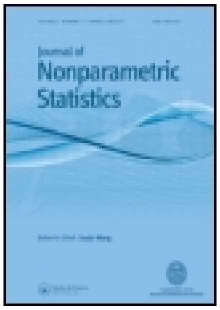

Journal of Nonparametric Statistics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gnst20

Dimension spectra for multbfractal measures with connections to nonparametric density estimation

Attila Frigyesi [a] & Ola Hössjer [a]
[a] Lund University , Sweden
Published online: 12 Apr 2007.

Nonparametric Statistics, Vol. 13, pp. 351–395
Reprints available directly from the publisher
Photocopying permitted by license only

DIMENSION SPECTRA FOR MULTIFRACTAL MEASURES WITH CONNECTIONS TO NONPARAMETRIC DENSITY ESTIMATION*

ATTILA FRIGYESI† and OLA HÖSSJER

Lund University, Sweden

(Received 22 December 1998; In final form 16 May 2000)

We consider relations between Rényi's and Hentschel–Procaccia's definitions of generalized dimensions of a probability measure μ, and give conditions under which the two concepts are equivalent/different. Estimators of the dimension spectrum are developed, and strong consistency is established. Particular cases of our estimators are methods based on the sample correlation integral and box counting.

Then we discuss the relation between generalized dimensions and kernel density estimators $\hat{f}$. It was shown in Frigyesi and Hössjer (1998), that $\int \hat{f}^{1+q}(x)dx$ diverges with increasing sample size and decreasing bandwidth if the marginal distribution μ has a singular part and $q > 0$. In this paper, we show that the rate of divergence depends on the qth generalized Rényi dimension of μ.

Keywords and Phrases: Kernel density estimates; Fractal dimension estimation; Generalized dimensions; Rényi dimension; Hentschel–Proccacia dimension; Correlation integral; Box counting

1. INTRODUCTION

Let μ be a probability measure on the Borel subsets $\mathcal{B}(\mathbb{R}^p)$ on $\mathbb{R}^p$. The concepts of generalized Rényi dimension and dimension spectrum of μ were introduced by Rényi (1970). The dimension spectrum measures

*Research sponsored in part by Office of Naval Research Grant N00014 93 1 0043.
†Corresponding author.

not only the fractal dimension of supp(μ), but also the extent to which μ has mass concentrations on sets of smaller dimension.

Rényi's definition (often referred to as the information-theoretic approach) is as follows: Define, for each $i=(i_1,\ldots,i_p)\in\mathbb{Z}^p$ and $h\in(0,\infty)$, the hypercube $\lambda_{hi}=(i_1h,(i_1+1)h]\times\cdots\times(i_ph,(i_p+1)h]$ of side length h. Thus $\Gamma_h=\{\lambda_{hi}\}_{i\in\mathbb{Z}^p}$ is a partition of $\mathbb{R}^p$. With $I(A)$ the indicator of the set A, let $\Gamma_h(x,y)=\sum_i I((x,y)\in\lambda_{hi}\times\lambda_{hi})$ signify whether x and y belong to the same hypercube.[1] Then introduce the functional[2]

$$\Lambda_\mu(q;\Gamma_h)=\begin{cases}\exp(\int\log(\int\Gamma_h(x,y)d\mu(y))d\mu(x)), & q=0,\\ (\int(\int\Gamma_h(x,y)d\mu(y))^q d\mu(x))^{1/q}, & q\neq 0.\end{cases}\qquad(1)$$

Using this functional, which, viewed as a function of q, measures how the probability mass on scale h varies with location, we arrive at:

Definition 1 The qth generalized Rényi dimension of μ is defined as

$$d_\mu(q;\Gamma)=\lim_{h\to 0+}\frac{\log\Lambda_\mu(q;\Gamma_h)}{\log h},\quad q\in\mathbb{R},\qquad(2)$$

whenever the limit exists. If the limit does not exist we instead consider lim inf and lim sup and thus define $d_\mu^-(q;\Gamma)$ and $d_\mu^+(q;\Gamma)$.

Here Γ represents the sequence of kernels $\{\Gamma_h\}_{h>0}$. If the dimension spectrum exists and is nonconstant, μ is called a *multifractal measure*.

A somewhat different formulation of these ideas was given in Hentschel and Procaccia (1983). First define the functional $\Lambda_\mu(q;\Psi_h)$ as in (1), replacing Γ_h with the kernel $\Psi_h(x,y)=I(\|x-y\|\leq h)$.[3] Then, in analogy with Definition 1, we get:

Definition 2 The qth Hentschel and Procaccia (HP) dimension is defined by

$$d_\mu(q;\Psi)=\lim_{h\to 0+}\frac{\log\Lambda_\mu(q;\Psi_h)}{\log h},\quad q\in\mathbb{R},\qquad(3)$$

[1]We use Γ_h to denote both the grid defined by the hypercubes *and* the associated kernel function. Since there is a one-to-one correspondence between the two concepts, we hope this will cause no confusion.

