \sum_{j=1}^{d+1} \alpha^{(n)}(\boldsymbol{\omega})\tilde{\omega}_j^{(n)}(c_{j,1}^{(n)}, \ldots, c_{j,d+1}^{(n)}), \tag{47}$$

where $\tilde{\omega}_j^{(n)}$ are the components of the vector $\tilde{\boldsymbol{\omega}}^{(n)}$, i.e., $\tilde{\boldsymbol{\omega}}^{(n)} = (\tilde{\omega}_1^{(n)}, \ldots, \tilde{\omega}_{d+1}^{(n)})$. Since $\tilde{\omega}_j^{(n)} \geq 0$, $1 \leq j \leq d+1$, and $\alpha^{(n)}(\boldsymbol{\omega}) > 0$, (47) implies the Lemma. □

The following Corollaries are immediate consequences of Lemma 3.

Corollary 2. *Suppose that (P5) and (P8) hold. If Δ_n is defined, i.e. $\min_{1\leq j\leq d+1} c_{j,1}^{(n)} > 0$, then $\boldsymbol{\omega} \in \Delta_n$.*

Proof. Consider the rays $\Gamma_j = \{(z_1, \ldots, z_{d+1}) = \lambda(c_{j,1}^{(n)}, \ldots, c_{j,d+1}^{(n)}) : \lambda > 0\}$, $1 \leq j \leq d+1$. Notice that the vertices $\boldsymbol{x}_j^{(n)}$ of Δ_n correspond to the points of intersection of Γ_j with the hyperplane $z_1 = 1$. Since $\tilde{\boldsymbol{\omega}} = (1, \boldsymbol{\omega})$ belongs to Σ_n it follows that $\boldsymbol{\omega} \in \Delta_n$. □

Corollary 3. *(P1)-(P6), (P8) imply (P7).*

Proof. Since $\boldsymbol{\omega} \in \Delta_n$, we have for almost all $\boldsymbol{\omega}$

$$\max_{1\leq j\leq d+1} \|\boldsymbol{\omega} - \boldsymbol{x}_j^{(n)}\| \leq \operatorname{diam} \Delta_n \leq \gamma^n, \quad 0 < \gamma < 1.$$

□

Corollary 4. *(P7) holds for the OJPA.*

Remark. It follows from property (P7) that, for almost all $\boldsymbol{\omega}$, the sequence of approximations $\boldsymbol{x}_n$ exponentially converges to $\boldsymbol{\omega}$ in the weak sense. Indeed, since $\|\boldsymbol{\omega} - \boldsymbol{x}_n\| \leq \gamma^n$ and $q(n) = \exp(\lambda_1 n(1 + o(1)))$, we have $\|\boldsymbol{\omega} - \boldsymbol{x}_n\| \leq q(n)^{\frac{\log\gamma}{\lambda_1(1+o(1))}}$, which implies exponential convergence in the weak sense.

5 Analysis of Lyapunov exponents

It follows from Theorem 1 that in order to have exponential strong convergence almost everywhere one has to show that $\lambda_2 < 0$.

It is well known that the calculation of Lyapunov exponents is a hard problem. There are no general formula to calculate them. However, there are several general methods which give rigorous estimates of Lyapunov exponents.

We start with a method which is based on the extension of the matrices A to their action on s-forms. It is well known (see [4, 10]) that the largest Lyapunov exponent corresponding to such an extension is equal to $\lambda_1 + \cdots + \lambda_s$ where $\lambda_1, \ldots, \lambda_s$ are the s largest Lyapunov exponents for $(T, A(\boldsymbol{\omega}))$. Below we will formulate this more precisely for the case $s = 2$.

Consider the linear transformation in $\mathbb{R}^{d+1}$ generated by $A(\boldsymbol{\omega})$, i.e.,

$$\boldsymbol{x} \mapsto A(\boldsymbol{\omega})\boldsymbol{x}.$$

Recall that a 2-form on $\mathbb{R}^{d+1}$ is a bilinear skew-symmetric function on a pair of vectors, $\alpha(\boldsymbol{x}, \boldsymbol{y})$. Denote by $\mathcal{F}$ the $(d+1)d/2$-dimensional linear space of 2-forms on $\mathbb{R}^{d+1}$. Consider the natural basis in $\mathcal{F}$, $e_{i,j} = x_i \wedge x_j$ for $1 \leq i < j \leq d+1$. We can define a linear transformation in $\mathcal{F}$ generated by $A(\boldsymbol{\omega})$, namely

$$\hat{A}(\boldsymbol{\omega})\alpha(\boldsymbol{x}, \boldsymbol{y}) = \alpha(A(\boldsymbol{\omega})\boldsymbol{x}, A(\boldsymbol{\omega})\boldsymbol{y}).$$

