Ergod. Th. & Dynam. Sys. (2021), **41**, 2294–2310
doi:10.1017/etds.2020.49
© *The Author(s) 2020. Published by Cambridge University Press*

Approximating integrals with respect to stationary probability measures of iterated function systems

ITALO CIPRIANO† and NATALIA JURGA‡

† *Facultad de Matemáticas, Pontificia Universidad Católica de Chile (PUC), Avenida Vicuña Mackenna 4860, Santiago, Chile*
(e-mail: icipriano@gmail.com)
‡ *Mathematical Institute, St Andrews, KY16 9SS, UK*
(e-mail: naj1@st-andrews.ac.uk)

(*Received* 8 *July* 2019 *and accepted in revised form* 20 *April* 2020)

Abstract. We study fast approximation of integrals with respect to stationary probability measures associated to iterated function systems on the unit interval. We provide an algorithm for approximating the integrals under certain conditions on the iterated function system and on the function that is being integrated. We apply this technique to estimate Hausdorff moments, Wasserstein distances and Lyapunov exponents of stationary probability measures.

Key words: low-dimensional dynamics, iterated function systems, transfer operator, stationary measures
2020 Mathematics Subject Classification: 37M25, 37E05, 37D20 (Primary); 37D35 (Secondary)

1. *Introduction*

Let $\mathcal{I} := \{1, \ldots, N\}$ for some $N \in \{2, 3, \ldots\}$ and let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ be an iterated function system consisting of Lipschitz contractions $\phi_i : [0, 1] \to [0, 1]$. It is well known that there exists a unique, non-empty and compact set $\Lambda \subset [0, 1]$ such that

$$\Lambda = \bigcup_{i\in\mathcal{I}} \phi_i(\Lambda).$$

This set is called the attractor of Φ. Given a probability vector $\mathbf{p} = (p_1, p_2, \ldots, p_N)$ (meaning that each $0 < p_i < 1$ and $\sum_{i\in\mathcal{I}} p_i = 1$), there exists a unique probability measure $\mu = \mu^{(\Phi,\mathbf{p})}$ such that

$$\int \varphi \, d\mu = \sum_{i\in\mathcal{I}} p_i \int \varphi \circ \phi_i \, d\mu \tag{1}$$

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

for every continuous function $\varphi : [0, 1] \to \mathbb{R}$. This probability measure is called the stationary probability measure associated to Φ and $\mathbf{p}$. Its support is the attractor of Φ.

Iterated function systems were first studied by Hutchinson [**14**] who proved the existence and uniqueness of the attractor and stationary probability measure. Since the late 1980s, the problem of algorithmically estimating stationary probability measures has been a topic of active research; see [**4**, **11**, **12**, **19**, **22**].

In these notes we focus on the related problem of algorithmically estimating the *integrals* of functions with respect to stationary measures. In many areas of analysis, dynamical systems and probability theory, important quantities can be expressed in terms of such integrals; see §6. For example, in fractal geometry, the Hausdorff dimension of the attractor Λ can be bounded or sometimes explicitly calculated in terms of an expression involving the Lyapunov exponent of particular stationary measures; see §6.3. On the other hand, in dynamical systems, equilibrium measures often take the form of a stationary probability measure and therefore it is a natural problem to study the approximation of integrals with respect to them. Apart from some special cases, for instance whenever the stationary measure is absolutely continuous with an explicit known density, it is impossible to analytically calculate these integrals and so one must turn to approximating them. We use an operator-theoretic approach similar to [**15**, Theorem 1], where Jenkinson and Pollicott constructed an algorithm yielding accelerated approximations of integrals of functions with respect to Lebesgue measure. Indeed, our main theorem (Theorem 2.3) can be seen as an analogue of their result. In particular, we provide:

- a sufficient condition on the iterated function system Φ; and
- a class $\mathcal{C}_\Phi$ of functions $g : [0, 1] \to \mathbb{R}$

such that the integral

$$\int_0^1 g \, d\mu$$

can be approximated efficiently, where μ denotes any stationary probability measure associated to Φ. In particular, in Theorem 2.3 we provide an algorithm which produces approximations $\mu_k(g)$ to $\int g \, d\mu$ such that the approximations approach the integral at a super-exponential rate. We demonstrate the performance of this algorithm by using it to estimate Hausdorff moments, Wasserstein distances and Lyapunov exponents.

2. *Main results*

We introduce the following contraction condition motivated by the definition introduced in [**2**, Definition 2.2].

Definition 2.1. A family of maps $\Phi = \{\phi_i\}_{i \in \mathcal{I}}$, $\phi_i : [0, 1] \to [0, 1]$ is called *complex contracting* if there exists a non-empty, bounded, connected and open set $\mathcal{D} \subset \mathbb{C}$ such that $[0, 1] \subset \mathcal{D}$ and:

(1) each ϕ_i extends holomorphically to $\mathcal{D}$;
(2) ϕ_i' is continuous on the boundary of $\mathcal{D}$;
(3) $\sup_{i \in \mathcal{I}} \sup_{z \in \mathcal{D}} |\phi_i'(z)| < 1$.

The significance of the above definition is that if Φ is a complex contracting family of maps then there exists some complex domain $D \subset \mathcal{D}$ such that $[0, 1] \subset D$, each ϕ_i extends

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

holomorphically to D and $\overline{\phi_i(D)} \subset D$; see [**2**, Lemma 2.4]. In this case we say that D is an *admissible domain* for Φ. The existence of such a domain is what is actually required in order to construct our algorithm; however in many cases it is easier to check that a family is complex contracting than to check that such a domain D exists.

Definition 2.2. Given a complex contracting iterated function system $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ on the unit interval, we define the class of functions

$$\mathcal{C}_\Phi := \bigcup_D \{g : [0,1] \to \mathbb{R} : g \text{ has an extension to some bounded holomorphic function on } D\},$$

where the union is taken over all admissible domains for Φ.

In particular, if Φ is complex contracting, then all sufficiently small Euclidean ε-neighbourhoods of $[0, 1]$ are admissible; see [**2**, Remark 2.5]. Therefore, any function which is bounded and holomorphic on some Euclidean ε-neighbourhood of $[0, 1]$ belongs to $\mathcal{C}_\Phi$.

We introduce some notation. Let $\mathcal{I}^n = \{i_1 \cdots i_n : i_j \in \mathcal{I}\}$ and $\phi_{i_1\ldots i_n} := \phi_{i_1} \circ \cdots \circ \phi_{i_n}$ for $n \in \mathbb{N}$. Define $\mathcal{I}^* = \bigcup_{n\in\mathbb{N}} \mathcal{I}^n$. For $\mathbf{i} = i_1 \ldots i_n \in \mathcal{I}^*$ and $1 \leqslant k \leqslant n-1$, we define

$$\sigma^k\mathbf{i} := i_{k+1} \ldots i_n i_1 \ldots i_k,$$
$$\tilde{\sigma}^k\mathbf{i} := i_{k+1} \ldots i_n \quad \text{and}$$
$$p_\mathbf{i} := \prod_{i=1}^{n} p_i.$$

Define $\Pi : \mathcal{I}^{\mathbb{N}} \to \Lambda$ by

$$\Pi(i_1 i_2 \ldots) := \lim_{n\to\infty} \phi_{i_1} \circ \cdots \circ \phi_{i_n}([0,1]).$$

For each $\mathbf{i} \in \mathcal{I}^*$, let $(\mathbf{i})^\infty$ denote the periodic point $(\mathbf{i})^\infty = \mathbf{iii} \ldots \in \mathcal{I}^{\mathbb{N}}$ and let $z_\mathbf{i}$ denote the fixed point $z_\mathbf{i} = \Pi((\mathbf{i})^\infty)$ of $\phi_\mathbf{i}$. Finally, we say that an iterated function system is non-overlapping if $\{\phi_i((0,1))\}_{i\in\mathcal{I}}$ are pairwise disjoint.

