Adv. Appl. Prob. **27**, 800–820 (1995)
Printed in N. Ireland

SOLVING THE INVERSE PROBLEM FOR MEASURES USING ITERATED FUNCTION SYSTEMS: A NEW APPROACH

B. FORTE* AND
E. R. VRSCAY,** *University of Waterloo*

Abstract

We present a systematic method of approximating, to an arbitrary accuracy, a probability measure μ on $x = [0, 1]^q$, $q \geqq 1$, with invariant measures for iterated function systems by matching its moments. There are two novel features in our treatment. 1. An infinite set of fixed affine contraction maps on X, $\mathcal{W} = \{w_1, w_2, \cdots\}$, subject to an '$\epsilon$-contractivity' condition, is employed. Thus, only an optimization over the associated probabilities p_i is required. 2. We prove a *collage theorem for moments* which reduces the moment matching problem to that of minimizing the collage distance between moment vectors. The minimization procedure is a standard quadratic programming problem in the p_i which can be solved in a finite number of steps. Some numerical calculations for the approximation of measures on $[0, 1]$ are presented.

COLLAGE DISTANCE; QUADRATIC PROGRAMMING; DATA COMPRESSION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 28A80
SECONDARY 41A20; 58A40

1. Introduction

This paper is concerned with the approximation of probability measures on a compact metric space X by invariant measures for iterated function systems (IFS): systems of contraction mappings on X, $\mathbf{w} = \{w_1, w_2, \cdots, w_N\}$, with associated probabilities $\mathbf{p} = \{p_1, p_2, \cdots, p_N\}$, introduced in [19] and developed further in [4] and [5]. The approximation of measures and functions with IFS and related methods has received much interest, especially in the context of *data compression.* As reported in the case of image processing [21], it is desirable to be able to represent a target measure or function with a rather small number of IFS parameters, thus achieving a large 'compression factor'.

The *inverse problem of measure construction* using IFS may be posed as follows: Let (X, d) denote a compact metric space and $\mathcal{M}(X)$ the set of probability measures on $\mathcal{B}(X)$, the σ-algebra of Borel subsets of X. Then, given a *target*

Received 19 April 1993; revision received 19 April 1994.
* Present address: Facoltà di Scienze MM. FF. e NN. a Cà Vignal, Università degli Studi di Verona, Strada Le Grazie, 37134 Verona, Italy.
** Postal address: Department of Applied Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

measure $\nu \in \mathcal{M}(X)$ and an $\epsilon > 0$, find an IFS $(\boldsymbol{w}, \boldsymbol{p})$ whose invariant measure μ satisfies $d_{\mathrm{H}}(\mu, \nu) < \epsilon$, where d_{H} denotes the Hutchinson metric (defined in Equation (2.1) below).

Most of the work on the inverse problem with IFS (see the papers cited in the list of references, for example) has been based on a knowledge of the moments of the target measure ν. Some form of 'moment matching' is applied, roughly as follows. Given a target measure ν (with $X \subset \mathbb{R}$ for simplicity) with moments $h_k \equiv \int x^k \, d\nu$, $k = 1, 2, 3, \cdots$, find an IFS invariant measure μ whose respective moments $g_k \equiv \int x^k \, d\mu$ are 'close' to the h_k. In practical applications, moment matching is performed on a finite sequence of moments. For example, given an $M > 0$, one minimizes the distance between the vectors $(g_1, g_2, \cdots, g_M)$ and $(h_1, h_2, \cdots, h_M)$. In the case of an IFS with affine maps, the moments g_k of its invariant measure μ may be computed recursively from the coefficients of the affine contraction maps as well as the associated probabilities. Hence moment matching becomes an optimization problem in terms of the IFS parameters.

The method described in this paper yields a systematic algorithm to approximate measures with IFS invariant measures to arbitrary accuracy. Our method differs from previous efforts in two significant aspects:

1. We first begin with an infinite set $\mathcal{W} = \{w_1, w_2, \cdots\}$ of *fixed* affine contraction maps $w_i : X \to X$ which must satisfy an 'ϵ-contractivity condition'. As such, we consider the w_i to form a basis for the representation of compact subsets of X. From this set we construct sequences of N-map IFS with probabilities $(\boldsymbol{w}^N, \boldsymbol{p}^N)$. For each such IFS, only an optimization over the probabilities p_i^N, $i = 1, 2, \cdots, N$ is required. The probabilities p_i can be loosely considered as 'Fourier coefficients' of the basis functions w_i.

2. The moment matching is accomplished by means of a *collage theorem for moments*. This is in contrast to a minimization of distances between moment vectors of respectively, target and approximating measures as was done, for example, in [24]–[26]. Since the IFS maps are fixed, the minimization of the collage distance in moment space need only be performed with respect to the probabilities p_i. Moreover, the squared collage distance between moment vectors is quadratic in the p_i and the minimization becomes a quadratic programming problem which can be numerically solved in a finite number of steps. In many cases, the minimum collage distance is achieved on a boundary point of the simplex $\Pi^N = \{(p_1, \cdots, p_N) \mid \sum_{i=1}^N p_i = 1\}$, which means that one or more of the p_i are zero. In such cases, superfluous maps w_i are essentially eliminated from the set $\mathcal{W}$. A density theorem ensures that as $N \to \infty$, the collage distance in moment space tends to zero.

The layout of this paper is as follows. In Section 2, after a brief glossary of notation, we discuss affine IFS and introduce the idea of an infinite set of contraction maps which satisfy the ϵ-contractivity condition mentioned above. In Section 3, we derive the collage theorem for moments and then prove that the method can be used to approximate measures to arbitrary accuracy. Section 4 contains some applications and numerical computations. In Section 5, a few final remarks are made.

2. Iterated function systems and their invariant measures

2.1. *Notation.* In this paper, the following notation will be employed:

(X, d) a compact metric space. (In applications, where X is the 'base space' of the IFS, X will be a compact subset of $\mathbb{R}^n$, e.g. $[0,1]$, $[0,1]^2$.)

$C(X)$ $=\{f: X \to \mathbb{R}, f \text{ is continuous}\}$.

$\mathrm{Lip}(X)$ $=\{f: X \to \mathbb{R}, |f(x)-f(y)| \leqq d(x, y), x, y \in X\}$.

$\mathrm{Con}(X)$ $=\{w: X \to X, d(w(x), w(y)) \leqq cd(x, y)$, for some $c \in [0, 1), \forall x, y \in X\}$: the set of contraction maps on X. We shall refer to c as the *contractivity factor* of w.

$\mathscr{H}(X)$ the set of non-empty compact subsets of X.

h Hausdorff metric on $\mathscr{H}(X)$: Let the distance between a point $x \in X$ and a set $A \in \mathscr{A}(X)$ be given by

$$d(x, A) = \inf_{y \in A} d(x, y).$$

Then for $A, B \in \mathscr{H}(X)$,

$$h(A, B) = \max\left[\sup_{x \in A} d(x, B), \sup_{y \in B} d(y, A)\right].$$

$(\mathscr{H}(X), h)$ is a complete metric space.

$\mathscr{M}(X)$ the set of probability measures on $\mathscr{B}(X)$, the σ-algebra of Borel subsets of X.

d_{H} a metric on $\mathscr{M}(X)$, often referred to in the IFS literature as the *Hutchinson metric* due to its use in [19]:

$$d_{\mathrm{H}}(\mu, \nu) = \sup_{f \in \mathrm{Lip}(X)} \left[\int_X f\, d\mu - \int_X f\, d\nu\right], \qquad \mu, \nu \in \mathscr{M}(X) \tag{2.1}$$

$(\mathscr{M}(X), d_{\mathrm{H}})$ is a complete metric space.

2.2. *Affine* IFS *and infinite sets of contraction maps with* ϵ*-contractivity.* We shall let $(\boldsymbol{w}, \boldsymbol{p})$ denote an N-map contractive IFS on X with probabilities, that is, a set of N contraction maps, $\boldsymbol{w} = (w_1, w_2, \cdots, w_N)$, $w_i \in \mathrm{Con}(X)$, with associated probabilities $\boldsymbol{p} = (p_1, p_2, \cdots, p_N)$, $p_i \geqq 0$, $i = 1, 2, \cdots, N$, and $\sum_{i=1}^N p_i = 1$. The contractivity factor of the IFS is defined as

$$c = \max_{1 \leqq i \leqq N} c_i < 1. \tag{2.2}$$

As usual [4], [5], [19], define a set-valued mapping $\hat{\boldsymbol{w}}: \mathscr{H}(X) \to \mathscr{H}(X)$ as follows.