We write $\hat{A}_i(\boldsymbol{\omega}) = \hat{A}(T^{i-1}\boldsymbol{\omega})$ and $A_i(\boldsymbol{\omega}) = A(T^{i-1}\boldsymbol{\omega})$. Define

$$\hat{C}_n(\boldsymbol{\omega}) = \hat{A}_1(\boldsymbol{\omega}) \circ \cdots \circ \hat{A}_n(\boldsymbol{\omega}).$$

Clearly,

$$\hat{C}_n(\boldsymbol{\omega})\alpha(\boldsymbol{x}, \boldsymbol{y}) = \alpha(A_n(\boldsymbol{\omega}) \cdots A_1(\boldsymbol{\omega})\boldsymbol{x}, A_n(\boldsymbol{\omega}) \cdots A_1(\boldsymbol{\omega})\boldsymbol{y}).$$

Obviously the operator $\hat{C}_n(\boldsymbol{\omega})$ has the property

$$\hat{C}_{n+m}(\boldsymbol{\omega}) = \hat{C}_n(\boldsymbol{\omega}) \circ \hat{C}_m(T^n\boldsymbol{\omega}). \tag{48}$$

Take an arbitrary norm in $\mathcal{F}$. This norm obviously generates a norm in $\mathcal{L}(\mathcal{F}, \mathcal{F})$. Denote

$$f_n(\boldsymbol{\omega}) = \log \|\hat{C}_n(\boldsymbol{\omega})\|.$$

Property (48) implies that $f_n(\boldsymbol{\omega})$ is a subadditive function, i.e.,

$$f_{n+m}(\boldsymbol{\omega}) \leq f_n(\boldsymbol{\omega}) + f_m(T^n\boldsymbol{\omega}).$$

It follows from (P4) that $f_1 \in L^1(\Omega, d\mu)$. Hence by Kingman's Subadditive Ergodic Theorem, for almost all $\boldsymbol{\omega}$, the following limit exists:

$$\lim_{n\to\infty} \frac{f_n(\boldsymbol{\omega})}{n} = \lim_{n\to\infty} \frac{1}{n} \log \|\hat{C}_n(\boldsymbol{\omega})\| = F.$$

Moreover

$$F = \lim_{k\to\infty} \frac{1}{k} \int_\Omega f_k(\boldsymbol{\omega})\, \mu(d\boldsymbol{\omega})$$

and, for arbitrary $k > 0$,

$$F \leq \frac{1}{k} \int_\Omega f_k(\boldsymbol{\omega})\, \mu(d\boldsymbol{\omega}).$$

Note that F does not depend on the choice of norm on $\mathcal{F}$ since all norms in $\mathcal{L}(\mathcal{F}, \mathcal{F})$ are equivalent.

The following theorem is well known (see [4, 10]). However we will provide a proof.

Theorem 3. *Suppose (P1), (P3) and (P4) hold. Then*

$$F = \lambda_1 + \lambda_2.$$

Proof. Take a measurable orthonormal basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_{d+1}\}$ of $\mathbb{R}^{d+1}$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_n \cdots A_1 \boldsymbol{e}_i\| = \lambda_i.$$

(We have dropped the dependence on $\boldsymbol{\omega}$ in the notation). The existence of such a basis follows from Oseledec's Multiplicative Ergodic Theorem. Here $\{\boldsymbol{e}_2, \dots, \boldsymbol{e}_{d+1}\}$ is a basis of the space

$$E_2 = \{\boldsymbol{x} : \lim_{n\to\infty} \frac{1}{n} \log \|A_n \cdots A_1 \boldsymbol{x}\| \le \lambda_2\}.$$