[2]We use the convention $\log 0=0$ and $0^q=0$, $\forall q\in\mathbb{R}$.

[3]Throughout the paper, $\|x\|=\max(|x_1|,\ldots,|x_p|)$ denotes the supremum norm of a vector.

whenever the limit exists. If the limits do not exist we instead consider lim inf and lim sup and thus define $d_\mu^-(q;\Psi)$ and $d_\mu^+(q;\Psi)$.

An advantage of the HP dimension spectrum is that no grid needs to be specified. In Section 2, we examine the relation between the Rényi and HP definitions of generalized dimensions. It turns out that the two concepts are identical when $q > -1$, but for $q \leq -1$, this is not the case.

How do we interpret the dimension spectrum? Letting $B(x,h) = \{y; \|y-x\| \leq h\}$ denote the closed hypercube with center x and side lengths $2h$, and $\mu_h(x) = \mu(B(x,h))$, we find that

$$\mu_h(x) = \int \Psi_h(x,y)d\mu(y),$$

and consequently,

$$E(\mu_h(X)^q) \approx h^{qd_\mu(q;\Psi)},$$

where X is a random variable with distribution μ. The exponent $d_\mu(q;\Psi) \in [0,\infty]$ characterizes a measure in different ways when q is varied. When $q < 0$, $d_\mu(q)$ is affected by the more rarified regions of the measure ($\mu_h(x)$ small).[4] In particular, $d_\mu(-\infty)$ is determined by the *most rarified* region of μ and $d_\mu(-1)$ is closely related to the fractal dimension of the support of μ.

Analogously, $d_\mu(q)$ is determined by the denser regions of the measure when $q > 0$ ($\mu_h(x)$ large). In particular $d_\mu(\infty)$ is affected by the *densest* region and $d_\mu(1)$ gives the well known correlation dimension

$$d_\mu(1) = \lim_{h\to 0} \frac{\log P(\|X-Y\| \leq h)}{\log h},$$

where X and Y are two independent random variables with distribution μ, and $\Lambda_\mu(1;\Psi_h) = P(\|X-Y\| \leq h) := C(h)$ is referred to as the correlation integral. Finally, $d_\mu(0)$ is most easily described by introducing

$$\alpha_\mu(x) = \lim_{h\to 0} \frac{\log \mu_h(x)}{\log h},$$

[4]When discussing generalized dimensions in general, we will often omit the second argument of d_μ, which, in view of the results in Section 2, is no restriction when $q > -1$.

the pointwise dimension of μ at x. Then, under certain regularity conditions,

$$d_\mu(0) = E\alpha_\mu(X),$$

cf. Cutler (1993), Theorem 3.3.10. That is, $d_\mu(0)$ equals the average point-wise dimension, which has contributions from both rarified and dense regions of μ.

One has, as a simple consequence of Jensen's inequality, that

$$d_\mu(q') \leq d_\mu(q) \quad \text{for } q' > q. \tag{4}$$

Beck (1990) proved that multifractals cannot be arbitrarily non-uniform, in that

$$d_\mu(r) \geq \frac{r+1}{r}\frac{s}{s+1}d_\mu(s),$$

for $r > s > 0$ and $-1 > r > s$. Multifractals with dimension spectra obeying $d_\mu(q) = d_\mu(q')$ are termed *maximum uniform*, whereas a scaling $sd_\mu(s)/(s+1) = rd_\mu(r)/(r+1)$ is termed *minimum uniform*.

According to the *multifractal formalism*, dimension spectra are intimately related with the widely used singularity spectra in physics through the Legendre transform (under certain regularity conditions). This has increased the importance of having good estimates of dimension spectra from data. Suppose we have observations $X_1, \ldots, X_n \in \mathbb{R}^p$ from a stationary and ergodic stochastic process with marginal distribution μ. Based on $\{X_i\}_1^n$, we will develop estimates of the dimension spectra. Natural 'plug-in' estimators of $\Lambda_\mu(q; \Gamma_h)$ and $\Lambda_\mu(q; \Psi_h)$ are presented in Section 4, essentially replacing μ with the empirical distribution. We prove consistency of these estimators, and the accompanying dimension estimators are treated in Section 5.