For a subset $S \in \mathscr{H}(X)$ denote $w_i(S) = \{w_i(x), x \in S\}$, $i = 1, 2, \cdots, N$ and let the action of $\hat{\boldsymbol{w}}$ on $\mathscr{H}(X)$ be given by

$$\hat{\boldsymbol{w}}(S) \equiv \bigcup_{i=1}^{N} w_i(S). \tag{2.3}$$

Then there exists a unique compact set $A \in \mathscr{H}(X)$, the *attractor* of $\boldsymbol{w}$ (independent of $\boldsymbol{p}$), such that

$$A = \hat{\boldsymbol{w}}(A) = \bigcup_{i=1}^{N} w_i(A) \tag{2.4}$$

and $h(\hat{\boldsymbol{w}}^n(S), A) \to 0$ as $n \to \infty$ for all $S \in \mathscr{H}(X)$. Now define an operator $M: \mathscr{M}(X) \to \mathscr{M}(X)$ (often called the 'Markov operator') as follows. For $\mu \in \mathscr{M}(X)$, let

$$M\mu = \sum_{i=1}^{N} p_i \mu \circ w_i^{-1}. \tag{2.5}$$

In [19] it was shown that M is a contraction mapping on $(\mathscr{M}(X), d_H)$: For all $\mu, \nu \in \mathscr{M}(X)$, $d_H(M\mu, M\nu) \leqq c d_H(\mu, \nu)$. Thus there exists a unique measure $\bar{\mu} \in \mathscr{M}(X)$, the *invariant measure* of the IFS, for which $M\bar{\mu} = \bar{\mu}$.

In this paper, we consider only IFS with affine maps on X. For example, on $\mathbb{R}$ these maps have the general form

$$w_i(x) = s_i x + a_i, \qquad |s_i| < 1, \qquad s_i, a_i \in \mathbb{R}, \qquad i = 1, 2, \cdots, N. \tag{2.6}$$

That there is no loss of generality in this restriction is shown by the following theorem, proved in [10].

Theorem 2.1. Let (X, d) denote a compact metric space and $\mathscr{M}_{\mathrm{AIFS}}(X) \subset \mathscr{M}(X)$ the subset of invariant measures of affine IFS on X. Then $\mathscr{M}_{\mathrm{AIFS}}(X)$ is dense in $(\mathscr{M}(X), d_H)$.

Theorem 2.1 is a rather trivial consequence of the following result.

Proposition 2.2 [23]. Let (X, d) be a compact metric space and $\mathscr{M}_f(X) = \{\mu \in \mathscr{M}(X) \mid \mu$ has finite support$\}$. Then $\mathscr{M}_f(X)$ is dense in $\mathscr{M}(X)$.

In our formal solution to the inverse problem, we shall be constructing sequences of N-map IFS with probabilities, denoted as $(\boldsymbol{w}^N, \boldsymbol{p}^N)$, with $N \to \infty$, where the IFS maps in $\boldsymbol{w}^N$ are chosen from a fixed, infinite set $\mathscr{W}$ of contraction maps. A condition must be placed on this set, according to the following definition.

Definition 2.3. An infinite set of contraction maps $\mathscr{W} = \{w_1, w_2, \cdots\}$, $w_i \in \mathrm{Con}(X)$ is said to satisfy an ϵ-contractivity condition on X if:
for each $x \in X$ and any $\epsilon > 0$, there exists an $i^* \in \{1, 2, \cdots\}$ such that

$w_{i^*}(X) \subset N_\epsilon(x)$, where $N_\epsilon(x) = \{y \in X \mid d(x, y) < \epsilon\}$ denotes the ϵ-neighbourhood of x.

If $\mathcal{W}$ satisfies the ϵ-contractivity condition on X, then $\inf_{1 \leq i < \infty} c_i = 0$. As a result, $\mathcal{W}$ can provide N-map IFS with arbitrarily small degrees of refinement on (X, d). A useful set of affine IFS maps on $X = [0, 1]$ satisfying the ϵ-contractivity condition is given by the following 'wavelet-type' basis functions (it is convenient to use two indices):

$$w_{ij}(x) = \frac{1}{2^i}[x + j - 1], \qquad i = 1, 2, \cdots, \quad j = 1, 2, \cdots, 2^i. \tag{2.7}$$

For each $i^* \geqq 1$, the set of maps $\{w_{i^*j}, j = 1, 2, \cdots, 2^{i^*}\}$ provides 2^{-i^*}-contractions of $[0, 1]$ in an obvious way. Another possible set of functions is given by

$$w_{ij}(x) = \frac{1}{i}[x + j - 1], \qquad i = 2, \cdots, \quad j = 2, \cdots, i.$$

In either of the above cases, for a given $i^* \geqq 1$, the sets $\{w_{i^*j}(X), j = 1, 2, \cdots, j_{\max}\}$ overlap with each other only at single points. This is not necessary and, in special cases, it might be advantageous to let these sets overlap on subintervals of X.

2.2. *Moment relations for affine IFS.* A primary motivation for the use of affine IFS maps is the simplicity of relations involving moments of probability measures. Given an N-map IFS $(\boldsymbol{w}, \boldsymbol{p})$ with associated Markov operator M, let $\mu \in \mathcal{M}(X)$ and $\nu = M\mu$. Then from Equation (2.5), for any continuous function $f: X \to \mathbb{R}$,

$$\begin{aligned} \int_X f(x)\, d\nu(x) &= \int_X f(x)\, d(M\mu)(x) \\ &= \sum_{i=1}^N p_i \int_X (f \circ w_i)(x)\, d\mu(x). \end{aligned} \tag{2.8}$$

For the remainder of this section, we consider only $X \subset \mathbb{R}$. Let the moments of μ and ν be denoted by

$$g_n = \int_X x^n\, d\mu, \qquad h_n = \int_X x^n\, d\nu, \qquad n = 0, 1, 2, \cdots, \tag{2.9}$$

where $g_0 = h_0 = 1$. For affine IFS maps on $\mathbb{R}$, set $f(x) = x^n$ in Equation (2.8) to give

$$h_n = \sum_{k=0}^n \binom{n}{k} \left[\sum_{i=1}^N p_i s_i^k a_i^{n-k}\right] g_k, \qquad i = 1, 2, \cdots,. \tag{2.10}$$

Now define

$$\begin{aligned} D(X) \equiv \bigg\{ \boldsymbol{g} = (g_0, g_1, g_2, \cdots) \mid g_n = \int_X x^n\, d\mu, \\ n = 0, 1, 2, \cdots, \mu \in \mathcal{M}(X) \bigg\}, \end{aligned} \tag{2.11}$$

i.e. the set of all (infinite) moment vectors for probability measures in $\mathcal{M}(X)$. Then for each N-map affine IFS $(\boldsymbol{w}, \boldsymbol{p})$ on (X, d), i.e. to each Markov operator $M: \mathcal{M}(X) \to \mathcal{M}(X)$, there corresponds a linear operator $A: D(X) \to D(X)$. In the standard basis $\{\boldsymbol{e}_i = (0, 0, \cdots, 0, 1, 0, \cdots)\}_{i=0}^{\infty}$, A is represented by a lower-triangular matrix.

In the case where $\mu = \nu = M\mu$, i.e. μ is the invariant measure of the IFS, then $h_n = g_n$, $n = 0, 1, 2, \cdots$, and $\boldsymbol{g} \in D(X)$ is the fixed point of A. A rearrangement of Equation (2.10) produces the following well-known relation:

$$(2.12) \qquad \left[1 - \sum_{i=1}^{N} p_i s_i^n\right] g_n = \sum_{k=0}^{n-1} \binom{n}{k} \left[\sum_{i=1}^{N} p_i s_i^k a_i^{n-k}\right] g_k, \qquad n = 1, 2, \cdots.$$

which allows the moments g_n to be computed recursively in terms of the IFS parameters s_i, a_i, p_i.

Finally, if any of the IFS maps w_i in $\boldsymbol{w}$ are polynomial of degree $\geqq 2$, it is not difficult to see that the linear operator $A: D(X) \to D(X)$ is *not* represented by a lower-triangular matrix. As well, the relations for moments of the IFS invariant measure, unlike Equation (2.12), are not complete, and the moments may not be computed recursively.

3. A collage theorem for moments and the inverse problem

Moment matching for the approximation of measures on $[0, 1]^q$, $q = 1, 2, \cdots$, can be justified by the fact that the convergence of moments is equivalent to the weak convergence of measures. Since we are working on compact spaces, the latter convergence is equivalent to convergence in Hutchinson metric d_{H}. This is summarized in the following theorem which is formally proved in [10].

Theorem 3.1. For $X = [0, 1]$, let μ and $\mu^{(j)} \in \mathcal{M}(X)$, $j = 1, 2, 3, \cdots$, with power moments defined as follows:

$$g_n = \int_X x^n \, d\mu, \qquad g_n^{(j)} = \int_X x^n \, d\mu^{(j)}, \qquad n = 0, 1, 2, \cdots.$$

Then the following are equivalent:

(i) $g_n^{(j)} \to g_n$ as $j \to \infty$, $\forall k$,

(ii) the sequence of measures $\mu^{(j)}$ converges weak* to μ, i.e. for any $f \in C(X)$, $\int f \, d\mu^{(j)} \to \int f \, d\mu$, as $j \to \infty$,

(iii) $d_{\mathrm{H}}(\mu^{(j)}, \mu) \to 0$ as $j \to \infty$.