Fix a norm in $\mathcal{F}$ generated by the norm in $\mathbb{R}^{d+1} \times \mathbb{R}^{d+1}$ given by

$$\|(\boldsymbol{x}, \boldsymbol{y})\| = \max(\|\boldsymbol{x}\|, \|\boldsymbol{y}\|)$$

where $\|\boldsymbol{x}\| = \sqrt{x_1^2 + \cdots + x_{d+1}^2}, \|\boldsymbol{y}\| = \sqrt{y_1^2 + \cdots + y_{d+1}^2}$. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{d+1}$ be arbitrary vectors with $\|\boldsymbol{x}\| = 1, \|\boldsymbol{y}\| = 1$. They can be uniquely written as $\boldsymbol{x} = a\boldsymbol{e}_1 + \boldsymbol{x}_1, \boldsymbol{y} = b\boldsymbol{e}_1 + \boldsymbol{y}_1$ where $\boldsymbol{x}_1, \boldsymbol{y}_1 \in E_2$ and $\|\boldsymbol{x}_1\| \le 1, \|\boldsymbol{y}_1\| \le 1, |a| \le 1, |b| \le 1$. Then, since $\alpha(A_n \cdots A_1 a\boldsymbol{e}_1, A_n \cdots A_1 b\boldsymbol{e}_1) = 0$, we have

$$\begin{aligned}
|\alpha(A_n \cdots A_1 \boldsymbol{x}, A_n \cdots A_1 \boldsymbol{y})| &\le |\alpha(A_n \cdots A_1 a\boldsymbol{e}_1, A_n \cdots A_1 \boldsymbol{y}_1)| \\
&\quad + |\alpha(A_n \cdots A_1 \boldsymbol{x}_1, A_n \cdots A_1 b\boldsymbol{e}_1)| \\
&\quad + |\alpha(A_n \cdots A_1 \boldsymbol{x}_1, A_n \cdots A_1 \boldsymbol{y}_1)| \\
&\le |a|\|\alpha\|\|A_n \cdots A_1 \boldsymbol{e}_1\|\|A_n \cdots A_1 \boldsymbol{y}_1\| \\
&\quad + |b|\|\alpha\|\|A_n \cdots A_1 \boldsymbol{x}_1\|\|A_n \cdots A_1 \boldsymbol{e}_1\| \\
&\quad + \|\alpha\|\|A_n \cdots A_1 \boldsymbol{x}_1\|\|A_n \cdots A_1 \boldsymbol{y}_1\| \\
&\le e^{\lambda_1 n(1+o(1))} e^{\lambda_2 n(1+o(1))}.
\end{aligned}$$

The last inequality is uniform in $\alpha, \boldsymbol{x}, \boldsymbol{y}$ with $\|\boldsymbol{x}\| = 1, \|\boldsymbol{y}\| = 1, \|\alpha\| = 1$. Hence

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \log \|\hat{C}_n\| &= \lim_{n\to\infty} \frac{\sup_{\|\alpha\|=1} \sup_{\|\boldsymbol{x}\|=1} \sup_{\|\boldsymbol{y}\|=1} \log |\alpha(A_n \cdots A_1 \boldsymbol{x}, A_n \cdots A_1 \boldsymbol{y})|}{n} \\
&\le \lim_{n\to\infty} (\lambda_1 + \lambda_2)(1 + o(1)) = \lambda_1 + \lambda_2.
\end{aligned}$$

To prove the inequality from the other side, consider the vectors

$$\boldsymbol{e}_i(n) = A_n \cdots A_1 \boldsymbol{e}_i, \quad 1 \le i \le d+1.$$

Let Π_n be the parallelepiped generated by $(\boldsymbol{e}_1(n), \dots, \boldsymbol{e}_{d+1}(n))$ and let P_n be the parallelogram defined by $\boldsymbol{e}_1(n), \boldsymbol{e}_2(n)$. Denote the $(d+1)$-dimensional volume of Π_n by V_n and the 2-dimensional area of P_n by S_n. It follows from Oseledec's Theorem that

$$V_n = \det(A_n \cdots A_1) = e^{(\lambda_1 + \cdots + \lambda_{d+1})n(1+o(1))}.$$

Since

$$V_n \le S_n \|\boldsymbol{e}_3(n)\| \dots \|\boldsymbol{e}_{d+1}(n)\|,$$

we have

$$S_n \geq \frac{V_n}{\prod_{i=3}^{d+1} \|\boldsymbol{e}_i(n)\|} = \frac{e^{(\lambda_1+\cdots+\lambda_{d+1})n(1+o(1))}}{\prod_{i=3}^{d+1} e^{\lambda_i n(1+o(1))}} = e^{(\lambda_1+\lambda_2)n(1+o(1))}.$$