The following is our main result.

THEOREM 2.3. *Let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ be a complex contracting iterated function system on the unit interval and $\mathbf{p} = (p_1, \ldots, p_N)$ a probability vector. Given $g \in \mathcal{C}_\Phi$, define*

$$t_m := \sum_{\mathbf{i}\in\mathcal{I}^m} p_\mathbf{i} \frac{1}{1 - \phi'_\mathbf{i}(z_\mathbf{i})}$$

and

$$\tau_m := \sum_{\mathbf{i}\in\mathcal{I}^m} p_\mathbf{i} \frac{g(z_\mathbf{i}) + g(z_{\sigma\mathbf{i}}) + \cdots + g(z_{\sigma^m\mathbf{i}})}{1 - \phi'_\mathbf{i}(z_\mathbf{i})}.$$

Also define

$$\alpha_n := \sum_{l=1}^{n} \frac{(-1)^l}{l!} \sum_{n_1+\cdots+n_l=n} \sum_{j=1}^{l} \frac{\tau_{n_j}}{n_j} \prod_{1\leqslant m\leqslant l, m\neq j} \frac{t_{n_m}}{n_m}$$

and

$$a_n := \sum_{l=1}^{n} \frac{(-1)^l}{l!} \sum_{n_1+\cdots+n_l=n} \prod_{i=1}^{l} \frac{t_{n_i}}{n_i}.$$

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

Then, for

$$\mu_k(g) := \frac{\sum_{n=0}^{k} \alpha_n}{\sum_{n=0}^{k} n a_n}$$

and μ *the stationary probability measure associated to* Φ *and* $\mathbf{p}$*,*

$$\left| \int_0^1 g \, d\mu - \mu_k(g) \right| < C \exp(-\lambda k^2) \tag{2}$$

for some constants $C, \lambda > 0$ *which are independent of* k*.*

Note that while our approximations are non-effective, in some special cases the constants C, λ appearing in (18) can be made explicit. For example, when each ϕ_i takes the form of a linear fractional transformation, that is, $\phi_i(x) = (a_i x + b_i)/(c_i x + d_i)$ for some constants a_i, b_i, c_i, d_i (which in particular includes the important class of self-similar iterated function systems), the constants C and λ can be bounded similarly to [**16**, **17**]. Also note that since the time taken to process n steps of the algorithm is exponential in n, the error decreases super-polynomially fast in time.

Finally, we observe that by using stationarity of the measure μ, Theorem 2.3 can also be used to obtain approximations of integrals of some *piecewise analytic* functions. In particular, suppose that for some $K \in \mathbb{N}$, $\{\phi_{\mathbf{i}}([0, 1])\}_{\mathbf{i} \in \mathcal{I}^K}$ are pairwise disjoint and let $g : [0, 1] \to \mathbb{R}$ be a function (not necessarily continuous) such that each composition function $g \circ \phi_{\mathbf{i}} : [0, 1] \to \mathbb{R}$ belongs to $\mathcal{C}_\Phi$ for every $\mathbf{i} \in \mathcal{I}^K$. By continuous continuation, there is a continuous function $\tilde{g} : [0, 1] \to \mathbb{R}$ with $\tilde{g}|_{\phi_{\mathbf{i}}([0,1])} = g|_{\phi_{\mathbf{i}}([0,1])}$ for every $\mathbf{i} \in \mathcal{I}^K$. Thus, by stationarity,

$$\begin{aligned} \int_0^1 g(x) \, d\mu(x) &= \int_0^1 \tilde{g}(x) \, d\mu(x) \\ &= \sum_{\mathbf{i} \in \mathcal{I}^K} p_{\mathbf{i}} \int_0^1 \tilde{g} \circ \phi_{\mathbf{i}}(x) \, d\mu(x) \\ &= \sum_{\mathbf{i} \in \mathcal{I}^K} p_{\mathbf{i}} \int_0^1 g \circ \phi_{\mathbf{i}}(x) \, d\mu(x); \end{aligned}$$

so, since $\phi_{\mathbf{i}}(D) \subset D$, Theorem 2.3 can be applied to approximate each integral $\int g \circ \phi_{\mathbf{i}} \, d\mu$.

3. *Preliminaries*

3.1. *Trace-class operators, determinants and approximation numbers.* Given a compact operator $L : H \to H$ on a Hilbert space H, its nth *approximation number* is defined as

$$s_n(L) = \inf\{\|L - K\| : \operatorname{rank}(K) \leqslant n - 1\}.$$

A bounded linear operator on a complex separable Hilbert space H is called *trace-class* if $\sum_{n=1}^{\infty} s_n(L) < \infty$. Given a trace-class operator L, the *trace* is defined as

$$\operatorname{tr}(L) = \sum_{n=1}^{\infty} \langle L e_n, e_n \rangle_H,$$

where $\{e_n\}$ is any orthonormal basis and $\langle , \rangle_H$ is the inner product for the Hilbert space H.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

Given a compact operator L, we denote by $\{\lambda_n(L)\}_{n\in\mathbb{N}}$ the monotone decreasing sequence of non-zero eigenvalues of L, listed with algebraic multiplicity. If L is trace-class, then it is compact and its sequence of eigenvalues $\lambda_n(L)$ is absolutely summable.

For a trace-class operator L, the *Fredholm determinant* of L can be defined as

$$\det(\mathrm{Id} - zL) = \prod_{n=0}^{\infty}(1 - z\lambda_n(L)), \tag{3}$$

which is an entire function of z [**21**, Theorem 3.3], so in particular there exist $a_n \in \mathbb{C}$ such that

$$\det(\mathrm{Id} - zL) = \sum_{n=0}^{\infty} a_n z^n.$$

Note that by (3) the roots of $\det(\mathrm{Id} - zL)$ are precisely the reciprocals of the eigenvalues of L, and the degree of each zero is given by the multiplicity of the corresponding eigenvalue. Moreover, each coefficient a_n can be expressed in terms of the traces of L^m for $1 \leqslant m \leqslant n$:

$$a_n = \sum_{m=1}^{n} \frac{(-1)^m}{m!} \sum_{\substack{n_1,\dots,n_m\in\mathbb{N}^m \\ n_1+\cdots+n_m=m}} \prod_{i=1}^{m} \frac{\mathrm{tr}L^{n_i}}{n_i}; \tag{4}$$

for a proof, see for instance [**17**, Proposition 3.2].