The results of this theorem can easily be extended to $X = [0, 1]^q$, $q \geqq 2$.

The idea of using IFS and moment matching for the inverse problem of fractal/measure construction was first suggested in [5], Section 3.3, where the method was applied to a 2-map affine IFS in the complex plane. In the case $X = [0, 1]$ and a target measure $\nu \in \mathcal{M}(X)$ with moments $h_n = \int_X x^n \, d\nu$, Diaconis and Shahshahani [13] proposed the following method. For a fixed number $N>0$ of affine IFS maps on $\mathbb{R}$ with probabilities, cf. Equation (2.6), and a given number $M>0$ of moments to be matched, impose the conditions

$$g_n(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{p}) = h_n, \qquad n = 1, 2, \cdots, M, \tag{3.1}$$

and use the moment recursion relations in Equation (2.12) to solve for the IFS parameters s_i, a_i, p_i directly. However, the g_n are complicated non-linear functions of the IFS parameters and approximation schemes such as the Newton–Kantorovich method are unstable. In [24]–[26] moment matching was performed by minimizing the following truncated l^2 distance between target and IFS moment vectors:

$$D_M^N(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{p}) = \sum_{n=1}^{M} (g_n(\boldsymbol{s}, \boldsymbol{a}, \boldsymbol{p}) - h_n)^2. \tag{3.2}$$

Since the g_n are differentiable with respect to the IFS parameters s_i, a_i, p_i, gradient methods for optimization can be used. This method was also tested for target measures or images in $\mathbb{R}^2$. In the one-dimensional case, the method works reasonably well, although a considerable amount of computation may be required for the optimization of the $3N$ IFS parameters. These difficulties are further enhanced in the two-dimensional case. As well, the graph of the function D_M^N can be very complicated, especially as N or M increases. Local methods such as gradient schemes are not guaranteed to converge to global minima or even reasonable minima.

The modified moment matching approach which we now outline represents a significant improvement because of two major changes:

(i) The IFS maps w_i, hence the parameters s_i and a_i, $i = 1, 2, \cdots, N$, are fixed. Moment matching is done only with respect to the p_i.

(ii) Instead of using Equation (3.2), which involves complicated expressions of the IFS moments g_n in terms of the probabilities p_i, we use a collage distance for moments which involves only quadratic terms in the p_i.

Let us now recall a simple consequence of the Banach fixed point theorem which, in the IFS literature, is referred to as the collage theorem [4], [7].

Theorem 3.2 (Collage theorem). Let (Y, d_Y) be a complete metric space. Given a $y \in Y$, suppose that there exists a map $f \in \mathrm{Con}(Y)$ with contractivity factor $0 \leqq c < 1$, so that $d_Y(y, f(y)) < \epsilon$. If $\bar{y}$ is the fixed point of f, i.e. $f(\bar{y}) = \bar{y}$, then $d_Y(\bar{y}, y) < \epsilon/(1-c)$.

In other words, suppose there exists a 'target' y that we wish to approximate with a fixed point $\bar{y}$ of an unknown mapping f. The inverse problem reduces to finding an f which minimizes the collage distance $d_Y(y, f(y))$. This idea was first used for the geometric approximation of sets with IFS attractors [4], [7] as well as for more generalized IFS-type methods used for image representation [21]. The inverse problem for measures using IFS may now be posed as follows:

given a target measure $\nu \in \mathcal{M}(X)$ and a $\delta > 0$, find an IFS $(\boldsymbol{w}, \boldsymbol{p})$ with associated Markov operator M such that $d_{\mathrm{H}}(\nu, M\nu) < \delta$.

For ease of presentation, unless otherwise indicated, $X = [0, 1]$ is the IFS base space for the remainder of this section. The extension to $[0, 1]^q$, $q \geqq 2$ is straightforward. Now define

$$l^2(\boldsymbol{N}) = \left\{ \boldsymbol{c} = (c_0, c_1, c_2, \cdots) \mid c_i \in \mathbb{R}, \quad \|\boldsymbol{c}\|_{l^2}^2 \equiv \sum_{k=0}^{\infty} c_k^2 < \infty \right\}. \tag{3.3}$$

As well, define the following weighted Banach space of half-infinite sequences:

$$\bar{l}^2(\boldsymbol{N}) = \left\{ \boldsymbol{c} = (c_0, c_1, c_2, \cdots) \mid c_i \in \mathbb{R}, \quad \|\boldsymbol{c}\|_{\bar{l}^2}^2 \equiv c_0^2 + \sum_{k=1}^{\infty} \frac{1}{k^2} c_k^2 < \infty \right\}. \tag{3.4}$$

We shall consider, in particular, the following subset:

$$\bar{l}_0^2(\boldsymbol{N}) = \{\boldsymbol{c} \in \bar{l}^2(\boldsymbol{N}) \mid c_0 = 1\} \subset \bar{l}^2(\boldsymbol{N}).$$

Now recall $D(X)$, the set of moment vectors for all $\mu \in \mathcal{M}(X)$, defined in Equation (2.11). Clearly, $D(X) \subset \bar{l}_0^2(\boldsymbol{N})$.

Proposition 3.3. Let $X = [0, 1]$ and $\mu, \nu^{(n)} \in \mathcal{M}(X)$, $n = 1, 2, 3, \cdots$, with associated moment vectors $\boldsymbol{g}, \boldsymbol{g}^{(n)} \in D(X)$. Then $\|\boldsymbol{g} - \boldsymbol{g}^{(n)}\|_{\bar{l}^2} \to 0$ as $n \to \infty$ iff $d_{\mathrm{H}}(\mu, \nu^{(n)}) \to 0$ as $n \to \infty$.

Proof. The proof follows from the results of Theorem 3.1.

Proposition 3.4. Let $X = [0, 1]$. Define the following metric on $D(X)$: for $\boldsymbol{u}, \boldsymbol{v} \in D(X)$, $\bar{d}_2(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|_{\bar{l}^2}$. Then $(D(X), \bar{d}_2)$ is a complete metric space.

The proof of this proposition is given in the Appendix.

Recall that for each Markov operator $M: \mathcal{M}(X) \to \mathcal{M}(X)$ associated with an N-map IFS $(\boldsymbol{w}, \boldsymbol{p})$, there exists a linear operator $A: D(X) \to D(X)$, whose action is given in Equation (2.10).

Proposition 3.5. The linear operator A is contractive in $(D(X), \bar{d}_2)$.

Proof. In the standard basis $\{\boldsymbol{e}_i = (0, 0, \cdots, 0, 1, 0, \cdots)\}_{i=0}^{\infty}$, the (infinite) matrix representation of A is lower triangular. Hence, A has eigenvalues

$$\lambda_0 = a_{00} = 1, \qquad \lambda_n = a_{nn} = \sum_{i=1}^{N} p_i s_i^n, \qquad n \geqq 1. \tag{3.5}$$

Thus, $|\lambda_n| = |a_{nn}| < c^n < 1$ for $n \geqq 1$. A little algebra shows that for any $\boldsymbol{u}$, $\boldsymbol{v} \in D(X)$, $\|A(\boldsymbol{u} - \boldsymbol{v})\|_{\bar{l}^2} \leqq c \; \|\boldsymbol{u} - \boldsymbol{v}\|_{\bar{l}^2}$, which implies the contractivity of A.

Corollary 3.6. The operator A has a unique (attractive) fixed point $\bar{\boldsymbol{g}} \in D(X)$.

The components g_n of $\bar{\boldsymbol{g}}$ are the moments of $\bar{\mu}$, the invariant measure of the IFS $(\boldsymbol{w}, \boldsymbol{p})$, cf. Equation (2.12). We have now arrived at the major result of this section.

Corollary 3.7 (collage theorem for moments). Let (X, d) be a compact metric space and $\mu \in \mathcal{M}(X)$ with moment vector $\boldsymbol{g} \in D(X)$. Let $(\boldsymbol{w}, \boldsymbol{p})$ be an N-map IFS, with contractivity factor $0 \leqq c < 1$, such that $\bar{d}_2(\boldsymbol{g}, \boldsymbol{h}) = \|\boldsymbol{g} - \boldsymbol{h}\|_{\bar{l}^2} < \epsilon$, where $\boldsymbol{h} \in D(X)$ is the moment vector corresponding to $\nu = M\mu$. Then

$$\bar{d}_2(\boldsymbol{g}, \bar{\boldsymbol{g}}) < \frac{\epsilon}{1 - c}, \tag{3.6}$$

where $\bar{\boldsymbol{g}}$ is the moment vector corresponding to $\bar{\mu}$, the invariant measure of the IFS $(\boldsymbol{w}, \boldsymbol{p})$.

