DIRECT CALCULATION OF THE $f(\alpha)$ FRACTAL DIMENSION SPECTRUM FROM HIGH-DIMENSIONAL CORRELATION-INTEGRAL PARTITIONS

M. Potter and W. Kinsner

University of Manitoba
Dept. of Electrical and Computer Engineering
Winnipeg, MB R3T 5V6
{m.potter|kinsner}@ieee.org

ABSTRACT

Fractal dimension spectra have been used to characterize the complexity of dynamical time series since the 1980s. Calculation of these spectra are traditionally based on fixed-size methods that are grid-based, such as the histogram technique, or sample-based, such as the correlation-integral method. This paper extends the Chhabra and Jensen direct approach on histogram-binned data by deriving the direct calculation of the $f(\alpha)$ spectrum of scaling indices from correlation-integral based partition functions. That is, the canonical correlation-integral approach to $f(\alpha)$ is defined. The benefit of this novel method is that the extended dynamical range of the correlation-integral can be used to generate the compact $f(\alpha)$ spectrum from high-dimensional embeddings without resorting to the Legendre transform. A comparison of spectra results on the Ikeda attractor are presented.

***Index Terms*—** Fractals, Signal Analysis, Nonlinear Systems, Multidimensional Signal Processing

1. INTRODUCTION

The *fractal dimension spectrum* is an important invariant for characterization of a dynamical attractor. It is a feature invariant to smooth topological transformations induced by measurement functions [1] and can be used as a distinguishing feature for the comparison of models to data recorded from natural systems [2]. The scaling properties of dynamical attractors must be characterized simultaneously by a multiplicity (or vector) of fractal dimensions in order to capture the inhomogeneities in the attractor density, and thus distinguish between a monofractal and a multifractal [1]. In particular, the scaling properties of attractors defined by multiplicative cascades of finitely many affine transformations are completely described by a fractal dimension spectrum [3].

1.1. Two Multifractal Spectra

A *multifractal characterization* is presented in one of two equivalent fractal dimension spectra: (i) the *Rényi fractal dimension spectrum* (RFDS), D_q, or (ii) the spectrum of scaling indices, $f(\alpha)$, [4], which here is called the *Mandelbrot fractal dimension spectrum* (MFDS). The RFDS [4] is important because it (i) is the historic unification and extension of the various historical fractal dimensions defined previously by Mandelbrot and others, (ii) has a geometric interpretation for positive integral q regarding q-tuple correlations [5], and (iii) has the nice property of existing in the $L^{\infty}(\mathbb{R})$ function space. The second spectrum, the MFDS [4], (i) describes a multifractal as a union of interwoven monofractal sets, and therefore has a nicer interpretation, and (ii) also provides a functional form with compact support. Though the MFDS variables, α and $f(\alpha)$, have a nice interpretation, Halsey *et al.* considered them directly unobservable and calculated them through the RFDS via a Legendre transform identity [4]. Chhabra and Jensen [6] later derived a direct form for the calculation of the MFDS, effectively describing a thermodynamic "canonical" form complementary to the "microcanonical" approach of the RFDS. Via their method, the MFDS variables, $\alpha(q)$ and $f(\alpha(q))$, can be considered global weighted averages of the local singular behaviour of the attractor measure.

1.2. Two Partitions

Fractal dimension spectra must be calculable from finite time series data. Once the multidimensional points of the attractor are reconstructed (likely by lag-embedding [7][2]), the fractal dimension spectra are defined based on a *partitioning* of the attractor according to, commonly, one of two different approaches. The first partitioning approach [8][2] is to use non-overlapping fixed-size boxes. This effectively develops a histogram estimate (of size ϵ) of the probability density function for the attractor. The simplicity and speed of this approach has made it quite natural for the origin and instruction of multifractal feature extraction. However, this method tends to become impractical, especially for attractors embedded in high-dimensional spaces [9].

Grassberger and Proccacia developed an independent partitioning approach based on the correlation-integral [10] which uses overlapping fixed-sized cells centred on the sample points of the attractor. Though the original implementation focused solely on calculating D_2, Pawelzik and Schuster [11] generalized the approach to calculate the entire RFDS. It has been shown that the dynamic range of the correlation-integral method outperforms the histogram method [2], and is the more popular approach in research. The complication of the correlation-integral approach to partition calculation is the $\mathcal{O}(N^2)$ algorithm complexity and significant bias for $q < 1$ [12, Sec. 11.3.1].

Thanks to the Natural Sciences and Engineering Research Council (NSERC) of Canada for funding.