On the other hand, by finding the coefficient of z^n in (3), we see that

$$a_n = (-1)^n \sum_{i_1<\cdots<i_n} \lambda_{i_1}(L)\dots\lambda_{i_n}(L)$$

and therefore

$$|a_n| \leqslant \sum_{i_1<\cdots<i_n} |\lambda_{i_1}(L)\dots\lambda_{i_n}(L)|. \tag{5}$$

3.2. *Bergman space.* We will use results of Bandtlow and Jenkinson [**2**, **3**] on operators acting on the Bergman space. Let D be any connected, non-empty and open subset of $\mathbb{C}$. We define the Bergman space $A^2(D)$ by

$$A^2(D) := \left\{ f : D \to \mathbb{C} \text{ holomorphic} : \int_D |f(z)|^2 \, dV(z) < \infty \right\},$$

where V denotes two-dimensional Lebesgue measure, normalized so that the unit ball has unit mass. $A^2(D)$ is a Hilbert space with inner product

$$\langle f, g\rangle_{A^2} = \int_D f(z)\overline{g(z)}\, dV(z).$$

Consider the operator

$$Lf(z) = \sum_{i=1}^{k} w_i(z) f(\psi_i(z)),$$

where ψ_i are holomorphic functions on D that satisfy $\bigcup_{i=1}^{k} \overline{\psi_i(D)} \subset D$ and w_i are holomorphic bounded functions on D. Using the work of Bandtlow and Jenkinson [**2**, **3**], we have the following result.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

PROPOSITION 3.1. *The operator L preserves the Bergman space $A^2(D)$. Moreover:*

(1) *L is a trace-class operator and there exist constants $C, \lambda > 0$ which are independent of N and w_i such that its eigenvalues satisfy*

$$\lambda_n(L) \leqslant C\left(\sum_{i=1}^{k} \|w_i\|_\infty\right) \exp(-\lambda n),$$

where $\|w_i\|_\infty := \sup_{z\in D} |w_i(z)|$;

(2) *the trace of L^n is given by*

$$\operatorname{tr}(L^n) = \sum_{\mathbf{i}\in\mathcal{I}^n} \frac{w_{\mathbf{i}}(z_{\mathbf{i}})}{1 - \psi'_{\mathbf{i}}(z_{\mathbf{i}})},$$

where $z_{\mathbf{i}}$ is the unique fixed point of $\psi_{\mathbf{i}}$ and

$$w_{\mathbf{i}}(z) := w_{i_n}(z) w_{i_{n-1}}(\phi_{i_n} z) \cdots w_{i_1}(\phi_{i_2\ldots i_n} z).$$

Proof. The first part is a special case of [**3**, Theorems 5.9 and 5.13]. The second part follows by [**2**, Theorem 4.2]. □

3.3. *Analytic perturbation theory.* We say that a bounded linear operator L on a Banach space has a *spectral gap* if $L = \lambda P + N$, where P is a rank-one projection (so $P^2 = P$ and $\dim(\operatorname{Im}(P)) = 1$), N is a bounded operator with spectral radius $\rho(N) < |\lambda|$ and $PN = NP = 0$. L does not need to be compact in order to have a spectral gap; however; if the operator L *is* compact and has a simple leading eigenvalue† and no other eigenvalues with the same absolute value, it has a spectral gap.

We can use the standard techniques of perturbation theory [**18**] to relate $\int g\, d\mu$ to the spectral properties of an appropriate operator. The following perturbation theorem is presented in a more general form in [**13**, Theorem 3.8].

THEOREM 3.2. (Analytic perturbation theorem) *Let $\{L_t\}_{t\in\mathbb{C}}$ be a family of bounded linear operators on a Banach space such that $t \mapsto L_t$ is holomorphic and L_0 has a spectral gap. Then there exists an open neighbourhood $U \subset \mathbb{C}$ of zero for which L_t has a spectral gap for all $t \in U$. Moreover, there exist $\lambda(t)$, P_t, N_t which are holomorphic families on U such that:*

(a) $L_t = \lambda(t) P_t + N_t$;

(b) $N_t P_t = P_t N_t = 0$;

(c) *P_t is a bounded rank-one projection and has the form*

$$P_t = \frac{1}{2\pi i} \int_\gamma (s\mathrm{Id} - L_t)^{-1}\, ds$$

for some small circle γ around λ which separates it from the rest of the spectrum of L_0;

(d) *$\rho(N_t) < |\lambda(t)| - \varepsilon$ for some $\varepsilon > 0$ which is independent of t.*

† Throughout the paper we say that an eigenvalue is simple if it is *algebraically simple*, that is, the eigenvalue has a one-dimensional generalized eigenspace.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

4. *Transfer operator*

Let Φ and g satisfy the hypothesis of Theorem 2.3. By assumption, there exists an admissible domain D for Φ such that g has a bounded and holomorphic extension to D. Abusing notation slightly, we also denote this extension by g. Fix a probability vector $\mathbf{p}$, $s \in \mathbb{C}$ and $f \in A^2(D)$. Define the operator $\mathcal{L}_s$ as

$$\mathcal{L}_s f(z) = \sum_{i=1}^{N} p_i \exp(sg(\phi_i(z)))f(\phi_i(z)).$$

Then, by Proposition 3.1, $\mathcal{L}_s f \in A^2(D)$. Observe that the iterates of $\mathcal{L}_s$ are given by

$$\mathcal{L}_s^n f(z) = \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} \exp(s(g(\phi_{\mathbf{i}}z) + g(\phi_{\tilde{\sigma}\mathbf{i}}z) + \cdots + g(\phi_{\tilde{\sigma}^{n-1}\mathbf{i}}z)))f(\phi_{\mathbf{i}}z).$$

The following proposition summarizes some of the spectral properties of $\mathcal{L}_s$.

PROPOSITION 4.1. *$\mathcal{L}_s : A^2(D) \to A^2(D)$ is a trace-class operator with a decreasing sequence of eigenvalues $\{\lambda_n(s)\}_{n\in\mathbb{N}}$. Moreover:*

(1) *there exist constants $C, \lambda > 0$ which are independent of s such that the eigenvalues admit the bound*
$$\lambda_n(\mathcal{L}_s) \leqslant C \exp(-\lambda n^2)$$
for any $s \in \mathbb{C}$ with $|s| \leqslant 1$;

(2) *$\lambda_1(0) = 1$ is a simple eigenvalue of $\mathcal{L}_0$ of maximum modulus;*

(3) *$\lambda_1(s)$ is analytic in s in a neighbourhood of zero; and*

(4) *$\int g\, d\mu = (d/ds)\lambda_1(s)|_{s=0}$.*

Proof. The fact that $\mathcal{L}_s$ is trace-class and the uniform bounds on $\lambda_n(\mathcal{L}_s)$ follow from Proposition 3.1.

Next we prove (2). Clearly 1 is an eigenvalue of $\mathcal{L}_0$ with eigenfunction $\mathbf{1}$. To see that it is geometrically simple, suppose that $f \in A^2(D)$ is a fixed point of $\mathcal{L}_0$ and $f \neq 0$. We will show that f must be a constant function. First, observe that

$$|f(z)| = |\mathcal{L}_0 f(z)| \leqslant \sum_{i=1}^{k} p_i |f \circ \phi_i(z)| \leqslant \sup_{z'\in\bigcup_{i=1}^{k}\overline{\phi_i(D)}} |f(z')|,$$

where the right-hand side is finite because $\bigcup_{i=1}^{k}\overline{\phi_i(D)}$ is a compact subset of D. Therefore,

$$\sup_{z\in D}|f(z)| \leqslant \sup_{z'\in\bigcup_{i=1}^{k}\overline{\phi_i(D)}} |f(z')| = |f(z_0)|$$

for some $z_0 \in \bigcup_{i=1}^{k}\overline{\phi_i(D)}$. By the maximum-modulus principle, f is constant on D. Using the same argument, we can establish that 1 is the only eigenvalue of modulus one.