Thus, given a target measure μ with moment vector $\boldsymbol{g}$, the inverse problem becomes one of finding an IFS $(\boldsymbol{w}, \boldsymbol{p})$ such that the collage distance $\bar{d}_2(\boldsymbol{g}, \boldsymbol{h})$, where $\boldsymbol{h} = A\boldsymbol{g}$, is small. Our main result—Theorem 3.9 below—ensures that for IFS constructed from a set $\mathcal{W}$ which satisfies the ϵ-contractivity property the collage distance may be made arbitrarily small. For its proof, we shall make use of Proposition 2.2 and the following result.

Proposition 3.8 [12]. Let (X, d) be a compact metric space and let $(\boldsymbol{u}, \boldsymbol{p})$ denote an N-map IFS on (X, d) with contractivity factor c_u and invariant measure $\mu \in \mathcal{M}(X)$. Given an $\epsilon > 0$, suppose that there exists another N-map IFS, $(\boldsymbol{v}, \boldsymbol{p})$, with identical probabilities, such that

$$d_w^N(\boldsymbol{u}, \boldsymbol{v}) \equiv \max_{1 \leqq i \leqq N} \sup_{x \in X} d(u_i(x), v_i(x)) < \epsilon(1 - c_u). \tag{3.7}$$

Then $d_{\mathrm{H}}(\mu, \nu) < \epsilon$, where $\nu \in \mathcal{M}(X)$ is the invariant measure of $(\boldsymbol{v}, \boldsymbol{p})$.

Now let $\mathcal{W} = \{w_1, w_2, \cdots\}$ be an infinite set of affine contraction maps on $X = [0, 1]$ which satisfies the ϵ-contractivity condition. Let

$$\boldsymbol{w}^N = \{w_1, w_2, \cdots, w_N\}, \qquad N = 1, 2, \cdots, \tag{3.8}$$

denote N-map truncations of $\mathcal{W}$. As well, let

$$\Pi^N = \left\{ \boldsymbol{p}^N = (p_1, p_2, \cdots, p_N) \mid p_i \geqq 0, \sum_{i=1}^{N} p_i = 1 \right\} \tag{3.9}$$

denote the set of all probability N-vectors for $\boldsymbol{w}^N$. Note that $\Pi^N \subset \mathbb{R}^N$ is compact in the natural topology on $\mathbb{R}^N$. Now let $\mu \in \mathcal{M}(X)$ be a target measure with moment vector $\boldsymbol{g} \in D(X)$. For a $\boldsymbol{p}^N \in \Pi^N$, let M^N be the Markov operator corresponding to the N-map IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$. Also let $\nu_N = M^N\mu$, with associated moment vector $\boldsymbol{h}_N \in D(X)$. The collage distance between the moment vectors of μ and ν_N will be denoted as

$$\Delta^N(\boldsymbol{p}^N) \equiv \|\boldsymbol{g} - \boldsymbol{h}_N\|_{l^2}. \tag{3.10}$$

Since $\Delta^N : \Pi^N \to \mathbb{R}$ is continuous, it attains an absolute minimum value, to be denoted as $\Delta^N_{\min}$, on Π^N. The following theorem ensures that the collage distance may be made arbitrarily small.

Theorem 3.9. $\Delta^N_{\min} \to 0$ as $N \to \infty$.

Proof. We first show that $\Delta^{\mathrm{N}}_{\min}$ is non-increasing with respect to N, i.e. $\Delta^{n_1}_{\min} \leqq \Delta^{n_2}_{\min}$ for $n_1 > n_2 \geqq 1$.

Let $\boldsymbol{q}^N = (q^N_1, q^N_2, \cdots, q^N_N) \in \Pi^N$ be an absolute minimum point for Δ^N, $N \geqq 2$. Also let $\boldsymbol{p}^N \in \Pi^N$ with the restriction (without loss of generality) $p^N_N = 0$. Then $\Delta^N(\boldsymbol{p}^N) \geqq \Delta^N(\boldsymbol{q}^N) = D^N_{\min}$. Now let $\boldsymbol{q}^{N-1} = (q^{N-1}_1, \cdots, q^{N-1}_{N-1}) \in \Pi^{N-1}$ be an absolute minimum for Δ^{N-1} and set $p^N_i = q^{N-1}_i$, $i = 1, 2, \cdots, N-1$. Then $\Delta^{N-1}_{\min} = \Delta^N(\boldsymbol{p}^N) \geqq \Delta^N_{\min}$. Thus $\{\Delta^N_{\min}\}^\infty_{N=1}$ is a non-increasing sequence of non-negative numbers. Hence, there exists a limit, $L \geqq 0$, of this sequence. We now show that $L = 0$.

If $\lim_{N\to\infty} d_{\mathrm{H}}(\mu, M^N\mu) = 0$, for *one* sequence of finite IFS $\{(\boldsymbol{w}^N, \boldsymbol{p}^N)\}$, where $\boldsymbol{p} \in \Pi^N$, $N = 1, 2, \cdots$, then it follows from Theorem 3.1 that $\lim_{N\to\infty} \Delta^N_{\min} = 0$. Now, from Proposition 2.2, given any $\epsilon_1 > 0$, there exists a $\mu_f \in \mathcal{M}_f(X)$, such that $d_{\mathrm{H}}(\mu, \mu_f) < \epsilon_1$. Then $\mu_f = \sum^{n_f}_{i=1} \alpha_i \delta_{\bar{x}_i}$ for some $n_f \geqq 1$, where $\alpha_i > 0$, $1 \leqq i \leqq n_f$ and $\sum^{n_f}_{i=1} \alpha_i = 1$: Here, δ_y denotes a point-mass measure at $y \in X$. Note that μ_f is the invariant measure for the n_f-map IFS $(\boldsymbol{u}, \boldsymbol{p})$, where

$$u_i(x) = \bar{x}_i, \qquad p_i = \alpha_i, \qquad 1 \leqq i \leqq n_f. \tag{3.11}$$

The contractivity factor of this IFS is $c_u = 0$.

For any $\epsilon_2 > 0$, let $U_i = N_{\epsilon_2}(\bar{x}_i)$, the ϵ_2-neighbourhood of $\bar{x}_i$, $1 \leqq i \leqq n_f$. From our refinement assumption, there exist affine maps $v_i = w_{k_i} \in \mathcal{W}$, $1 \leqq i \leqq n_f$, such that $v_i(X) \subset U_i$. Now let $(\boldsymbol{v}, \boldsymbol{p})$ be the n_f-map IFS,

$$v_i(x) = w_{k_i}(x) = \beta_i x + \gamma_i, \qquad p_i = \alpha_i, \qquad 1 \leqq i \leqq n_f. \tag{3.12}$$

The contractivity factor of this IFS is $\beta \equiv \max_{1 \leqq i \leqq n_f} |\beta_i| < 2\epsilon_2$. Denote its Markov operator as M_f and its invariant measure as ν_f, i.e. $M_f \nu_f = \nu_f$. From the above construction, we have $\sup_{x\in X} d(u_i(x), v_i(x)) < 2\epsilon_2$, $1 \leqq i \leqq n_f$. Hence, from Proposition 3.8, with $c_u = 0$, we have

$$d_{\mathrm{H}}(\mu_f, \nu_f) < 2\epsilon_2. \tag{3.13}$$

Let N be the smallest integer for which $\{v_i\}^{n_f}_{i=1} \in \boldsymbol{w}^N$. Also let M^N be the Markov operator for the IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$, where $p_{k_i} = \alpha_i$, $1 \leqq i \leqq n_f$, and all other $p_j = 0$. Then $M^N = M_f$. Now, given the target measure $\mu \in \mathcal{M}(X)$, consider the inequality

$$d_H(\mu, M^N\mu) \leqq d_H(\mu, \mu_f) + d_H(\mu_f, \nu_f) + d_H(\nu_f, M^N\mu). \tag{3.14}$$

Note that

$$\begin{aligned} d_H(\nu_f, M^N\mu) &= d_H(M^N\nu_f, M^N\mu) \\ &\leqq \beta d_H(\nu_f, \mu) \\ &\leqq 2\epsilon_2[d_H(\nu_f, \mu_f) + d_H(\mu_f, \mu)]. \end{aligned} \tag{3.15}$$

Substitution of inequality (3.15) into inequality (3.14) yields

$$\begin{aligned} d_H(\mu, M^N\mu) &\leqq (1 + 2\varepsilon_2)[d_H(\mu, \mu_f) + d_H(\mu_f, \nu_f)] \\ &\leqq (1 + 2\epsilon_2)(\epsilon_1 + 2\epsilon_2). \end{aligned}$$

Given $\epsilon = 1/2^n$, we find $\epsilon_1, \epsilon_2 > 0$ such that $(1 + 2\epsilon_2)(\epsilon_1 + 2\epsilon_2) < 1/2^n$ and a finite IFS $(\boldsymbol{w}^{N_n}, \boldsymbol{p}^{N_n})$ for which $d_H(\mu, M^{N_n}\mu) < 1/2^n$. Hence $\lim_{n\to\infty} \inf d_H(\mu, M^{N_n}\mu) = 0$. Thus, $L = 0$ and the theorem is proved.