Notice that

$$\sum_{1\leq i<j\leq d+1} |x_i \wedge x_j(\boldsymbol{e}_1(n), \boldsymbol{e}_2(n))|^2 = S_n^2.$$

Hence for at least one $x_i \wedge x_j$

$$|x_i \wedge x_j(A_n \cdots A_1 \boldsymbol{e}_1, A_n \cdots A_1 \boldsymbol{e}_2)| \geq \frac{S_n}{\sqrt{\frac{(d+1)d}{2}}}.$$

Clearly $\|x_i \wedge x_j\| = 1$. Thus

$$\|\hat{C}_n\| \geq \frac{\sqrt{2}S_n}{\sqrt{d(d+1)}},$$

which implies

$$\lim_{n\to\infty} \frac{\log \|\hat{C}_n\|}{n} \geq \lim_{n\to\infty} \frac{\log e^{(\lambda_1+\lambda_2)n(1+o(1))}}{n} = \lambda_1 + \lambda_2.$$

□

We next show that if (P8) holds then λ_1 can be found explicitly. This fact is also well known (see [14]).

Lemma 4. *Under conditions (P1), (P3)-(P5), (P8),*

$$\lambda_1 = -\int_\Omega \log(L_0(\boldsymbol{\omega}) + b_0)\,\mu(d\boldsymbol{\omega}). \tag{49}$$

Proof. Consider again the sequence of vectors $\boldsymbol{\omega}^{(i)} = T^i\boldsymbol{\omega}$, $i \geq 0$, and the vectors $\tilde{\boldsymbol{\omega}}^{(i)} = (1, \boldsymbol{\omega}^{(i)})$. Formula (45) implies that

$$\prod_{i=0}^{n-1} \frac{1}{L_0(\boldsymbol{\omega}^{(i)}) + b_0} = \sum_{j=1}^{d+1} \tilde{\omega}_j^{(n)} c_{j,1}^{(n)} = c_{1,1}^{(n)} + \sum_{j=2}^{d+1} \omega_j^{(n)} c_{j,1}^{(n)}. \tag{50}$$

Since $c_{1,1}^{(n)} = \max_{1\leq j\leq d+1} c_{j,1}^{(n)} = \exp(\lambda_1 n(1+o(1)))$, and $0 \leq \omega_j^{(n)} \leq 1$, $1 \leq j \leq d+1$, it follows from (50) that

$$\prod_{i=0}^{n-1} \frac{1}{L_0(\boldsymbol{\omega}^{(i)}) + b_0} = \exp(\lambda_1 n(1+o(1))). \tag{51}$$

Properties (P3)-(P5) imply that $\frac{1}{L_0(\boldsymbol{\omega})+b_0} \in L^1(\Omega, d\mu)$. Hence, by Birkhoff's Ergodic Theorem,

$$\prod_{i=0}^{n-1} \frac{1}{L_0(\boldsymbol{\omega}^{(i)}) + b_0} = \exp(\lambda n(1+o(1))) \tag{52}$$

where

$$\lambda = \int_\Omega \log\left(\frac{1}{L_0(\boldsymbol{\omega}) + b_0}\right) \mu(d\boldsymbol{\omega}) = -\int_\Omega \log(L_0(\boldsymbol{\omega}) + b_0)\,\mu(d\boldsymbol{\omega}).$$

Comparing (51) and (52) we get $\lambda_1 = \lambda$. □

Summarising the results of this section we can prove

Theorem 4. *Suppose (P1)-(P6), (P8) hold. Then the algorithm is exponentially strongly convergent almost everywhere if and only if there exists k such that*

$$\frac{1}{k}\int_{\Omega} f_k(\boldsymbol{\omega})\,\mu(d\boldsymbol{\omega}) < \lambda_1 = -\int_{\Omega} \log(L_0(\boldsymbol{\omega}) + b_0)\,\mu(d\boldsymbol{\omega}). \tag{53}$$

Proof. If (53) holds then obviously

$$\lambda_1 + \lambda_2 = F \le \frac{1}{k}\int_{\Omega} f_k(\boldsymbol{\omega})\,\mu(d\boldsymbol{\omega}) < \lambda_1.$$

Hence $\lambda_2 < 0$, which implies exponentially strong convergence.