1.3. Incomplete Theory

The RFDS was developed initially for calculation using histogram partitioning [1], and was generalized for calculation using correlation-integral partitioning [11]. The MFDS was developed initially for calculation from the completed RFDS via a transform (and therefore applies to either partitioning) [4], and then generalized to a direct approach [6]. However, the correlation-integral redefinition of the partition function invalidates the direct MFDS formula derived in [6]. As such, the theory remains incomplete. Can a correlation-integral partition be used to calculate $f(\alpha)$ directly?

This paper adds the missing piece to the "mosaic" of multifractal theory by deriving a direct formula for the MFDS under the assumption of a correlation-integral partition. This technique maintains the sound interpretation of the canonical approach, but allows the dynamical range of the correlation integral to enhance the calculation from finite data in higher-dimensional embeddings.

The derivation of interest is performed in Sec. 3, after a brief description of the formulae of the other methods in Sec. 2 for comparison. The new method is applied to the Ikeda attractor in Sec. 4 and compared to the performance of the other algorithms.

2. BACKGROUND

2.1. Attractor Characterization by Rényi Dimensions

2.1.1. Histogram Partitioning

Consider a dynamical attractor in $\mathbb{R}^d$, so that the natural measure μ of the attractor acts as a functional on the subsets of $\mathbb{R}^d$. By applying a fixed-size grid $\mathcal{V}$ of size $\epsilon > 0$ onto $\mathbb{R}^d$, a partition $\mathcal{P}$ of μ is induced according to

$$\mathcal{P}_\mu(\epsilon) = \left\{ p_i(\epsilon) = \frac{\mu(V_i)}{\mu(\mathbb{R}^d)} \middle| \ \forall V_i \in \mathcal{V}(\epsilon) \right\} \tag{1}$$

From time series data drawn from the attractor by time-delay reconstruction [7][2][12][8], this partition can be approximated easily by the empirical frequency ratio,

$$p_i(\epsilon) = \frac{N_i(\epsilon)}{N} \tag{2}$$

where N is the total number of points drawn from the attractor distribution and $N_i(\epsilon)$ is the number of points contained in the cell V_i of the grid $\mathcal{V}(\epsilon)$. The Rényi generalized entropy of order q of the discrete partition is then, [13][14],

$$H_q = \frac{1}{1-q} \log \sum_i p_i^q \tag{3}$$

which is a generalization of Shannon entropy, H_{S}, since [13][1]

$$\lim_{q \to 1} H_q = H_{\mathrm{S}} = -\sum_i p_i \log p_i \tag{4}$$

This feature of μ is a function over q. Since the partition $\mathcal{V}$ is of size ϵ, it is also scale-dependent and should be written $H_q(\epsilon)$. The RFDS is the set of Hentschel and Proccaccia generalized dimensions defined as the function $D_q : \mathbb{R} \mapsto \mathbb{R}$ such that

$$D_q = \lim_{\epsilon \to 0} \frac{-H_q(\epsilon)}{\log \epsilon} \tag{5}$$

As such, the general interpretation of the RFDS is that it quantifies the exponential behaviour of Rényi's generalized entropies for the discrete probability measures induced by a scaling partition of the attractor. Nice interpretations exist for subsets of the D_q, in particular for the capacity dimension $q = 0$, the information dimension $q = 1$, and the correlation dimension $q = 2$ [1][2].

To evaluate D_q, the limit is avoided and instead a *scaling region* is determined in the right hand side of (5) by evaluating $-H_q(\epsilon)$ as a function of $\log(\epsilon)$. (Note that this requires the evaluation of $\mathcal{P}_\mu(\epsilon)$ by histogram binning at several scales) Linear regions of this plot (similar to Fig. 1) indicate a scaling, and the linear slope is extracted as an estimate for D_q.

2.1.2. Correlation-Integral Partitioning

The Grassberger-Procaccia correlation-integral formalism uses the interpretation

$$\sum_i p_i^q = \sum_i p_i p_i^{(q-1)} = \mathcal{E}\left\{ p_i^{(q-1)} \right\} \tag{6}$$

to redefine the partition of the attractor measure μ. Again considering an attractor in $\mathbb{R}^d$, a covering $\tilde{\mathcal{V}}(\epsilon)$ of overlapping balls of size ϵ centred on the known points $\mathbf{x}_i$ of the attractor are taken. These balls induce a discrete partition,