Remarks

1. Theorem 3.9 is a density result establishing that the set of invariant measures for all N-map IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$, where $\boldsymbol{p}^N \in \Pi^N$, $N = 1, 2, \cdots$, is dense in $(\mathcal{M}(X), d_H)$. This result can be extended to $[0, 1]^q$, $q \geqq 2$.

2. Although not explicitly stated in the proof, the collage distances Δ^N and, in particular, the sequence $\Delta^N_{\min}$, are also dependent on the ordering of the w_i maps in the infinite set $\mathcal{W}$. However, at this point, we are not interested in any questions about the 'optimal' ordering of the maps in $\mathcal{W}$ nor how N-map subsets $\boldsymbol{w}^N$ should be chosen.

3.1. *The inverse problem as a quadratic programming problem.* Let us now consider the square of the collage distance, cf. Equation (3.10), between the moment vector $\boldsymbol{g}$ of the target measure $\mu \in \mathcal{M}(X)$ and the moment vector $\boldsymbol{h}_N$ of the measure $\nu_N = M^N\mu$, where, as above, M^N is the Markov operator associated with the truncated IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$:

$$S^N(\boldsymbol{p}^N) \equiv (\Delta^N)^2(\boldsymbol{p}^N) = \sum_{n=1}^{\infty} \frac{1}{n^2}(h_n - g_n)^2. \tag{3.16}$$

Let $A^N: D(X) \to D(X)$ denote the linear operator associated with M^N. Then $\boldsymbol{h}_N = A^N \boldsymbol{g}$ and from Equation (2.10),

$$h_n = \sum_{i=1}^{N} A_{ni}^N p_i^N, \qquad n = 1, 2, 3, \cdots, \tag{3.17}$$

where

$$\begin{aligned} A_{ni}^N &= \int_X (w_i x + a_i)^n \, d\mu \\ &= \sum_{k=0}^{n} \binom{n}{k} s_i^k a_i^{n-k} g_k. \end{aligned} \tag{3.18}$$

The function $S^N(\boldsymbol{p}^N)$ may be written in the following form:

$$S^N(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}} \boldsymbol{x} + c, \qquad \boldsymbol{x} \in \mathbb{R}^N, \tag{3.19}$$

where $\boldsymbol{x} = \boldsymbol{p}^N = (p_1^N, p_2^N, \cdots, p_N^N)$. The elements of the symmetric matrix $\boldsymbol{Q}$ are given by

$$q_{ij} = \sum_{n=1}^{\infty} \frac{1}{n^2} A_{ni}^N A_{nj}^N, \qquad i, j \in \{1, 2, \cdots, N\}. \tag{3.20}$$

As well,

$$b_i = -2 \sum_{n=1}^{\infty} \frac{1}{n^2} g_n A_{ni}^N, \qquad i = 1, 2, \cdots, N \tag{3.21}$$

and

$$c = \sum_{n=1}^{M} \frac{g_n^2}{n^2}. \tag{3.22}$$

Since $0 \leqq A_{ni}^N \leqq 1$, the infinite sums in Equations (3.20) and (3.21) converge. The minimization of the moment collage distance $S^N(\boldsymbol{x})$ becomes the following standard quadratic programming problem with linear constraints:

$$\text{minimize } S^N(\boldsymbol{x}), \qquad \sum_{i=1}^{N} x_i = 1, \qquad x_i \geqq 0. \tag{3.23}$$

This represents a significant simplification of the moment matching problem since quadratic programming problems can be solved computationally in a finite number of steps.

4. Applications and numerical computations

In the applications to be described below, $X = [0, 1]$. The moments of the target measure μ are denoted by $\{g_n\}_{n=0}^{\infty}$. The approach begins with the selection of an infinite set $\mathcal{W}$ of affine IFS maps, cf. Equation (2.6), satisfying the refinement condition of Definition 2.3. We then truncate this set to produce an N-map IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$ and solve the quadratic programming problem in Equation (3.23). In practical calculations, it is possible to match only a finite number of moments, so we consider the minimization of the following function:

$$(4.1) \qquad \begin{aligned} S_M^N(\boldsymbol{p}^N) &= (\Delta_M^N)^2(\boldsymbol{p}^N) \\ &= \sum_{n=1}^{M} \frac{1}{n^2}\left(\sum_{i=1}^{N} A_{ni}^N p_i - g_n\right)^2, \qquad M = 1, 2, 3, \cdots, \end{aligned}$$

subject to the linear constraints in Equation (3.23).

The minimization of the function $S_M^N(\boldsymbol{x})$ in Equation (4.1) was performed with a quadratic programming (QP) algorithm developed by Best and Ritter [9]. (See the acknowledgements at the end of this paper.) The QP method is superior to gradient projection (GP) methods of minimizing our objective function (such as the Davidon method employed in [24]–[26]) for the following reasons. 1. QP locates the minimum of S^N on the simplex Π^N in a finite number of steps, whereas GP converges only to a local minimum of S^N and is sensitive to the initial point from where the search begins. Furthermore, the convergence of GP may be extremely slow, especially when the graph of the objective function is quite flat near a minimum. 2. In many of the problems we have studied, the minimum of S^N is achieved on a boundary point of Π^N, which implies that one or more probabilities p_i^N are zero. This, in turn, implies that the IFS maps w_i assocated with these vanishing probabilities are superfluous. QP essentially discards these maps. In general, we have found that GP rarely converges to such a minimum on the boundary of Π^N. As a result, many of these superfluous IFS maps are kept in the set.

We show below not only the minimum (truncated) collage distances Δ_M^N achieved for a particular truncation $(\boldsymbol{w}^N, \boldsymbol{p}^N)$ but also the following (truncated) distances in $D(X)$:

$$(4.2) \qquad \Gamma_M^N = \left[\sum_{n=1}^{M} \frac{1}{n^2}(\bar{g}_n - g_n)^2\right]^{\frac{1}{2}},$$

where, as above, $\boldsymbol{g}$ denotes the moment vector of the target measure μ and $\bar{\boldsymbol{g}}_N$ denotes the moment vector of the invariant measure $\bar{\mu}_N$ of the IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$. In the limit $M \to \infty$, it follows from the collage theorem for moments, cf. Equation (3.6), that

$$(4.3) \qquad \Gamma^N \equiv \bar{d}_2(\boldsymbol{g}, \bar{\boldsymbol{g}}) < \frac{1}{1-c}\Delta^N,$$

where c is the contractivity factor of the IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$.

In the following calculations, we used the 'wavelet-type' affine maps of Equation (2.7). The truncated IFS map vectors $\boldsymbol{w}^N$ were constructed by arranging the w_{ij} maps as follows:

$$w_{1,1}, w_{1,2}, w_{2,1}, \cdots, w_{2,4}, w_{3,1}, \cdots, w_{3,8}, \cdots. \tag{4.4}$$

The vector $\boldsymbol{w}^{N(i^*)}$, where $N(i^*) = \sum_{i=1}^{i^*} 2^i$, contains affine maps w_{ij} with contraction factors 2^{-i}, $i = 1, 2, \cdots, i^*$. (For $i^* = 1, 2, 3, 4$, $N(i^*) = 2, 6, 14, 30$, respectively.) In all cases, the contractivity factor of $\boldsymbol{w}^N$ is $c = \frac{1}{2}$. This appears to be a natural ordering of the maps since the refinement afforded by the IFS $\boldsymbol{w}^{N(i^*)}$ increases with i^*.

4.1. *A simple target measure with continuous density.* We first consider a target measure with continuous probability density function, namely, $\rho(x) = 6x(1-x)$. The moments of this measure are

$$g_n = \int_0^1 x^n \rho(x)\, dx = \frac{6}{(n+2)(n+3)}, \qquad n = 0, 1, 2, \cdots. \tag{4.5}$$

The convergence of IFS measures to the target measure will be demonstrated not only in terms of moments but also with regard to convergence to the distribution function $F(x)$, defined as

$$F(x) = \int_0^x d\mu(t) = \int_0^x \rho(t)\, dt. \tag{4.6}$$

In this case, $F(x) = x^2(3-2x)$. The distribution function corresponding to the invariant measure $\bar{\mu}_N$ of the IFS $(\boldsymbol{w}^N, \boldsymbol{p}^N)$ will be denoted as $\bar{F}_N(x)$, i.e.

$$\bar{F}_N(x) = \int_0^x d\bar{\mu}_N(t). \tag{4.7}$$

In order to compute $\bar{F}_N(x)$, we generate a discrete approximation $\bar{\mu}_N^K$ to $\bar{\mu}_N$ on subintervals I_k, $k = 1, 2, \cdots, K$, formed by the equipartition on $[0, 1]$ generated by the points $x_i = i/K$, $i = 0, 1, 2, \cdots, K$. In these calculations $K = 1000$. (For further details of this procedure of obtaining a discrete measure, see [25].) The discrete measure is represented by an array $\bar{M}_k$, $k = 1, 2, \cdots, K$, where $\bar{M}_k = \bar{\mu}_N^K(I_k)$. Then $F(x_k) = \sum_{i=1}^k \bar{M}_k$.