Somewhat surprisingly, there is a link between the abovementioned estimators and kernel density estimators, defined as

$$\hat{f}(x) = \frac{1}{nh^p}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right), \tag{5}$$

with K is a non-negative kernel function integrating to one and h the bandwidth. Usually, $\hat{f}$ is regarded as an estimator of a density f, when $d\mu = fdx$ is absolutely continuous (*cf.* *e.g.*, Wand and Jones, 1995 for an introduction to kernel methods applied to density estimation). However, it was found in Frigyesi and Hössjer (1998), that $\hat{f}$ makes sense even when μ has a singular part.[5] For instance, if $q > 0$, the functional $\int \hat{f}^{1+q}(x)dx$ increases to infinity as $n \rightarrow \infty$ and $h \rightarrow 0$, when μ has a singular part. On the other hand, $\int \hat{f}^{1+q}(x)dx$ converges to $\int f^{1+q}(x)dx$ for absolutely continuous μ. This was used by Frigyesi and Hössjer (1998) to devise a test, which discriminates between measures having a singular part and a large class of absolutely continuous measures (*i.e.*, those with $\int f^{1+q}(x)dx \leq C < \infty$). Using some properties of a new functional Λ'_μ (closely related to Λ_μ), we find in Section 6 that the speed of divergence of $\int \hat{f}^{1+q}(x)dx$ is determined by the qth generalized Rényi dimension. Similarly, one can replace kernel estimates with histogram estimators. The corresponding functional will diverge at the same speed.

We classify our estimator of $\Lambda_h(q; \Psi_h)$ as a kernel method, whereas the estimator of $\Lambda_h(q; \Gamma_h)$ can be viewed as a histogram method. Well known estimators fall into our framework, *e.g.*, box counting is a histogram method with $q = -1$, and the sample correlation integral a kernel method with $q = 1$.

We advocate the kernel methods, since they do not depend on the choice of a grid. When dealing with convergence rates of the estimators, the kernel methods seem to yield more suitable expressions for asymptotic distributions and confidence bands. We hope to explore this in a forthcoming paper. *Cf.* Denker and Keller (1986), who used U-statistics theory for dependent data as a tool for obtaining the asymptotic distribution of the sample correlation integral.

Other nonparametric techniques have also been used in connection with fractal dimension estimation, such as wavelets (*cf. e.g.*, Muzy *et al.*, 1994) and nearest neighbour methods (van de Water and Scram, 1988).

Fractal dimensions (specifically the correlation dimension) have been used to estimate the embedding dimension of nonlinear time series (*cf. e.g.*, Cutler, 1994). On the other hand, Cheng and Tong (1992) argue that the embedding dimension should be estimated *before* the fractal dimension of the attractor. They propose a method

[5]By μ having a singular part we mean the existence of a set $A \in \mathcal{B}(\mathbb{R}^p)$ such that $\mu(A) > 0$ and $\lambda(A) = 0$, with $d\lambda(x) = dx$ the Lebesgue measure.

involving nonparametric regression kernel smoothing and cross validation for estimating the embedding dimension.

For an overview of fractal dimension estimation, we refer to Cutler (1993), whereas much of the mathematical theory of fractals can be found in Falconer (1997).

2. PARTIAL EQUIVALENCE OF HP AND RÉNYI DEFINITIONS

In this section, we will investigate under which circumstances the generalized HP and Rényi dimensions agree. As a preparation for this, we need to introduce some regularity concepts of probability measures. The first such concept, introduced by Federer (1969), is frequently used as a regularity condition on measures in multifractal analysis:

DEFINITION 3 (Diametric regularity.) A measure μ is *diametrically regular* or a *Federer measure* if, for each $A > 1$, there exists a constant $c > 0$, such that for all sufficiently small $h > 0$ and every x we have

$$\mu_h(x) \geq c\mu_{Ah}(x) \tag{6}$$

It is easily seen that it suffices to verify (6) for one particular value of $A > 1$. However, diametrically regularity is not well suited for measures defined on $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$, as the following result shows:

PROPOSITION 1 *Any compactly supported measure on* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ *is not diametrically regular.*

The proof of Proposition 1 is based on finding points x slightly *outside* $\text{supp}(\mu)$ for which (6) fails. We will find below (Example 2) measures violating (6) even if we consider points $x \in \text{supp}(\mu)$ only.

Let us now introduce weaker versions of diametric regularity that will be of interest for our subsequent comparison between generalized Rényi and HP dimensions. Given $A > 1$ and $c > 0$, define

$$\Omega_{ch}^A = \{x; \mu_h(x) > c\mu_{Ah}(x)\}.$$

Clearly, diametric regularity means that for some $c > 0$ and all $h > 0$ sufficiently small, $\Omega_{ch}^A = \mathbb{R}^p$.