$$\tilde{\mathcal{P}}_\mu(\epsilon) = \left\{ \tilde{p}_i(\epsilon) = \frac{\mu(\tilde{V}_i)}{\mu(\mathbb{R}^d)} \middle| \ \forall \tilde{V}_i = B_{\mathbf{x}_i}(\epsilon) \in \tilde{\mathcal{V}}(\epsilon) \right\} \tag{7}$$

but the Rényi entropies (3) have a revised form where the summation is replaced according to (6) [1][11]. The values for the $\tilde{p}_i$ are estimated from the attractor data $\{\mathbf{x}_j\}$ based on the number of points $\tilde{N}_i$ that appear in $\tilde{V}_i$ and are not temporally correlated to the point $\mathbf{x}_i$. [15] That is, the points $\mathbf{x}_j$, $|j - i| < W$ are discarded from $\tilde{p}_i$, because the pair-correlation should be based on the dynamics, and not temporal correlation. References [15][2][12] can provide insight into proper selection of W. It is clear however, that $W \geq 1$ is necessary, at least to discard the self-counting of $\mathbf{x}_i$ from $\tilde{p}_i$. Since there are fewer valid points, the normalization in the correlation-integral approach is slightly different from (2)

$$\tilde{p}_i(\epsilon) = \frac{\tilde{N}_i(\epsilon)}{N - W} \tag{8}$$

It is typical [11][5] to use the Heaviside function Θ to express $\tilde{N}_i$ as

$$\tilde{N}_i(\epsilon) = \sum_{|j-i| \geq W} \Theta(\epsilon - \|\mathbf{x}_j - \mathbf{x}_i\|) \tag{9}$$

Finally, if the points on the attractor are the result of an ergodic trajectory (i.e., a single initial condition) [11], each point is equally likely under μ (i.e., the density of the $\mathbf{x}_i$ is the same as the density of μ) and the expectation can use simple averaging, and thus the result of Pawelzik and Schuster [11]

$$H_q(\epsilon) = \frac{1}{1-q} \log \left(\frac{1}{N} \sum_{i=1}^{N} \tilde{p}_i^{(q-1)}(\epsilon) \right) \tag{10}$$

$$D_q = \lim_{\epsilon \to 0} \frac{-H_q(\epsilon)}{\log \epsilon} \tag{11}$$

is obtained. Again, experimental estimates are obtained by the slope of a scaling region of (11).

2.2. Attractor Characterization by $f(\alpha)$

Given the RFDS D_q vs. q, the MFDS (with its elegant multifractal interpretation) can be calculated by the Legendre transform [4] [2] as

$$\alpha = -\frac{\partial}{\partial q}(1-q)D_q \tag{12}$$

$$f(\alpha) = q\alpha + (1-q)D_q \tag{13}$$

For the explanation of how the interpretation follows from the derivation, consult references [4] [2].

2.3. Direct Canonical Approach to $f(\alpha)$

Consider again, a histogram partition $\mathcal{P}_\mu(\epsilon) = \{p_i(\epsilon)\}$ of an attractor in $\mathbb{R}^d$ with measure μ. Chhabra and Jensen [6] derived their formula and its interpretation from an independent set of first principles based on the Shannon entropies of a q-ordered family of measures $\mu_i(q,\epsilon) = p_i^q(\epsilon)/\sum_i p_i^q(\epsilon)$. It follows from this definition that their result must be consistent with the Legendre transformations, where the derivative actions pass inside the sums and limits. Alternatively, the Legendre identities (12) and (13) can be exploited to obtain from (5) the direct MFDS form

$$\alpha(q) = -\frac{\partial}{\partial q}(1-q)\lim_{\epsilon\to 0}\frac{-H_q(\epsilon)}{\log\epsilon} \tag{14}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\frac{\partial}{\partial q}\log\sum_i p_i^q(\epsilon) \tag{15}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\sum_i \frac{p_i^q(\epsilon)}{Z_q(\epsilon)}\log p_i(\epsilon) \tag{16}$$

and

$$f(q) = q\alpha(q) + (1-q)\lim_{\epsilon\to 0}\frac{-H_q(\epsilon)}{\log\epsilon} \tag{17}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\left(\sum_i \frac{p_i^q(\epsilon)}{Z_q(\epsilon)} q\log p_i(\epsilon) - Z_q(\epsilon)\right) \tag{18}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\left(\sum_i \frac{p_i^q(\epsilon)}{Z_q(\epsilon)}\log\frac{p_i^q(\epsilon)}{Z_q(\epsilon)}\right) \tag{19}$$

where (3) is used for substitution and $Z_q = \sum_i p_i^q(\epsilon)$ is used as shorthand. The MFDS relationship is then experimentally determined by finding coincident scaling regions of the right hand sides of (16) and (19) and using slopes to identify the values for $\alpha(q)$ and $f(q)$. The MFDS is thus the parametric relationship of α and $f(\alpha)$ via q.