In Table 1, we summarize the results of our minimization procedure, using $M = 30$ moments. The collage distance Δ_M^N as well as the distance Γ_M^N between moments of the target measure and the approximating IFS measure are given. In each case, we list N, the number of maps in the truncation $\boldsymbol{w}^N$ over which the QP optimization was performed, as well as $\bar{N}_{\text{QP}} \leqq N$, the actual number of non-zero probabilities at the minimum point in Π^N. For purposes of comparison, we also list the minimum collage distances Δ_M^N yielded by GP as well as $\bar{N}_{\text{GP}}$, the number of non-zero probabilities at the minimum. For small values of N, the GP results are consistent with QP. However, as N increases, the GP method does not converge to minima found by

TABLE 1

Results of moment matching via minimization of the collage distance S_M^N in Equation (4.1) to approximate the measure with probability density function $\rho(x) = 6x(1-x)$. The 'wavelet-type' IFS maps of Equation (2.8) were used and the truncated IFS map vectors $\mathbf{w}^N$ were constructed according to (4.4) in the text. Columns 2–4 show the results for minimization via quadratic programming (QP). Rows which are missing correspond to values of N for which the QP method produced no decrease in the collage distance D_M^N, i.e. the probability assigned to map w_N was zero. The value $\bar{N}_{QP}$ denotes the number of maps in $\mathbf{w}^N$ with non-zero probabilities. The column under $\mathbf{w}^{\bar{N}_{QP}}$ lists the indices of those maps. Results obtained from the GP method of minimization are shown in columns 5 and 6, for comparison. $\bar{N}_{GP}$ denotes the number of maps with non-zero probabilities as determined by the GP method.

	QP				GP	
N	D_M^N	Γ_M^N	$\bar{N}_{QP}$	$\mathbf{w}^{\bar{N}_{QP}}$	D_M^N	$\bar{N}_{GP}$
2	$2{\cdot}13\times10^{-2}$	$3{\cdot}12\times10^{-2}$	2	(1, 2)	$2{\cdot}13\times10^{-2}$	2
4	$3{\cdot}58\times10^{-3}$	$4{\cdot}08\times10^{-3}$	3	(2-4)	$3{\cdot}58\times10^{-2}$	3
5	$7{\cdot}72\times10^{-5}$	$7{\cdot}72\times10^{-5}$	5	(1-5)	$7{\cdot}72\times10^{-5}$	5
7	$7{\cdot}60\times10^{-5}$	$7{\cdot}59\times10^{-5}$	5	(1, 2, 4, 5, 7)	$2{\cdot}13\times10^{-5}$	7
9	$6{\cdot}89\times10^{-5}$	$6{\cdot}74\times10^{-5}$	6	(2–5, 7, 9)	$6{\cdot}90\times10^{-5}$	7
12	$3{\cdot}66\times10^{-5}$	$3{\cdot}52\times10^{-5}$	6	(2–4, 8, 10, 12)	$3{\cdot}66\times10^{-5}$	10
13	$1{\cdot}05\times10^{-6}$	$1{\cdot}03\times10^{-6}$	8	(2–5, 8, 10, 11, 13)	$3{\cdot}28\times10^{-6}$	13

QP to lie on the boundary of Π^N. As a result, (i) more IFS maps are required for the approximation of the target measure and (ii) the accuracy of the approximation, in terms of moment distance, is poorer, especially as N gets larger.

In Figure 1 are presented some approximations to the distribution function $F(x) = x^2(3-2x)$ yielded by the IFS invariant measures $\bar{\mu}_N$. The convergence of the $\bar{F}_N(x)$ to $F(x)$ with increasing N is evident.

4.2. *IFS reconstruction of the spectral measure of an FCC crystal lattice.* We now apply our approximation method to a typical problem from theoretical physics which requires the computation of integrals over a measure which is not explicitly known. In such problems, the moments of these measures are usually available and a standard approach is to employ Padé approximants to reconstruct the measures.

Bessis and Demko [8] first showed that IFS invariant measures could also be used to approximate the measure for a specific problem involving crystal lattices. The problem is to determine thermodynamic averages of crystal models as integrals over distributions. The simple model which they studied was the face-centered cubic (FCC) lattice. The zero-point energy of this lattice is given by the integral

$$u_0 = \frac{1}{2}\int_0^1 \sqrt{x}\,G(x)\,dx, \tag{4.8}$$

where $G(x)\,dx$ is the fraction of vibrational modes in the interval $[x, x+dx]$. The function $G(x)$ is not known in closed form. However, the moments g_n of $G(x)$ were

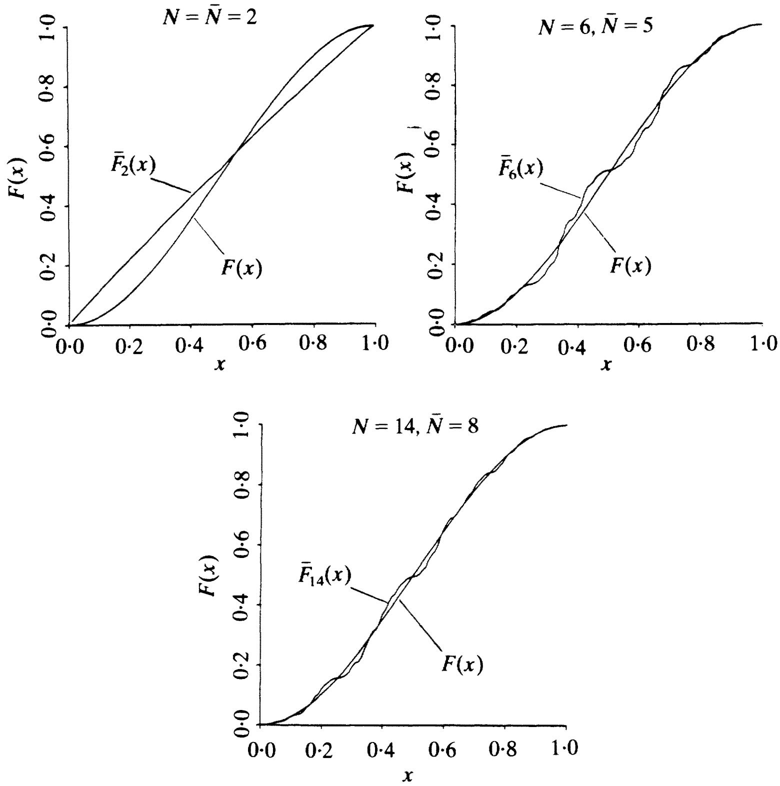

Figure 1. Estimates $\bar{F}_N(x)$ to the distribution function $F(x) = x^2(3 - 2x)$ as yielded by IFS invariant measures $\bar{\mu}_N$, $N = 2, 6, 14$, cf. Equation (4.7). The value $\bar{N}$ denotes the actual number of maps in the w^N vector with non-zero probabilities (cf. Table 1).

first calculated to order $n = 34$ by Isenberg [20]. Wheeler and Gordon [27] then constructed Padé approximants from these moments in order to numerically approximate this integral. Using 30 moments, they obtained the bounds

$$0.3408807 < u_0 \leqq 0.3408883. \tag{4.9}$$

With only 10 moments, their approximation was correct to 1 part in 10^4. Using additional information about $G(x)$, they were able to improve the bounds in Equation (4.9):

$$0.3408872177 < u_0 \leqq 0.3408872204. \tag{4.10}$$

Bessis and Demko's method of polynomial sampling uses *homogeneous* IFS maps on $[0, 1]$, $w_i(x) = sx + a_i$ with probabilities p_i, $i = 1, 2, \cdots, N$. The $2N + 1$ parameters s, a_i, p_i are then determined so that M target moments g_n, $n = 1, 2, \cdots, M$, are matched. Bessis and Demko then computed the integral in Equation (4.8) over the approximating IFS invariant measure. (The method of computing this integral is described below.) For the case $N = 4$ and $M = 9$, they obtained the approximation

$$\bar{u}_0 = \frac{1}{2}\int_0^1 \sqrt{x}\, d\mu = 0.340899. \tag{4.11}$$

The relative error of this approximation is 3.5×10^{-5}, an improvement over the Padé bounds using 10 moments. In [24], these results were improved slightly by minimizing the function in Equation (3.2) with $M = 9$, using gradient optimization. However, the use of a higher number of moments, e.g. $M = 20$ or 30, was not investigated at that time.

In Table 2, we list the results of minimizing the moment collage distance S_M^N in Equation (4.1), using $M = 30$ moments. The 'wavelet-type' IFS functions of Equation (2.7) were again used. In each case, the collage distance S_M^N as well as the moment distance Γ_M^N are presented along with the estimate $\bar{u}_0$ afforded by the IFS invariant measure. The relative errors of these estimates are also shown.