On the contrary, if

$$\frac{1}{k}\int_{\Omega} f_k(\boldsymbol{\omega})\,\mu(d\boldsymbol{\omega}) \ge -\int_{\Omega} \log(L_0(\boldsymbol{\omega}) + b_0)\,\mu(d\boldsymbol{\omega}),$$

for all k then

$$\lambda_1 + \lambda_2 = F = \lim_{k\to\infty}\frac{1}{k}\int_{\Omega} f_k(\boldsymbol{\omega})\,\mu(d\boldsymbol{\omega}) \ge \lambda_1$$

which implies $\lambda_2 \ge 0$. □

6 Numerical results in dimension 3

In this section we will present numerical results which imply that in dimension 3 the OJPA is exponentially strongly convergent with probability 1. As explained in the introduction, this scheme can in principle be used to prove exponential strong convergence in any dimension.

Basically, one has to find k such that

$$\frac{1}{k}\int_{\Delta^3} \log\|\hat{C}_k(\boldsymbol{\omega})\|\,\mu(d\boldsymbol{\omega}) < -\int_{\Delta^3} \log(\omega_1)\,\mu(d\boldsymbol{\omega}). \tag{54}$$

The notation is that of the previous section applied to the OJPA in the case $d = 3$. Notice that it is enough to prove this inequality for an arbitrary norm in the space of 2-forms. Since $\mu(d\boldsymbol{\omega})$ is known explicitly and $\|\hat{C}_k(\boldsymbol{\omega})\|$ is a well defined function on Δ^3 one can find the integrals in (54) with arbitrary precision.

Lemma 5.

$$-\int_{\Delta^3} \log(\omega_1)\,\mu(d\boldsymbol{\omega}) = 0.4897219015 \pm (0.5 \times 10^{-9}).$$

Proof. The invariant measure is given by

$$\mu(d\boldsymbol{\omega}) = \frac{1}{K}\rho(\omega_1, \omega_2, \omega_3)\,d\omega_1 d\omega_2 d\omega_3$$

where

$$\rho(\omega_1,\omega_2,\omega_3) = \frac{1}{1+\omega_1+\omega_2+\omega_3}\left(\frac{1}{(1+\omega_1)(1+\omega_1+\omega_2)}+\frac{1}{(1+\omega_1)(1+\omega_1+\omega_3)}\right.$$
$$+\frac{1}{(1+\omega_2)(1+\omega_2+\omega_1)}+\frac{1}{(1+\omega_2)(1+\omega_2+\omega_3)}$$
$$\left.+\frac{1}{(1+\omega_3)(1+\omega_3+\omega_1)}+\frac{1}{(1+\omega_3)(1+\omega_3+\omega_2)}\right)$$

and

$$K = \int_{\Delta^3} \rho(\omega_1,\omega_2,\omega_3)\, d\omega_1 d\omega_2 d\omega_3.$$

We have

$$\begin{aligned} -\int_{\Delta^3} \log(\omega_1)\,\mu(d\boldsymbol{\omega}) &= -\frac{1}{K}\int_0^1\int_0^{\omega_1}\int_0^{\omega_2} \log(\omega_1)\rho(\omega_1,\omega_2,\omega_3)\, d\omega_3 d\omega_2 d\omega_1 \\ &= -\frac{1}{0.1668786238} \times -0.08172411697 \\ &= 0.4897219015. \end{aligned}$$

(The calculations were performed using Maple V Release 5). $\square$

Consider the basis $e_{ij} = x_i \wedge x_j$, $1 \le i < j \le 4$ of the space $\mathcal{F}$ of 2-forms on $\mathbb{R}^4$:

$$e_{ij}(\boldsymbol{e}_i,\boldsymbol{e}_j) = 1,\; e_{ij}(\boldsymbol{e}_j,\boldsymbol{e}_i) = -1 \quad \text{and} \quad e_{ij}(\boldsymbol{e}_k,\boldsymbol{e}_l) = 0 \quad \text{if} \quad (k,l) \ne (i,j),(j,i)$$

where $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_4\}$ is the standard basis of $\mathbb{R}^4$. Any 2-form $\alpha \in \mathcal{F}$ can be written as

$$\alpha = \sum_{1\le i<j\le 4} \alpha_{ij} e_{ij}, \quad \alpha_{ij} \in \mathbb{R}.$$