3. DERIVATION

Now we are able to derive the novel direct MFDS formulation from the correlation-integral partition. We follow the same approach as Sec. 2.3, but utilize substitutions from Sec. 2.1.2. Thus, it is clear that (12) and (13) require that

$$\alpha(q) = -\frac{\partial}{\partial q}(1-q)\lim_{\epsilon\to 0}\frac{-\frac{1}{1-q}\log\left(\frac{1}{N}\sum_{i=1}^N \tilde{p}_i^{(q-1)}(\epsilon)\right)}{\log\epsilon} \tag{20}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\frac{\partial}{\partial q}\log\left(\frac{1}{N}\sum_{i=1}^N \tilde{p}_i^{(q-1)}(\epsilon)\right) \tag{21}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\left(\frac{1}{N}\sum_{i=1}^N \frac{\tilde{p}_i^{(q-1)}(\epsilon)}{\tilde{Z}_q}\log\tilde{p}_i(\epsilon)\right) \tag{22}$$

and

$$f(q) = q\alpha(q) + (1-q)\times \lim_{\epsilon\to 0}\frac{-1}{\log\epsilon}\frac{1}{1-q}\log\left(\frac{1}{N}\sum_{i=1}^N \tilde{p}_i^{(q-1)}(\epsilon)\right) \tag{23}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\frac{q}{N}\sum_{i=1}^N \frac{\tilde{p}_i^{(q-1)}(\epsilon)}{\tilde{Z}_q}\log\tilde{p}_i(\epsilon) - \log\tilde{Z}_q \tag{24}$$

$$= \lim_{\epsilon\to 0}\frac{1}{\log\epsilon}\left(\frac{1}{N}\sum_{i=1}^N \frac{\tilde{p}_i^{(q-1)}(\epsilon)}{\tilde{Z}_q}\log\frac{\tilde{p}_i^q(\epsilon)}{\tilde{Z}_q}\right) \tag{25}$$

now using $\tilde{Z}_q = \sum_i \tilde{p}_i^{(q-1)}(\epsilon)/N$ as shorthand. Equations (22) and (25) are our main result. It is particularly important to notice that the asymmetry in the exponents of the $\tilde{p}_i$ in (25) as compared to the histogram case (19) which is completely symmetric (having the form of a Shannon entropy).

4. EXPERIMENTAL RESULTS

We have applied the calculation of (22) and (25) to 96×2^{10} points drawn from the Ikeda map [12]. The real component of the map was lag-embedded into $\mathbb{R}^5$ with a lag of 1. The correlation integral was evaluated using a box-assist method [12] on the range 2^{-2} to 2^{-9} with the L^∞ metric. (The Theiler window is negligible here because the Ikeda map produces uncorrelated values.) Over the 3-octave scaling region shown in Fig. 1(a) and (b), numerical fitting to a line produces a good slope estimate for $\alpha(q)$ and $f(q)$ for $q = (-2,\ldots,10)$. The parametrically defined canonical correlation-integral MFDS is shown in Fig. 1(c) (stars). The Legendre transform of the RFDS calculated via a correlation-integral partition is also shown (squares) and is in agreement with the new approach.

5. CONCLUSIONS

This paper has derived the direct formula for the Mandelbrot fractal dimension spectrum under the assumption of a correlation-integral partition. This provides the missing "canonical correlation-integral" approach to the mosaic of canonical multifractal theory. This direct calculation of $f(\alpha)$ is suitable for application to high-dimensional data and situations that benefit from the dynamic range of the correlation-integral. Preliminary application of the direct formula on a lag-embedding of the Ikeda map has demonstrated an agreement with the Legendre transform of the Rényi fractal dimension spectrum.