TABLE 2

Results of moment matching via minimization of the collage distance S_M^N in Equation (4.1) applied to the FCC lattice problem of Section 4.2. $M = 30$ moments were used and the optimization was performed using a quadratic programming algorithm. The quantity $\bar{u}_0$ denotes the approximation to the integral in Equation (4.11) which is obtained by integrating over the IFS invariant measure. The next column lists the relative error of each approximation. The integrals were computed using the iteration scheme described in Equations (4.12–14). $\bar{N}$ denotes the number of maps in w^N with non-zero probabilities, i.e. the actual number of IFS maps used to generate the approxating measure. The final column lists the indices of those maps. The final row, designated BD, gives the moment distances and estimate $\bar{u}_0$ yielded by the four homogeneous IFS maps (nine parameters) obtained by Bessis and Demko [8].

$M = 30$						
N	D_M^N	Γ_M^N	$\bar{u}_0$	$\frac{\lvert\bar{u}_0 - u_0\rvert}{u_0}$	$\bar{N}$	$w^{\bar{N}}$
2	$1{\cdot}10 \times 10^{-2}$	$1{\cdot}64 \times 10^{-2}$	0·328708	$3{\cdot}6 \times 10^{-2}$	2	(1, 2)
4	$1{\cdot}19 \times 10^{-3}$	$1{\cdot}29 \times 10^{-3}$	0·340433	$1{\cdot}3 \times 10^{-3}$	3	(1, 2, 4)
5	$1{\cdot}16 \times 10^{-3}$	$1{\cdot}23 \times 10^{-3}$	0·340200	$2{\cdot}0 \times 10^{-3}$	4	(1, 2, 4, 5)
8	$9{\cdot}99 \times 10^{-4}$	$1{\cdot}03 \times 10^{-3}$	0·341317	$1{\cdot}3 \times 10^{-3}$	4	(2, 4, 5, 8)
9	$8{\cdot}53 \times 10^{-4}$	$8{\cdot}64 \times 10^{-4}$	0·341763	$2{\cdot}6 \times 10^{-3}$	4	(2, 5, 8, 9)
12	$7{\cdot}67 \times 10^{-4}$	$7{\cdot}59 \times 10^{-4}$	0·341901	$3{\cdot}0 \times 10^{-3}$	5	(2, 7, 8, 9, 12)
13	$1{\cdot}20 \times 10^{-4}$	$1{\cdot}18 \times 10^{-4}$	0·340992	$3{\cdot}0 \times 10^{-3}$	6	(2, 3, 6, 8, 10, 13)
16	$1{\cdot}19 \times 10^{-4}$	$1{\cdot}18 \times 10^{-4}$	0·341074	$5{\cdot}5 \times 10^{-4}$	6	(2, 6, 8, 10, 13, 16)
17	$1{\cdot}17 \times 10^{-4}$	$1{\cdot}16 \times 10^{-4}$	0·341181	$8{\cdot}6 \times 10^{-4}$	6	(2, 6, 8, 10, 13, 17)
22	$1{\cdot}17 \times 10^{-4}$	$1{\cdot}16 \times 10^{-4}$	0·341223	$9{\cdot}9 \times 10^{-4}$	7	(2, 6, 8, 10, 13, 17, 22)
28	$1{\cdot}86 \times 10^{-5}$	$1{\cdot}77 \times 10^{-5}$	0·340711	$5{\cdot}1 \times 10^{-4}$	7	(2, 14, 15, 18, 19, 22, 28)
BD	$2{\cdot}33 \times 10^{-5}$	$2{\cdot}33 \times 10^{-5}$	0·340899	$3{\cdot}5 \times 10^{-5}$	4	

The integral in Equation (4.11) was computed as in [8], using the property

$$T^n f(x_0) \to \int_X f\,d\mu, \qquad \text{as } n \to \infty, \qquad x_0 \in X. \tag{4.12}$$

The action of the operator $T: C(X) \to C(X)$ is given by [5]

$$(Tf)(x) = \sum_{i=1}^{N} p_i (f \circ w_i)(x). \tag{4.13}$$

The iterates in Equation (4.12) are given by the nested sums

$$(T^n f)(x) = \sum_{i_1=1}^{N} \cdots \sum_{i_n=1}^{N} p_1 \cdots p_n f(w_{i_1} \circ \cdots w_{i_n})(x). \tag{4.14}$$

The evaluation of this quantity involves the enumeration of an N-tree to n generations.

5. Final remarks

In this paper, we have proved a collage theorem for moments which can greatly simplify the calculations involved in moment matching procedures using IFS. Moment matching becomes the minimization of a moment collage distance. A further simplification is achieved by using a fixed, infinite set of affine IFS maps satisfying an ϵ-contractivity condition. In this case, the minimization is performed only with respect to the IFS probabilities p_i. Note from Equation (3.16) that the squared moment collage distance S^N (as well as its truncation, S_M^N in Equation (4.1)) is a quadratic function of the p_i. The optimization problem may be performed by quadratic programming (QP) which locates a minimum on the simplex Π^N in a finite number of steps. In many cases, the minimum occurs on the boundary of Π^N so that 'useless' IFS maps w_i are removed. The additional condition that we use an infinite number of IFS maps which satisfy an ϵ-contractivity condition ensures that the collage distance $S^N \to 0$ as $N \to \infty$.

We have paid little attention to the question of choosing a set of affine IFS maps satisfying the ϵ-contractivity condition and which may be 'optimal' for a given problem. The 'wavelet-type' functions of Equation (2.7) represent a convenient choice of IFS maps. The question of using other maps which may be better suited to particular problems is beyond the scope of this paper. Note that the use of a fixed set of IFS maps has already become a standard tool in image compression methods [21]. Our method differs in that it allows room for increasing degrees of refinement on the base space X, as guaranteed by the ϵ-contractivity condition. We have also not given much attention to the question of the ordering of the contraction maps w_i in the infinite set $\mathcal{W}$. The 'wavelet-type' IFS maps of Equation (2.7) admit a natural ordering. Nevertheless, one may wish to exclude maps representing certain regions of X or, alternatively, to insert maps to permit additional refinement in certain regions. From a practical perspective, it will be important to develop optimal algorithms which are based on the problem at hand.

The present work involving IFS with measures was motivated, in part, by an ongoing study of the inverse problem of function approximation using IFS-type methods. Our construction of IFS-type methods over function spaces began with iterated fuzzy set systems (IFZS) [11], [15]: a variation of IFS which is formulated over an appropriate subset of functions from the class of functions $\mathscr{F}^*(X) = \{u : X \to [0, 1]\}$, often referred to as the class of *fuzzy sets* on X. However, the IFZS approach still employs a Hausdorff metric which is very restrictive from practical as well as theoretical perspectives. By making two modifications to the IFZS approach [16], one arrives at an IFS with 'grey level maps' (IFSM) over the space $\mathscr{L}^1(X, \mu)$. This, in turn, serves as the motivation to formulate IFS over the general function spaces $\mathscr{L}^p(X, \mu)$. Our solution to the inverse problem for function and image approximation in $\mathscr{L}^p(X, \mu)$ employs a strategy similar to the one described in this paper—constructing sequences of finite IFSM whose IFS maps w_i are chosen from an infinite set of contraction maps $\mathscr{W}$ which satisfy a refinement condition on (X, d) with respect to a measure μ. The basic aspects of this theory as well as some very encouraging results involving function and image approximation have already been reported [17].

Appendix: Proof of Proposition 3.4

Proposition 3.4. Let $X = [0, 1]$. Now define the following metric on $D(X)$ (cf. Equation (2.12): for $\boldsymbol{u}, \boldsymbol{v} \in D(X)$, $\bar{\boldsymbol{d}}_2(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|_{\bar{l}^2}$. Then $(D(X), \bar{d}_2)$ is a complete metric space.