We use the standard norm on $\mathcal{F}$ which is defined by

$$\|\alpha\| = \sum_{1\le i<j\le 4} |\alpha_{ij}|.$$

By looking at the action of $\hat{A}$ on the basis of $\mathcal{F}$ we can obtain a matrix representation of $\hat{A}$. We write

$$\hat{A}e_{kl} = \sum_{1\le i<j\le 4} \beta_{ij} e_{ij}.$$

Note that $\beta_{ij} = \hat{A}e_{kl}(\boldsymbol{e}_i,\boldsymbol{e}_j)$. By the definition of $\hat{A}$, we find

$$\hat{A}e_{kl}(\boldsymbol{e}_i,\boldsymbol{e}_j) = e_{kl}(A\boldsymbol{e}_i, A\boldsymbol{e}_j) = a_{ki}a_{lj} - a_{li}a_{kj}.$$

Thus

$$\beta_{ij} = a_{ki}a_{lj} - a_{li}a_{kj}.$$

This allows us to find an explicit matrix representation of $\hat{A}$, and hence of $\hat{C}_k$.

Lemma 6. *For k=20,*

$$\frac{1}{k}\int_{\Delta^3} \log \|\hat{C}_k(\boldsymbol{\omega})\|\, \mu(d\boldsymbol{\omega}) = 0.44 \pm 0.01. \tag{55}$$

We have found the integral above using a Monte-Carlo integration scheme. Lemmas 5 and 6 indicate that condition (53) is satisfied. By changing the norm in (55), one can probably reduce the value of k for which (53) is satisfied from 20 to 15, but not much further. As it was mentioned in the introduction, the rigorous calculation of (55) will be carried out in [12]. Since Monte-Carlo methods do not provide rigorous bounds we formulate our result as a

"Theorem" 5. *In dimension 3, the OJPA is exponentially strongly convergent almost everywhere.*

7 Conclusions

In this paper we have presented a numerical scheme which in principle can provide a rigorous proof of exponentially strong convergence of the OJPA in arbitrary dimension. The scheme we present is simple from the theoretical point of view, although numerically it can be substantially improved (see [12]). The main advantage of the OJPA in this setting is the existence of a formula for the density of the invariant measure, but this is not the only advantage. As we will explain in [12] the OJPA shares many other nice properties with one dimensional continued fractions.

Our scheme could in principle be made to work for the JPA itself provided we have a good approximation to the invariant measure. At the same time, almost everywhere strong convergence of the JPA, OJPA and other algorithms seems to be a conceptual fact and ideally a proof of this should not rely on the evaluation of concrete numbers. Unfortunately, such a conceptual proof is lacking at present.

Lagarias has considered two exponents η and η^* which characterise the quality of approximations provided by multidimensional continued fraction algorithms [18]. Consider an integer vector $\bar{\boldsymbol{u}} = (q, p_1, \ldots, p_d) \in \mathbb{Z}_+^{d+1}$ and denote by $\boldsymbol{u}$ the corresponding rational vector $(\frac{p_1}{q}, \ldots, \frac{p_d}{q}) \in \mathbb{Q}^d$. The quality of the approximation by $\boldsymbol{u}$ of an irrational vector $\boldsymbol{\omega}$ can be characterised by the exponent

$$\eta(\boldsymbol{u}, \boldsymbol{\omega}) = -\frac{\log \|\boldsymbol{\omega} - \boldsymbol{u}\|}{\log q}. \tag{56}$$

Note that

$$\|\boldsymbol{\omega} - \boldsymbol{u}\| = q^{-\eta(\boldsymbol{u},\boldsymbol{\omega})}. \tag{57}$$

At the n^{th} step of any approximation scheme such as the OJPA one gets a simplex with rational vertices

$$\boldsymbol{u}_i(n) = \left(\frac{p_1^{(i)}(n)}{q^{(i)}(n)}, \ldots, \frac{p_d^{(i)}(n)}{q^{(i)}(n)}\right), \quad 1 \le i \le d+1.$$

Define

$$\eta_n(\boldsymbol{\omega}) = \max_{1\le i\le d+1} \eta(\boldsymbol{u}_i(n),\boldsymbol{\omega}) \quad \text{and} \quad \eta_n^*(\boldsymbol{\omega}) = \min_{1\le i\le d+1} \eta(\boldsymbol{u}_i(n),\boldsymbol{\omega}).$$