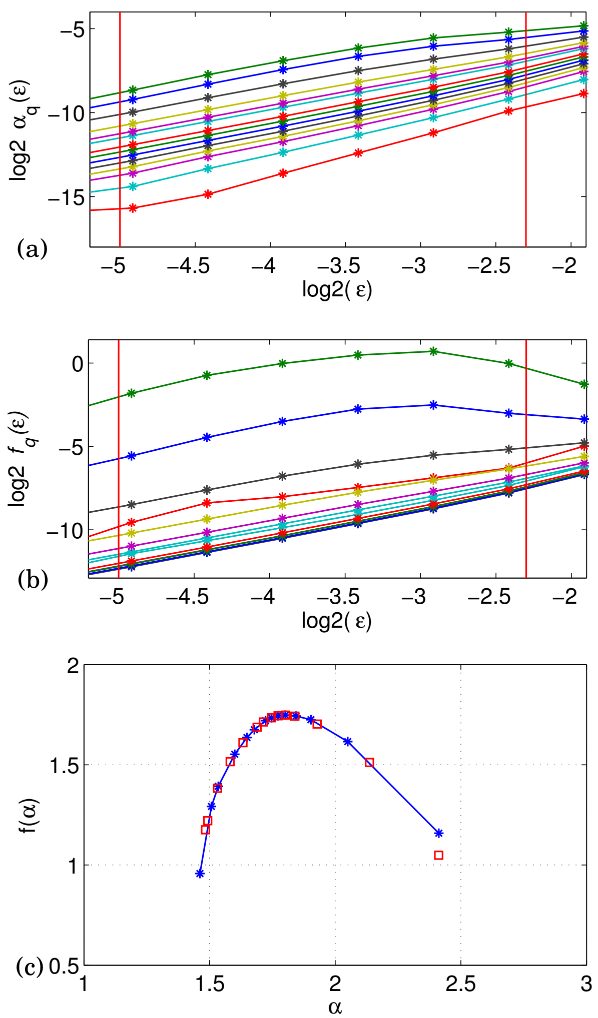

Fig. 1. Ikeda Attractor Results: scaling plots of (a) $\alpha(q)$ vs. $\log\epsilon$, and (b) $f(q)$ vs. $\log\epsilon$ over the scaling region $\epsilon \in [2^{-5}, 2^{-2.3}]$. Approximating (22) and (25) by fitting the slopes in (a) and (b) respectively, the MFDS is shown (stars) in (c). This is in agreement with the legendre transform of the RFDS (squares). The attractor is a 5-dimensional lag-embedding of 96×2^{10} real points drawn from the Ikeda map.

6. REFERENCES

[1] H. Hentschel and I. Procaccia, "The infinite number of generalized dimensions of fractals and strange attractors," *Physica D*, vol. 8, no. 3, pp. 435–444, 1983.

[2] J. Theiler, "Estimating fractal dimension," *J. Optical Society of America A*, vol. 7, no. 6, pp. 1055–1073, June 1990.

[3] B. B. Mandelbrot, C. J. Evertsz, and Y. Hayakawa, "Exactly self-similar left-sided multifractal measures," *Physical Review A*, vol. 42, no. 8, pp. 4528–4536, Oct. 1990.

[4] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, "Fractal measures and their singularities: The characterization of strange sets," *Physical Review A*, vol. 33, no. 2, pp. 1141–1151, Feb. 1986.

[5] W. Kinsner, "Entropy-based fractal dimensions: Probability and pair-correlation algorithms for e-dimensional images and strange attractors," University of Manitoba, Winnipeg, MB, Canada R3T 5V6, Tech. Rep. 94-5, June 1994, 44 pp.

[6] A. B. Chhabra and R. V. Jensen, "Direct determination of the $f(\alpha)$ singularity spectrum," *Physical Review Letters*, vol. 62, no. 12, pp. 1327–1330, Mar. 1989.

[7] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*, ser. Lecture Notes in Mathematics, D. A. Rand and L. S. Young, Eds. Springer, 1981, vol. 898, pp. 366–381.

[8] P. S. Addison, *Fractals and Chaos*. Philadelphia, PA, USA: Institute of Physics, 1997.

[9] H. S. Greenside, A. Wolf, J. Swift, and T. Pignataro, "Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors," *Physical Review A*, vol. 25, no. 6, pp. 3453–3456, June 1982.

[10] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D*, vol. 9, pp. 189–208, 1983.

[11] K. Pawelzik and H. G. Schuster, "Generalized dimensions and entropies from a measured time series," *Physical Review A*, vol. 35, no. 1, pp. 481–484, Jan. 1987.

[12] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, ser. Cambrdige Nonlinear Science Series. Cambridge, UK: Cambridge University Press, 1997, no. 7.

[13] A. Rényi, *Probability Theory*. Amsterdam: North-Holland, 1970.

[14] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 6th ed. Toronto: Halsted Press (John Wiley and Sons), 1994, vol. 1.

[15] J. Theiler, "Spurious dimension from correlation algorithms applied to limited time-series data," *Physical Review A*, vol. 34, pp. 2427–2432, Sept. 1986.