Proof. Let $\boldsymbol{g}^{(n)} = (g_0^{(n)}, g_1^{(n)}, \cdots) \in D(X)$ for $n = 1, 2, \cdots$ be a Cauchy sequence in $\bar{l}^2$, that is, for any $\epsilon > 0$, there exists an $N > 0$ such that

$$\|\boldsymbol{g}^{(n)} - \boldsymbol{g}^{(m)}\|_{\bar{l}^2} < \epsilon, \qquad \forall m, n > N. \tag{A.1}$$

Let $\nu^{(n)} \in \mathscr{M}(X)$, $n = 1, 2, \cdots$, be the probability measures whose moments are the components of the $\boldsymbol{g}^{(n)}$, i.e. for $n = 1, 2, \cdots$,

$$g_k^{(n)} = \int_X x^k \, d\nu^{(n)}, \qquad k = 0, 1, 2, \cdots. \tag{A.2}$$

Now consider the sequences $\boldsymbol{a}^{(n)} = (a_0^{(n)}, a_1^{(n)}, \cdots)$, $n = 1, 2, \cdots$, where $a_0^{(n)} = g_0^{(n)} = 1$ and

$$a_k^{(n)} = \frac{1}{k} g_k^{(n)}, \qquad k = 1, 2, \cdots. \tag{A.3}$$

Since $0 \leqq g_k^{(n)} \leqq 1$ for $k \geqq 0$, it follows that $\boldsymbol{a}^{(n)} \in l^2(\boldsymbol{N})$ for all $n \geqq 1$. Furthermore, from Equation (A.1), $\{\boldsymbol{a}^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence in $l^2(\boldsymbol{N})$. Hence, by the completeness of $l^2(\boldsymbol{N})$, there exists an $\boldsymbol{a} = (a_0, a_1, \cdots) \in l^2(\boldsymbol{N})$ such that

$$\text{(A.4)} \qquad \|\boldsymbol{a}^{(n)} - \boldsymbol{a}\|_{l^2} \to 0 \qquad \text{as } n \to \infty.$$

Now let $\boldsymbol{g} = (g_0, g_1, \cdots)$, where $g_0 = 1$ and $g_k = ka_k$, $k = 1, 2, \cdots$. From Equation (A.4), it follows that for each $k = 1, 2, \cdots$, $|a_k^{(n)} - a^{(n)}| \to 0$ as $n \to \infty$ which, in turn, implies that $|g_k^{(n)} - g_k| \to 0$ as $n \to \infty$. Since $\boldsymbol{a} \in l^2(\boldsymbol{N})$, then $\boldsymbol{g}$, the limit of the Cauchy sequence $\{\boldsymbol{g}^{(n)}\}$, is an element of $\bar{l}^2(\boldsymbol{N})$. We now show that $\boldsymbol{g} \in D(X)$.

A necessary and sufficient condition that an infinite set of real numbers $\boldsymbol{c} = (c_0, c_1, \cdots)$ be the moments of a unique probability measure $\mu \in \mathcal{M}(X)$, i.e. $c_n = \int_X x^n \, d\mu$, $n = 0, 1, 2, \cdots$, is that they satisfy the Hausdorff inequalities [3]:

$$\text{(A.5)} \qquad H_{ij}(\boldsymbol{c}) \equiv \sum_{m=0}^{j} (-1)^m \binom{j}{m} c_{i+m} \geqq 0, \qquad i, j \in \{0, 1, 2, \cdots\}.$$

Since for each fixed $n \geqq 1$ the $g_k^{(n)}$, $k = 0, 1, 2, \cdots$, are the moments of the measures $\nu^{(n)} \in \mathcal{M}(X)$, cf. Equation (A.2), they must satisfy the following relations:

$$\text{(A.6)} \qquad H_{ij}(\boldsymbol{g}^{(n)}) = \sum_{m=0}^{j} (-1)^m \binom{j}{m} g_{i+m}^{(n)} \geqq 0, \qquad i, j \in \{0, 1, 2, \cdots\}.$$

The limit as $n \to \infty$ of each of these inequalities may be now be taken:

$$\text{(A.7)} \qquad H_{ij}(\boldsymbol{g}) \equiv \sum_{m=0}^{j} (-1)^m \binom{j}{m} g_{i+m} \geqq 0, \qquad i, j \in \{0, 1, 2, \cdots\}.$$

The above inequalities are simply the Hausdorff inequalities for the sequence $\boldsymbol{g}$. This implies that $g_k = \int_X x^k \, d\nu$, $k = 0, 1, 2, \cdots$, for a unique measure $\nu \in \mathcal{M}(X)$. Thus $\boldsymbol{g} \in D(X)$, which completes the proof.

Acknowledgements

We wish to express our thanks to Professor M. J. Best, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo for many useful discussions on quadratic programming and for providing the FORTRAN QP subroutines which were used in the computations reported above. This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada, which are gratefully acknowledged.

References

[1] Abenda, S., Demko, S. and Turchetti, G. (1992) Local moments and inverse problem for fractal measures. *Inverse Problems* **8**, 737–750.

[2] Abenda, S. and Turchetti, G. (1989) Inverse problem of fractal sets on the real line via the moment method. *Nuovo Cim.* **104**B, 213–227.

[3] AKHIEZER, N. I. (1965) *The Classical Moment Problem and Some Related Questions in Analysis.* Hafner, New York.

[4] BARNSLEY, M. F. (1988) *Fractals Everywhere.* Academic Press, New York.

[5] BARNSLEY, M. F. AND DEMKO, S. (1985) Iterated function systems and the global construction of fractals. *Proc. R. Soc. London* A**399**, 243–275.

[6] BARNSLEY, M. F. AND ELTON, J. H. (1988) A new class of Markov processes for image encoding. *Adv. Appl. Prob.* **20**, 14–32.

[7] BARNSLEY, M. F., HARDIN, D., ERVIN, V. AND LANCASTER, J. (1986) Solution of an inverse problem for fractals and other sets. *Proc. Nat. Acad. Sci. USA* **83**, 1975–1977.

[8] BESSIS, D. AND DEMKO, S. (1991) Stable recovery of fractal measures by polynomial sampling. *Physica* **47**D, 427.

[9] BEST, J. J. AND RITTER, K. (1988) A quadratic programming algorithm. *Z. Operat. Res.* **32**, 271–297.

[10] CABRELLI, C. A., MOLTER, U. M. AND VRSCAY, E. R. (1992) "Moment matching" for the approximation of measures using iterated function systems. Preprint.

[11] CABRELLI, C. A., FORTE, B., MOLTER, U. M. AND VRSCAY, E. R. (1992) Iterated fuzzy set systems: A new approach to the inverse problem for fractals and other sets. *J. Math. Anal. Appl.* **171**, 79–100.

[12] CENTORE, P. AND VRSCAY, E. R. (1994) Continuity of attractors and invariant measures for iterated function systems. *Canad. Math. Bull.* **37**, 315–329.

[13] DIACONIS, P. AND SHAHSHAHANI, M. (1986) Products of random matrices and computer image generation. *Contemp. Math.* **50**, 173–182.

[14] ELTON, J. AND YAN, Z. (1989) Approximation of measures by Markov processes and homogeneous affine iterated function systems. *Constr. Approx.* **5**, 69–87.

[15] FORTE, B., LO SCHIAVO, M. AND VRSCAY, E. R. (1994) Continuity properties of attractors for iterated fuzzy set systems. *J. Austral. Math. Soc.* B **36**, 175–193.

[16] FORTE, B. AND VRSCAY, E. R. (1995) Solving the inverse problem for function and image approximation using iterated function systems. *Dynamics of Continuous, Discrete and Impulsive Systems.* To appear.

[17] FORTE, B. AND VRSCAY, E. R. (1994) Solving the inverse problem for function and image approximation using iterated function systems I. Theoretical basis. *Fractals* **2**, 325–334. II. Algorithm and computations. *Fractals* **2**, 335–346.

[18] HANDY, C. AND MANTICA, G. (1990) Inverse problems in fractal construction: moment method solution. *Physica* D**43**, 17–36.

[19] HUTCHINSON, J. (1981) Fractals and self-similarity. *Indiana Univ. J. Math* **30**, 713–747.

[20] ISENBERG, C. (1963) Moment calculations in lattice dynamics, I. FCC lattice with nearest-neighbour interactions. *Phys. Rev.* **132**, 2427–2433.

[21] JACQUIN, A. (1989) A Fractal Theory of Iterated Markov Operators with Applications to Digital Image Coding. Ph.D. Thesis, Georgia Institute of Technology.

[22] MANTICA, G. AND SLOAN, A. (1989) Chaotic optimization and the construction of fractals: solution of an inverse problem. *Complex Systems* **3**, 37–62.

[23] PARTHASARATHY, K. R. (1967) *Probability Measures on Metric Spaces.* Academic Press, New York.

[24] VRSCAY, E. R. (1991) Moment and collage methods for the inverse problem of fractal construction with iterated function systems. In *Fractals in the Fundamental and Applied Sciences,* ed. by H.-O. Peitgen, J. M. Henriques and L. F. Penedo, pp. 443–461. North-Holland, Amsterdam.

[25] VRSCAY, E. R. (1991) Iterated function systems: theory, applications and the inverse problem. In *Proceedings of the NATO Advanced Study Institute on Fractal Geometry and Analysis,* ed. J. Belair and S. Dubuc, pp. 405–468. Kluwer, Dordrecht.

[26] VRSCAY, E. R. AND ROEHRIG, C. (1989) Iterated function systems and the inverse problem of fractal construction using moments. In *Computers and Mathematics,* ed. E. Kaltofen and S. M. Watt, pp. 250–259. Springer-Verlag, New York.

[27] WHEELER, J. C. AND GORDON, R. G. (1970) Bounds for averages using moment constraints. In *The Padé Approximant in Theoretical Physics,* ed. G. A. Baker Jr and J. L. Gammel, pp. 99–128. Academic Press, New York.