Therefore, it remains to show that 1 is an algebraically simple eigenvalue. We need to show that $\ker(\mathcal{L}_0 - \mathrm{Id})^2$ is one dimensional (so only consists of the constant functions). For a contradiction, suppose that there exists $f \in A^2(D)$ for which $(\mathcal{L}_0 - \mathrm{Id})f \neq 0$ but $(\mathcal{L}_0 - \mathrm{Id})f \in \ker(\mathcal{L}_0 - \mathrm{Id})$. So, in particular, $(\mathcal{L}_0 - \mathrm{Id})f = c\mathbf{1}$ for some constant c.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

In particular, $c \neq 0$ since $(\mathcal{L}_0 - \mathrm{Id})f \neq 0$ and therefore by replacing f by $c^{-1}f$ we obtain that $(\mathcal{L}_0 - \mathrm{Id})f = \mathbf{1}$, that is, $\mathcal{L}_0 f = \mathbf{1} + f$. By induction, we see that

$$\mathcal{L}_0^n f = n\mathbf{1} + f. \tag{6}$$

On the other hand, define

$$\Gamma^n = \overline{\bigcup_{\mathbf{i}\in\mathcal{I}^n} \phi_{\mathbf{i}}(D)}$$

and define $\Gamma = \bigcap_{n=1}^{\infty} \Gamma^n$. Since Γ^n is a nested sequence of closed subsets of D, Γ is a compact subset of D. For any $z \in \Gamma$,

$$|\mathcal{L}_0^n f(z)| = \left| \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} f(\phi_{\mathbf{i}}(z)) \right| \tag{7}$$

$$\leqslant \sup_{z\in\Gamma} |f(z)|. \tag{8}$$

By (6), $|\mathcal{L}_0^n f(z)| = |n + f(z)| \geqslant n - |f(z)|$, implying that

$$n \leqslant 2 \sup_{z\in\Gamma} |f(z)|,$$

which is clearly a contradiction since f is bounded on Γ.

To see (3), since $\mathcal{L}_0$ is compact, $\lambda_1(0)$ is a simple eigenvalue of maximum modulus and $\{\mathcal{L}_s\}$ is clearly an analytic family of operators in s, we can apply the analytic perturbation theorem (Theorem 3.2) to deduce that $\lambda_1(s)$ is analytic in a neighbourhood U of 0.

Finally, to prove (4), put

$$\mathcal{P}_s = \frac{1}{2\pi i} \int_\gamma (t\mathrm{Id} - \mathcal{L}_s)^{-1}\, dt$$

as in (c) of Theorem 3.2. Parts (a)–(c) of Theorem 3.2 imply that the image of $\mathcal{P}_s$ is an eigenspace for the eigenvalue $\lambda_1(s)$ and that $h_s = \mathcal{P}_s\mathbf{1}$ is an eigenfunction for the eigenvalue $\lambda_1(s)$. Note that $h_0 = \mathbf{1}$. Since $s \mapsto \mathcal{P}_s$ is holomorphic, it immediately follows that $s \mapsto h_s$ is also holomorphic. We write $f_0 = (d/ds)h_s|_{s=0} \in A^2(D)$.

Fix some $z_0 \in \Gamma \cap (0, 1)$. Observe that for each $n > 1$ and $s \in U$,

$$\lambda_1(s)^n h_s(z_0) = (\mathcal{L}_s^n h_s)(z_0) = \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} \exp(s(g(\phi_{\mathbf{i}}z_0) + \cdots + g(\phi_{\tilde{\sigma}^n\mathbf{i}}z_0)))h_s(\phi_{\mathbf{i}}(z_0)).$$

Denote $S_n g(\phi_{\mathbf{i}}z_0) = g(\phi_{\mathbf{i}}z_0) + \cdots + g(\phi_{\tilde{\sigma}^n\mathbf{i}}z_0)$. Differentiating at $s = 0$, we obtain

$$\begin{aligned} n\lambda_1'(0) + f_0(z_0) &= \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} S_n g(\phi_{\mathbf{i}}z_0) + \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} f_0(\phi_{\mathbf{i}}(z_0)) \\ &= \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} S_n g(\phi_{\mathbf{i}}z_0) + \mathcal{L}_0^n f_0(z_0), \end{aligned}$$

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

where we used that $\lambda_1(0) = 1$ and $h_0 = \mathbf{1}$. Therefore, since $|\mathcal{L}_0^n f_0(z_0) - f_0(z_0)| \leqslant \sup_{z\in\Gamma} |f_0(z)| < \infty$,

$$\begin{aligned}\lambda_1'(0) &= \lim_{n\to\infty} \frac{1}{n} \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} S_n g(\phi_{\mathbf{i}} z_0) + \lim_{n\to\infty} \frac{1}{n} (\mathcal{L}_0^n f_0(z_0) - f_0(z_0)) \\ &= \lim_{n\to\infty} \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} g(\phi_{\mathbf{i}} z_0) = \int g \, d\mu.\end{aligned}$$

The fact that $\lim_{n\to\infty} (1/n) \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} S_n g(\phi_{\mathbf{i}} z_0) = \lim_{n\to\infty} \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} g(\phi_{\mathbf{i}} z_0)$ follows because ϕ_i are uniformly contracting and g has bounded derivative on $[0, 1]$; therefore, for any $\varepsilon, \delta > 0$ one can choose N sufficiently large so that for $n \geqslant N$ and all $k \leqslant (1 - \varepsilon)n$,

$$|g(\phi_{\mathbf{i}} z_0) - g(\phi_{\tilde{\sigma}^k \mathbf{i}} z_0)| \leqslant \delta. \qquad \square$$

5. *Determinants and the algorithm*

Since $\mathcal{L}_s$ is a trace-class operator for each $s \in \mathbb{C}$, we can define its determinant function $\det(\mathrm{Id} - z\mathcal{L}_s)$. The following proposition summarizes its properties.