We can now define two approximation exponents which characterise the best and the worst approximations of the OJPA or any other similar scheme. The best approximation exponent for $\boldsymbol{\omega}$ given by the OJPA is defined as

$$\eta_{OJP}(\boldsymbol{\omega}) = \limsup_{n\to\infty} \eta_n(\boldsymbol{\omega}).$$

The uniform approximation exponent is defined as

$$\eta_{OJP}^*(\boldsymbol{\omega}) = \liminf_{n\to\infty} \eta_n^*(\boldsymbol{\omega}).$$

Under conditions (P1)-(P4), (P7) plus the additional condition that

$$\int_\Omega k(\boldsymbol{\omega})\, \mu(d\boldsymbol{\omega}) < \infty, \tag{58}$$

Lagarias has shown that $\eta_{OJP}(\boldsymbol{\omega}) = \eta_{OJP}$ and $\eta_{OJP}^*(\boldsymbol{\omega}) = \eta_{OJP}^*$ are constant with probability 1. Moreover,

$$\eta_{OJP}^* = 1 - \frac{\lambda_2}{\lambda_1}.$$

It can be shown that condition (58) holds for the OJPA (see [12]). According to a classical result, for almost all $\boldsymbol{\omega}$

$$\eta_{OJP}(\boldsymbol{\omega}) \le 1 + \frac{1}{d}.$$

Obviously, $\eta_{OJP}^*(\boldsymbol{\omega}) \le \eta_{OJP}(\boldsymbol{\omega})$ which, provided $\lambda_2 < 0$, implies

$$\frac{|\lambda_2|}{\lambda_1} \le \frac{1}{d}.$$

Numerical simulations suggest that for the JPA λ_1 has a limit as $d \to \infty$ and in the case of the OJPA λ_1 scales like $\frac{1}{d}$. Based on our simulations the following conjecture seems to be reasonable:

Conjecture. *For the JPA,*

$$\lambda_1(d) \to \lambda_1^{JP} \quad \textit{as} \quad d \to \infty$$

and

$$\lambda_2(d)d \to \lambda_2^{JP} < 0 \quad \textit{as} \quad d \to \infty.$$

For the OJPA,

$$\lambda_1(d)d \to \lambda_1^{OJP} \quad \textit{as} \quad d \to \infty$$

and

$$\lambda_2(d)d^2 \to \lambda_2^{OJP} < 0 \quad \textit{as} \quad d \to \infty.$$

Numerical simulations give the following values of the exponents λ_1, λ_2 in dimension 3:

$$\lambda_1^{JP}(3) = 1.18,\ \lambda_2^{JP}(3) = -0.23,\ \lambda_1^{OJP}(3) = 0.49,\ \lambda_2^{OJP}(3) = -0.11.$$

Notice that

$$\frac{|\lambda_2^{OJP}(3)|}{\lambda_1^{OJP}(3)} > \frac{|\lambda_2^{JP}(3)|}{\lambda_1^{JP}(3)}.$$

We expect that the same holds in any dimension. It is an interesting question whether or not this is true in the limit as $d \to \infty$, i.e. whether

$$\frac{|\lambda_2^{OJP}|}{\lambda_1^{OJP}} > \frac{|\lambda_2^{JP}|}{\lambda_1^{JP}}.$$

References

[1] V. Baladi and A. Nogueira, Lyapunov exponents for non-classical multidimensional continued fraction algorithms, *Nonlinearity* **9** (1996), 1529-1546.

[2] P. R. Baldwin, A multidimensional continued fraction and some of its statistical properties, *Journal of Statistical Physics* **66** (1992), 1463-1505.

[3] P. R. Baldwin, A convergence exponent for multidimensional continued-fraction algorithms, *Journal of Statistical Physics* **66** (1992), 1507-1526.

[4] G. Benettin, L. Galgani, A. Giorgilli and J. -M. Strelcyn, Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; A method for computing all of them. Part 1: Theory, *Meccanica* **15** (1980), 9-20.

[5] L. Bernstein, *The Jacobi-Perron Algorithm-Its Theory and Application, Lecture Notes in Mathematics* **207**, Springer-Verlag (1971).