PROPOSITION 5.1. *For all $s \in \mathbb{C}$, $\det(\mathrm{Id} - z\mathcal{L}_s)$ is an entire function of z and can be written in the form*

$$\det(\mathrm{Id} - z\mathcal{L}_s) = \sum_{n=0}^{\infty} b_n(s) z^n$$

for $b_n(s) \in \mathbb{C}$, where $b_0(s) = 1$ for all s and, for $n \geqslant 1$, is defined as

$$b_n(s) = \sum_{m=1}^{n} \frac{(-1)^m}{m!} \sum_{\substack{n_1,\dots,n_m\in\mathbb{N}^m \\ n_1+\cdots+n_m=n}} \prod_{i=1}^{m} \frac{\mathrm{tr}\mathcal{L}_s^{n_i}}{n_i}. \tag{9}$$

Moreover, the trace of $\mathcal{L}_s^n$ is given by

$$\mathrm{tr}(\mathcal{L}_s^n) = \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} \frac{\exp(s(g(z_{\mathbf{i}}) + g(z_{\sigma\mathbf{i}}) + \cdots + g(z_{\sigma^{n-1}\mathbf{i}})))}{1 - \phi_{\mathbf{i}}'(z_{\mathbf{i}})}. \tag{10}$$

Proof. (9) follows from (4). To see (10), notice that by Proposition 3.1(2),

$$\mathrm{tr}(\mathcal{L}_s^n) = \sum_{\mathbf{i}\in\mathcal{I}^n} p_{\mathbf{i}} \frac{\exp(s(g(\phi_{\mathbf{i}}(z_{\mathbf{i}})) + g(\phi_{\tilde{\sigma}\mathbf{i}}(z_{\mathbf{i}})) + \cdots + g(\phi_{\tilde{\sigma}^k\mathbf{i}}(z_{\mathbf{i}}))))}{1 - \phi_{\mathbf{i}}'(z_{\mathbf{i}})}.$$

Since for any $1 \leqslant k \leqslant n$, $\phi_{\tilde{\sigma}^k\mathbf{i}}(z_{\mathbf{i}})$ is a fixed point of $\phi_{\sigma^k\mathbf{i}}$, it follows that $\phi_{\tilde{\sigma}^k\mathbf{i}}(z_{\mathbf{i}}) = z_{\sigma^k\mathbf{i}}$ and therefore (10) follows. $\square$

PROPOSITION 5.2. *Let $\det(\mathrm{Id} - z\mathcal{L}_s) = \sum_{n=0}^{\infty} b_n(s) z^n$ as before. Then there exist constants $C, \lambda > 0$ such that $|b_n(s)|, |b_n'(0)| \leqslant C \exp(-\lambda n^2)$ for all $s \in \mathbb{C}$, $|s| \leqslant 1$. Moreover,*

$$\int g \, d\mu = \lambda_1'(0) = \frac{\sum_{n=0}^{\infty} b_n'(0)}{\sum_{n=0}^{\infty} n b_n(0)}. \tag{11}$$

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

Proof. For the first part, note that by (5) and Proposition 4.1(1),

$$\begin{aligned}|b_n(s)| &\leqslant \sum_{i_1<\cdots<i_n} \lambda_{i_1}(\mathcal{L}_s)\cdots\lambda_{i_n}(\mathcal{L}_s)\\ &\leqslant C^n \sum_{i_1<\cdots<i_n} \exp(-\lambda(i_1+\cdots+i_n))\\ &\leqslant \frac{C^n \exp(-(n(n+1)/2)\lambda)}{\prod_{i=1}^n (1-\exp(-i\lambda))},\end{aligned}$$

where the last line follows by repeated geometric summation. The upper bound on $|b_n(s)|$ follows (for new constants C and λ). By the Cauchy integral formula,

$$b_n'(0) = \frac{1}{2\pi i}\int_{|s|=1} \frac{b_n(s)}{s^2}\,ds$$

therefore

$$|b_n'(0)| \leqslant \sup_{|s|=1} |b_n(s)|,$$

from which we obtain the bound on $|b_n'(0)|$.

Let U be the neighbourhood of zero on which $\lambda_1(s)$ is analytic. Observe that since the zeroes of the determinant $\det(\mathrm{Id} - z\mathcal{L}_s)$ are the reciprocals of the eigenvalues of $\mathcal{L}_s$,

$$\sum_{n=0}^{\infty} b_n(s)\lambda_1(s)^{-n} = 0. \tag{12}$$

Since $|b_n(s)| = O(\exp(-\lambda n^2))$ uniformly on U, we can apply the Cauchy integral formula to deduce that the partial sums $\sum_{n=1}^N b_n'(s) - nb_n(s)\lambda_1'(s)$ converge uniformly on compact subsets of U as $N \to \infty$. Therefore, we can differentiate (12) and take derivatives inside the summation to obtain

$$0 = \frac{d}{ds}\left(\sum_{n=0}^{\infty} b_n(s)\lambda_1(s)^{-n}\right)\Bigg|_{s=0} = \sum_{n=0}^{\infty} b_n'(0) - nb_n(0)\lambda_1'(0). \tag{13}$$

Since $(\lambda_1(0))^{-1}$ is a simple zero of $\det(\mathrm{Id} - z\mathcal{L}_0)$, it follows that $\sum_{n=0}^{\infty} nb_n(0) \neq 0$ and so by rearranging (13) we obtain

$$\lambda_1'(0) = \frac{\sum_{n=0}^{\infty} b_n'(0)}{\sum_{n=0}^{\infty} nb_n(0)},$$

which completes the proof. □

We define the kth approximation

$$\mu_k(g) := \frac{\sum_{n=0}^{k} b_n'(0)}{\sum_{n=0}^{k} nb_n(0)}. \tag{14}$$

Define

$$t_m := \mathrm{tr}(\mathcal{L}_0^m) = \sum_{\mathbf{i}\in\mathcal{I}^m} p_{\mathbf{i}} \frac{1}{1-\phi_{\mathbf{i}}'(z_{\mathbf{i}})}$$

and

$$\tau_m := \frac{d}{ds}\mathrm{tr}(\mathcal{L}_s^m)\Big|_{s=0} = \sum_{\mathbf{i}\in\mathcal{I}^m} p_{\mathbf{i}} \frac{g(z_{\mathbf{i}}) + g(z_{\sigma\mathbf{i}}) + \cdots + g(z_{\sigma^{n-1}\mathbf{i}})}{1-\phi_{\mathbf{i}}'(z_{\mathbf{i}})}.$$

We are now ready to prove the main theorem.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

Proof of Theorem 2.3. By (9), (14) and the definitions of t_n, τ_n, it follows that $\mu_k(g)$ is given by Theorem 2.3. By the bounds on $b_n(0)$ and $b_n'(0)$ given in Proposition 5.2 and (11), we obtain the desired bound on the error $|\int g\, d\mu - \mu_k(g)|$. □

6. *Applications*

In this section we demonstrate Theorem 2.3 by estimating moments and Lyapunov exponents of stationary probability measures and the first Wasserstein distance between stationary probability measures.

6.1. *Moments of stationary probability measures.*

The nth moment of a positive Borel measure μ on $\mathbb{R}$ is defined by

$$\gamma_n = \gamma_n[\mu] := \int_{-\infty}^{\infty} x^n \, d\mu(x), \quad n = 0, 1, \ldots.$$

Moments are particularly important in analysis, probability and statistics. In analysis, the classical moment problem [**1**] is connected with characterizing the image of the map

$$\mathcal{S} : \{\text{positive Borel measures } \mu \text{ on } \mathbb{R} : \mu(\mathbb{R} \setminus [-R, R]) = O(R^{-n}), \forall n\} \to \mathbb{R}^{\mathbb{N}},$$
$$\mu \mapsto (\gamma_n[\mu])_n,$$

which is connected to the problem of the extension of positive functionals. In probability and statistics, the method of moments [**7**] is useful for proving limit theorems and estimating distributions of samples.

Since moments of stationary probability measures can be computed analytically in some special cases (see Lemma 6.2), we can demonstrate the efficiency of the algorithm by using it to approximate moments.

COROLLARY 6.1. *Let* Φ *and* $\mathbf{p}$ *satisfy the assumptions of Theorem 2.3. Then, for each fixed* $m \in \mathbb{N}$, *Theorem 2.3 provides an algorithm with a super-exponential rate of convergence (18) that gives approximations* $w_k \in \mathbb{R}^{m+1}$ *to* $\gamma_{[0,m]} := (\gamma_0, \gamma_1, \ldots, \gamma_m)$ *for* $k = 1, 2, \ldots$.

In the setting where Φ is a non-overlapping iterated function system of similarities, it is possible to obtain the following analytic formula for the moments.

LEMMA 6.2. *Let* $\Phi = (\phi_1, \ldots, \phi_N)$, $\phi_i(x) = \rho_i x + t_i$ *with* $0 < \rho_i < 1$ *such that* $\phi_i(0, 1) \cap \phi_j(0, 1) = \emptyset$ *for every* $i \neq j$ *and* $\mathbf{p} = (p_1, \ldots, p_N)$ *a probability vector.*

Then $\gamma_0 = 1$ *and, for every* $n > 0$,

$$\gamma_n = \frac{\sum_{i=0}^{n-1} \binom{n}{i} \gamma_i \sum_{j=1}^{N} p_j \rho_j^i t_j^{n-i}}{1 - \sum_{j=1}^{N} p_j \rho_j^n}.$$

Proof. Directly from (17) and the definition of Φ. □

We can use Theorem 2.3 to compute approximate values for the moments and compare these with the exact values given by Lemma 6.2.

Example 6.3. Let $\Phi = \{\frac{1}{3}x, \frac{1}{3}x + \frac{2}{3}\}$ and $p = (\frac{1}{3}, \frac{2}{3})$. In Table 1 we compare the approximate values for the moments with the exact values given by the formula.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

TABLE 1. Approximate and actual values for the moments.

Order n	Approx. moment value ($k = 14$)	Actual moment value γ_n	Error
0	1	1	0
1	$0.\bar{6}$	2/3	0
2	0.555555555555555555555555555555555555555555554956	5/9	$< 10^{-48}$
3	0.495726495726495726495726495726495726495726495726296	58/117	$< 10^{-48}$
4	0.455270655270655270655270655270655270655272966	799/1755	$< 10^{-47}$
5	0.424681939833454984970136485288000439515591030742856	54110/127413	$< 10^{-48}$
6	0.400239922384444528966673488818010962533107055240313	662945/1656369	$< 10^{-46}$
7	0.380103465368104155158312752269065347061865131565893	2064430846/5431233951	$< 10^{-46}$
8	0.363174113035119080714419595189042981452860757191787	1213077397297/3340208879865	$< 10^{-46}$
9	0.348713510641572814140979346643070945142831570520821	764170684622650/2191399705783431	$< 10^{-46}$
10	0.336192206948823127238914429058461576988475565523927	16313445679660723325/48524163685162512633	$< 10^{-46}$

On the other hand, when Φ is made up of non-affine contractions, there is no analytic formula available. In the following example we approximate the moments of a stationary probability measure in this setting.

Example 6.4. Let $\Phi = \{1/(x+2),\ 1/(x+4)\}$ and $\mathbf{p} = (\frac{1}{2}, \frac{1}{2})$. We can compute the approximate values for the moments. The results are in Table 2, where all of the digits are stable†.

TABLE 2. Approximate values for the moments.

Order n	Approximate moment value ($k = 13$)
0	1
1	0.330469717526485534080138479518406828981534429410127592033533774023242108 . . .
2	0.119280377696054479896120024958182335914566318048955018663354958988339753 7 . . .
3	0.046120840185791531088127627408927410331302177673749474472036456394013489 1 . . .
4	0.018695667931940428854958531372337814847140638705275614423964219770309791 0 . . .
5	0.007807847977063560924555378015913035121359128312247536049064442681158589 3 . . .
6	0.003320110573231903768685936564676714748591723927498902758538761678623163 9 . . .
7	0.001427182118378503652418281093362766634002528335724754419216502382209445 21 . . .
8	0.000617598307785531412407227175919292282332872096283836068810742529448255746 . . .
9	0.000268421862695922651075727295017088906810341162763633320631290756501684091 . . .
10	0.000117017198360695628842316148053569471108768123919492595293008649132277667 . . .

6.2. *First Wasserstein distance between stationary probability measures.* Fraser [**10**] initiated the study of the interaction between fractal geometry and optimal transportation by proposing the problem of computing the Wasserstein distances between stationary probability measures. The Wasserstein distance is a metrization of the weak-$*$ topology of the space of probability Borel measures on a Polish space with metric d. For $m \in [1, \infty)$, the Wasserstein distance of order m between two probability measures μ and ν is defined by

$$W_m(\mu, \nu) = \inf\{[\mathbb{E}d(X, Y)^m]^{1/m} : \mathrm{law}(X) = \mu, \mathrm{law}(Y) = \nu\}.$$

In general, under no assumption on the iterated function system and associated stationary probability measure, there is no hope of finding an analytic formula. Under some very restrictive assumptions it has been proved in [**6**, **10**] that $W_1(\mu, \nu)$ can be obtained

† We say that a digit is stable if, from empirical observations, it appears to have converged to a stable value.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

explicitly. Using results announced in [**5**] and Theorem 2.3, we can provide an estimate of $W_1(\mu, \nu)$ when $d(x, y) = |x - y|$ and μ, ν are stationary probability measures. The following result was obtained in [**5**].

THEOREM 6.5. *Let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ be a non-overlapping iterated function system of differentiable Lipschitz contractions on the unit interval. Additionally, suppose that $d\phi_i/dx > 0$ for each $i \in \mathcal{I}$ and that $(\mathbf{p}, \mathbf{q})$ is a pair of probability vectors with the property that*

$$\sum_{j=1}^{i}(p_j - q_j)$$

does not change sign for $i = 1, \ldots, N$. Then

$$W_1(\mu^{(\Phi,\mathbf{p})}, \mu^{(\Phi,\mathbf{q})}) = \left|\int x\, d\mu^{(\Phi,\mathbf{p})} - \int x\, d\mu^{(\Phi,\mathbf{q})}\right|.$$

Using the above result and Theorem 2.3, we obtain the following corollary.

COROLLARY 6.6. *Let Φ and $\mathbf{p}, \mathbf{q}$ satisfy the hypothesis of Theorems 6.5 and 2.3. Then Theorem 2.3 provides an algorithm that gives approximations w_k $(k = 1, 2, \ldots)$ to $W_1(\mu^{(\Phi,\mathbf{p})}, \mu^{(\Phi,\mathbf{q})})$ and which converges at a super-exponential rate (18).*

Let $\Phi, \mathbf{p}, \mathbf{q}$ satisfy the assumptions of Corollary 6.6 and additionally suppose that $\phi_i(x) = \rho_i x + t_i$ with $0 < \rho_i < 1$ for each $i = 1, \ldots, N$. In [**5**], it was proved that

$$W_1(\mu^{(\Phi,p)}, \mu^{(\Phi,q)}) = \left|\frac{\sum_{i=1}^{N} p_i t_i}{1 - \sum_{i=1}^{N} p_i \rho_i} - \frac{\sum_{i=1}^{N} q_i t_i}{1 - \sum_{i=1}^{N} q_i \rho_i}\right|. \tag{15}$$

We can compare the approximate values for the Wasserstein distances given by Theorem 6.6 with the exact value given by the equation (15).