[6] A. J. Brentjes, *Multi-dimensional Continued Fraction Algorithms*, Mathematical Centre Tracts **145**, Mathematisch Centrum Amsterdam (1981).

[7] V. Brun, Algorithmes euclidiens pour trois et quatre nombres, in: *13 ième Congre. Math. Scand., Helsinki* (1957), 45-64.

[8] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge Tracts in Mathematics and Mathematical Physics **45**, Cambridge University Press (1957).

[9] R. L. Dobrushin, Conditions for the absence of phase transitions in one-dimensional classical systems, *Math. USSR Sbornik* **22** (1974), 28-48.

[10] J. -P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Mod. Physics* **57** (1985), 617-656.

[11] T. Fujita, S. Ito, M. Keane and M. Ohtsuki, On almost everywhere exponential convergence of the modified Jacobi-Perron algorithm: a corrected proof, *Ergod. Th. and Dyn. Sys.* **16** (1996), 1345-1352.

[12] D. M. Hardcastle and K. Khanin, Convergence of the ordered Jacobi-Perron algorithm, to be published.

[13] D. M. Hardcastle and K. Khanin, Continued fractions and the d-dimensional Gauss transformation, to be published.

[14] S. Ito, M. Keane and M. Ohtsuki, Almost everywhere exponential convergence of the modified Jacobi-Perron algorithm, *Ergod. Th. and Dyn. Sys.* **13** (1993), 319-334.

[15] C. G. J. Jacobi, Allgemeine theorie der kettenbruchähnlichen algorithmen, in welchen jede Zahl aus drei vorhergehenden gebildet wird, *J. Reine Angew. Math* **69** (1868), 29-64.

[16] K. Khanin, Talk at the International Workshop on Dynamical Systems, Porto, August 1992.

[17] D. V. Kosygin, Multidimensional KAM theory from the renormalization group viewpoint, in: *Dynamical Systems and Statistical Mechanics (Ed.: Ya. G. Sinai), Advances in Soviet Mathematics* **3** (1991), 99-129.

[18] J. C. Lagarias, The quality of the Diophantine approximations found by the Jacobi-Perron algorithm and related algorithms, *Mh. Math.* **115** (1993), 299-328.

[19] D. H. Mayer, Approach to equilibrium for locally expanding maps in $\mathbb{R}^k$, *Commun. Math. Phys.* **95** (1984), 1-15.

[20] R. Meester, A simple proof of the exponential convergence of the modified Jacobi-Perron algorithm, preprint (to appear in *Ergod. Th. and Dyn. Sys.*).

[21] R. E. A. C. Paley and H. D. Ursell, Continued fractions in several dimensions, *Proc. Cambridge Philos. Soc* **26** (1930), 127-144.

[22] O. Perron, Grundlagen für eine theorie des Jacobischen kettenbruchalgorithmus, *Math. Ann.* **64** (1907), 1-76.

[23] E. V. Podsypanin, A generalization of the algorithm for continued fractions related to the algorithm of Viggo Brunn, *Zap. Naucn. Sem. Leningrad Otdel. Mat. Inst. Steklov* **67** (1977), 184-194. English translation: *Journal of Soviet Math.* **16** (1981), 885-893.

[24] H. Poincaré, Sur une généralisation des fractions continues, *C. R. Acad. Sci. Paris* **99** (1884), 1014-16.

[25] F. Schweiger, *The Metrical Theory of the Jacobi-Perron Algorithm, Lecture Notes in Mathematics* **334**, Springer-Verlag (1973).

[26] F. Schweiger, A modified Jacobi-Perron algorithm with explicitly given invariant measure, in: *Ergodic Theory, Proceedings Oberwolfach, Germany 1978, Lecture Notes in Mathematics* **729**, 199-202, Springer-Verlag (1979).

[27] F. Schweiger, Invariant measures for maps of continued fraction type, *Journal of Number Theory* **39** (1991), 162-174.

[28] F. Schweiger, *Ergodic Theory of Fibred Systems and Metric Number Theory*, Oxford University Press (1995).

[29] F. Schweiger, The exponent of convergence for the 2-dimensional Jacobi-Perron algorithm, *Proceedings of the Conference on Analytic and Elementary Number Theory in Vienna 1996 (Ed.: W. G. Nowak and J. Schoissengeier)*, 207-213.

[30] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag (1982).