Example 6.7. Let $\Phi = \{\frac{1}{3}x, \frac{1}{2}x + \frac{1}{2}\}$ and $p = (\frac{1}{3}, \frac{2}{3})$, $q = (\frac{3}{4}, \frac{1}{4})$. We compare the approximate value of the Wasserstein distance w_k for $k = 8, 9, \ldots, 16$ with the exact value $W_1(\mu^{(\Phi,\mathbf{p})}, \mu^{(\Phi,\mathbf{q})}) = \frac{2}{5}$. The results are in Table 3.

TABLE 3. Approximate values for the Wasserstein distances.

Iteration k	Approx. Wasserstein distance w_k	Error
8	0.39999999995831699720326495795627945349 15	$< 10^{-11}$
9	0.40000000000006338801861112763551950418 8	$< 10^{-13}$
10	0.39999999999999995309731958058477648699 26	$< 10^{-16}$
11	0.40000000000000000001702043753208208516 71	$< 10^{-19}$
12	0.39999999999999999999999695513518752588 63	$< 10^{-23}$
13	0.40000000000000000000000000026945980385 64	$< 10^{-27}$
14	0.39999999999999999999999999999998817522 92	$< 10^{-31}$
15	0.40000000000000000000000000000000000010 67	$< 10^{-36}$
16	0.39999999999999999999999999999999999999999999 72	$< 10^{-41}$

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

On the other hand, when the contractions in Φ are non-affine, an analytic formula is not available.

Example 6.8. Let $\Phi = \{\sin(\pi x/4)/6 + \frac{1}{4}, \sin(\pi x/4)/3 + \frac{2}{3}\}$ and $\mathbf{p} = (\frac{1}{7}, \frac{6}{7})$, $\mathbf{q} = (\frac{1}{2}, \frac{1}{2})$. We can compute the approximate values for the Wasserstein distance between the stationary probability measures associated to $(\Phi, \mathbf{p})$ and $(\Phi, \mathbf{q})$. The results are in Table 4.

TABLE 4. Approximate values for the Wasserstein distances.

Iter. k	Approx. Wasserstein distance w_k
8	0.221045754222800998668664621411128440053831461650047504925067083971083092115702 8
9	0.221045754222800998668664664822322435005592621138097135184119027209843678362625
10	0.221045754222800998668664664822208327921834716857501578546431733679413474804132 8
11	0.221045754222800998668664664822208327924432328476185036762204108148528502575738
12	0.221045754222800998668664664822208327924432327918382319334332352348209033215652
13	0.221045754222800998668664664822208327924432327918382321725464968058605768813675
14	0.221045754222800998668664664822208327924432327918382321725464966008628827356344
15	0.221045754222800998668664664822208327924432327918382321725464966008628827708159

6.3. *Lyapunov exponents.* Let Φ be an iterated function system and μ be the stationary probability measure associated to $(\Phi, \mathbf{p})$. The Lyapunov exponent of Φ with respect to μ can be defined by

$$\chi_\mu := -\int \sum_{i=1}^{N} p_i \log |\phi_i'(x)| \, d\mu(x).$$

The Lyapunov exponent of μ describes the typical rate at which distances are contracted under the action of the maps in the iterated function system, from the point of view of the measure μ. Therefore, Lyapunov exponents play an important role in fractal geometry; for example, the Hausdorff dimension of μ is given in terms of its Lyapunov exponent, which in turn can provide bounds or even precise values for the Hausdorff dimension of the attractor Λ; see [**8**]. When Φ is non-overlapping, the Lyapunov exponent of μ has an equivalent interpretation as the μ-typical rate of expansion of the dynamical system whose branches are given by the inverse images of the maps in Φ and therefore Lyapunov exponents also play an important role in ergodic theory, particularly in thermodynamic formalism. We can use Theorem 2.3 to approximate the Lyapunov exponent of Φ with respect to μ.

COROLLARY 6.9. *Let $\mathbf{p} = (p_1, \ldots, p_N)$ a probability vector and let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$ be a complex contracting iterated function system on the unit interval such that*

$$\max_{i\in\mathcal{I}} \sup_{x\in\phi_i([0,1])} -\log |\phi_i'(x)| \leqslant M < \infty. \tag{16}$$

Then Theorem 2.3 provides an algorithm that gives approximations w_k $(k = 1, 2, \ldots)$ to the Lyapunov exponent of Φ with respect to μ which converges at the super-exponential rate (18).

Proof. Follows by applying Theorem 2.3 to the function $g(x) = -\sum_{i=1}^{N} p_i \log |\phi_i'(x)|$ which belongs to $\mathcal{C}_\Phi$ by (16) and the fact that Φ is complex contracting. □

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

Example 6.10. Let $\Phi = \{\sin(\pi x/4)/6 + \frac{1}{4}, \sin(\pi x/4)/3 + \frac{2}{3}\}$ and $\mathbf{p} = (\frac{1}{3}, \frac{2}{3})$. We compute the approximate values to the Lyapunov exponent of Φ with respect to μ. The results are in Table 5.

TABLE 5. Approximate values for the Lyapunov exponents.

Iter. k	Approx. Lyapunov exponent w_k
8	1.736720814737319877193356682331743524059401611924548203358300048781024559421702833505109933764374558137444965215
9	1.736720814737319877193356690960517046525531648451499154061905482695056109446427575301340179555447731310753282 48
10	1.736720814737319877193356690960513773359889733174090877545124888126670284381330691947369317070262443230792760 22
11	1.736720814737319877193356690960513773360205906010235778257780882278841470132753728359544474157505804029430726 48
12	1.736720814737319877193356690960513773360205906006376079898630983125233520891781830120797706797208143053055117 67
13	1.736720814737319877193356690960513773360205906006376079918873624797626637714646377891168442026187482503544453 94
14	1.736720814737319877193356690960513773360205906006376079918873624791932498453331374280452940254564000080656340808
15	1.736720814737319877193356690960513773360205906006376079918873624791932498455557168841101532523089304936037019
16	1.736720814737319877193356690960513773360205906006376079918873624791932498455557168841091045497369365056896754 19
17	1.736720814737319877193356690960513773360205906006376079918873624791932498455557168841091045497369651485495786 68
18	1.736720814737319877193356690960513773360205906006376079918873624791932498455557168841091045497369651485495786 25

7. *Further remarks*

7.1. *Implementation.* The algorithm was implemented in Python using the free library mpmath†. To estimate the fixed points $z_{\mathbf{i}}$ with arbitrary precision, the command

```
mpmath.findroot
```

was used and to estimate $\phi'_{\mathbf{i}}(z_{\mathbf{i}})$ with arbitrary precision the command

```
mpmath.diff
```

was used.

7.2. *Other applications.* Our methods can be used to estimate other functions in probability theory, such as pointwise estimates of the Fourier–Stieltjes transform of stationary probability measures $\hat{\mu}(x) := \int_0^1 e^{-2\pi i x y}\, d\mu(y)$ and coefficients of the Fourier series of singular measures arising as stationary probability measures.

7.3. *Extension to non-constant weight functions.* Let Φ be a complex contracting iterated function as before. We can extend our results to certain probability measures associated to *non-constant* weight functions.

In particular, suppose that there exist an admissible domain D for Φ and bounded holomorphic functions $\mathbf{p} = \{p_i\}_{i\in\mathcal{I}}$, $p_i : D \to \mathbb{C}$ with the property that:

(1) $\sum_{i\in\mathcal{I}} p_i(z) = 1$ for all $z \in D$;

(2) $p_i([0, 1]) \subset (0, 1)$ for all $i \in \mathcal{I}$.

Then we can obtain an analogue of Theorem 2.3 for the unique probability measure $\mu = \mu^{(\Phi,\mathbf{p})}$ supported on the attractor of Φ with the property that

$$\int \varphi(x)\, d\mu(x) = \sum_{i\in\mathcal{I}} \int p_i(x)\varphi \circ \phi_i(x)\, d\mu(x) \tag{17}$$

† Python library for real and complex floating-point arithmetic with arbitrary precision.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

for every continuous function $\varphi : [0, 1] \to \mathbb{R}$. This probability measure is called the stationary probability measure associated to Φ and $\mathbf{p}$; see [**9**]. It generalizes the stationary probability measure considered for constant weight functions.

By following the proof of Theorem 2.3 with the operator

$$\mathcal{L}_s f(z) = \sum_{i=1}^{N} p_i(z) \exp(sg(\phi_i(z))) f(\phi_i(z)),$$

one can obtain the following analogue of Theorem 2.3.

THEOREM 7.1. *Let $\Phi = \{\phi_i\}_{i\in\mathcal{I}}$, $\mathbf{p} = \{p_i\}_{i\in\mathcal{I}}$ and $\mu = \mu^{\Phi,\mathbf{p}}$ be as above. Given $g \in \mathcal{C}_\Phi$, define*

$$t_m := \sum_{\mathbf{i}\in\mathcal{I}^m} \frac{p_{\mathbf{i}}(z_{\mathbf{i}})}{1 - \phi'_{\mathbf{i}}(z_{\mathbf{i}})}$$

and

$$\tau_m := \sum_{\mathbf{i}\in\mathcal{I}^m} p_{\mathbf{i}}(z_{\mathbf{i}}) \frac{g(z_{\mathbf{i}}) + g(z_{\sigma\mathbf{i}}) + \cdots + g(z_{\sigma^m\mathbf{i}})}{1 - \phi'_{\mathbf{i}}(z_{\mathbf{i}})},$$

where

$$p_{\mathbf{i}}(z) := p_{i_m}(z) p_{i_{m-1}}(\phi_{i_m} z) \cdots p_{i_1}(\phi_{i_2\ldots i_m} z).$$

Define α_n, a_n and $\mu_k(g)$ as in Theorem 2.3. We have

$$\left| \int g(x)\, d\mu(x) - \mu_k(g) \right| < C \exp(-\lambda k^2) \tag{18}$$

for some constants $C, \lambda > 0$ which are independent of k.

Acknowledgements. IC would like to express his gratitude to Felipe Serrano and Franziska Schlösser for suggestions around the effective implementation of the algorithms. IC was partially supported by CONICYT PIA ACT172001. NJ was financially supported by the Leverhulme Trust (Research Project Grant No. RPG-2016-194) and by the EPSRC (Standard Grant EP/R015104/1). Both authors would like to thank Ian Morris for helpful discussions and suggestions.

REFERENCES

[**1**] N. I. Akhiezer. *The Classical Moment Problem and Some Related Questions in Analysis*. Hafner, New York, 1965. Translated by N. Kemmer.
[**2**] O. F. Bandtlow and O. Jenkinson. On the Ruelle eigenvalue sequence. *Ergod. Th. & Dynam. Sys.* **28**(6) (2008), 1701–1711.
[**3**] O. F. Bandtlow and O. Jenkinson. Explicit eigenvalue estimates for transfer operators acting on spaces of holomorphic functions. *Adv. Math.* **218**(3) (2008), 902–925.
[**4**] M. Blank. Perron–Frobenius spectrum for random maps and its approximation. *Mosc. Math. J.* **1**(3) (2001), 315–344, 470.
[**5**] I. Cipriano. The Wasserstein distance between stationary measures. *Preprint*, 2018, arXiv:1611.00092v3.
[**6**] I. Cipriano and M. Pollicott. Stationary measures associated to analytic iterated function schemes. *Math. Nachr.* **291**(7) (2018), 1049–1054.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press

[**7**] P. Diaconis. Application of the method of moments in probability and statistics. *Moments in Mathematics (San Antonio, TX, 1987) (Proceedings of Symposia in Applied Mathematics, 37)*. American Mathematical Society, Providence, RI, 1987, pp. 125–142.
[**8**] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*, 3rd edn. John Wiley & Sons, Chichester, 2014.
[**9**] A. H. Fan and K.-S. Lau. Iterated function system and Ruelle operator. *J. Math. Anal. Appl.* **231** (1999), 319–344.
[**10**] J. Fraser. First and second moments for self-couplings and Wasserstein distances. *Math. Nachr.* **288** (2015), 2028–2041.
[**11**] G. Froyland. Ulam's method for random interval maps. *Nonlinearity* **12** (1999), 1029–1052.
[**12**] S. Galatolo, M. Monge and I. Nisoli. Rigorous approximation of stationary measures and convergence to equilibrium for iterated function systems. *J. Phys. A* **49** (2016), 274001.
[**13**] H. Hennion and L. Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness.* Vol. 1766. Springer Science and Business Media, Berlin, 2001.
[**14**] J. Hutchinson. Fractals and self similarity. *Indiana Univ. Math. J.* **30** (1981), 713–747.
[**15**] O. Jenkinson and M. Pollicott. A dynamical approach to accelerating numerical integration with equidistributed points. *Proc. Steklov Inst. Math.* **256**(1) (2007), 275–289.
[**16**] O. Jenkinson and M. Pollicott. Rigorous effective bounds on the Hausdorff dimension of continued fraction Cantor sets: a hundred decimal digits for the dimension of E2. *Adv. Math.* **325** (2018), 87–115.
[**17**] N. Jurga and I. Morris. Effective estimates on the top Lyapunov exponent for random matrix products. *Nonlinearity* **32**(11) (2019), 4117–4146.
[**18**] T. Kato. *Perturbation Theory for Linear Operators (Classics in Mathematics)*. Reprint of the 1980 edn. Springer, Berlin, 1995.
[**19**] A. Öberg. Algorithms for approximation of invariant measures for IFS. *Manuscripta Math.* **116** (2005), 31–55.
[**20**] D. Ruelle. Zeta-functions for expanding maps and Anosov flows. *Invent. Math.* **34**(3) (1976), 231–242.
[**21**] B. Simon. *Trace Ideals and their Applications.* Vol. 120. American Mathematical Society, Providence, RI, 2010.
[**22**] R. Strichartz, A. Taylor and T. Zhang. Densities of self-similar measures on the line. *Exp. Math.* **4** (995), 101–128.

https://doi.org/10.1017/etds.2020.49 Published online by Cambridge University Press