DEFINITION 4 (Weak diametric regularity in HP sense.) Given $q < 0$ and $A > 1$, a measure is referred to as weakly diametrically regular in the (q, A)th HP sense if

$$\lim_{h \to 0} \frac{\log\left(\int \mu_h(x)^q d\mu(x) / \int_{\Omega_{ch}^A} \mu_h(x)^q d\mu(x)\right)}{\log h} = 0.$$

for some function $c = c(h) > 0$ such that $\log c/\log h \to 0$ as $h \to 0$.

Definition 4 applies to a constant function $c(h) \equiv c > 0$, but it also allows for $c \to 0$ (slowly) as $h \to 0$. For technical reasons, a Rényi variant of Definition 4 will be more useful for us. To begin with, it is straightforward to check that an equivalent formulation of the Rényi type functional is

$$\Lambda_\mu(q; \Gamma_h) = \begin{cases} \left(\sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q}\right)^{1/q}, & q \neq 0, \\ \exp\left(\sum_{\gamma \in \Gamma_h} \mu(\gamma) \log(\mu(\gamma))\right), & q = 0. \end{cases} \tag{7}$$

To each $\gamma \in \Gamma_h$ we may associate a vector $j = (j_1, \ldots, j_p) \in \mathbb{Z}^p$, with $(j_i, j_i + 1]h$ being the projection of γ onto the ith coordinate axis. We may hence define a distance between two cubes in Γ_h as $\|\gamma - \gamma'\| = \max_{1 \leq i \leq p} |j_i - j_i'|$, if $j' = (j_1', \ldots, j_p')$ is the vector corresponding to γ'. Given any positive integer m, we associate a neighbourhood

$$U_m(\gamma) = \{\gamma' \in \Gamma_h;\ \|\gamma' - \gamma\| \leq m\},$$

of all cubes in Γ_h with distance at most m from γ. In particular, for $m = 1$, we will write $U_1(\gamma) = U(\gamma)$. Then, for any $c > 0$, we introduce

$$\Gamma_{ch}^m = \left\{\gamma \in \Gamma_h; \mu(\gamma) > c \sum_{\gamma' \in U_m(\gamma)} \mu(\gamma')\right\}.$$

In particular we write $\Gamma_{ch}^1 = \Gamma_{ch}$ when $m = 1$. Thus all hypercubes in Γ_{ch} have neighbours to which μ assign probability at most of the same order as to γ.

DEFINITION 5 (Weak diametric regularity in Rényi sense.) Given $q < 0$ and a positive integer m, a measure is referred to as weakly diametrically regular in the (q,m)th Rényi sense if

$$\lim_{h\to 0} \frac{\log\left(\sum_{\gamma\in\Gamma_h}\mu(\gamma)^{1+q}/\sum_{\gamma\in\Gamma^m_{ch}}\mu(\gamma)^{1+q}\right)}{\log h} = 0 \tag{8}$$

for some function $c = c(h) > 0$ such that $\log c/\log h \to 0$ as $h \to 0$.

The next lemma shows that the concept of weak diametric regularity is meaningful only for $q \le -1$:

LEMMA 1 (Weak diametric regularity for $-1 < q < 0$.) *Suppose $-1 < q < 0$. Then all measures are weakly diametrically regular in both the (q,m)th Rényi (with $m \ge 1$ an integer) and (q,A)th HP sense (with $A > 1$).*

The relation between the various versions of diametric regularity when $q \le -1$ is furnished in the next proposition. We use the abbreviated notation DR for diametric regularity, $\mathrm{DR}^{\mathrm{HP}}_{(q,A)}$ for weak diametric regularity in the (q,A)th HP sense and so on.

PROPOSITION 2 (Relations between various types of diametric regularity.) *The following implications hold between the various types of diametric regularity when $q \le -1$:*

$$\begin{aligned} DR &\Rightarrow DR^{HP}_{(q,A)}, && \text{if } A > 1,\\ DR &\Rightarrow DR^{Rényi}_{(q,m)}, && \text{if } m \ge 1,\\ DR^{HP}_{(q,A_2)} &\Rightarrow DR^{HP}_{(q,A_1)}, && \text{if } A_2 > A_1 > 1,\\ DR^{Rényi}_{(q,m_2)} &\Rightarrow DR^{Rényi}_{(q,m_1)}, && \text{if } m_2 > m_1 \ge 1,\\ DR^{Rényi}_{(q,m)} &\Rightarrow DR^{HP}_{(q,A)}, && \text{if } A > 1 \text{ and } \lceil A\rceil \le m, \end{aligned} \tag{9}$$

where $\lceil A\rceil$ denotes the smallest integer greater than or equal to A.

We are now ready to formulate the main result of this section:

THEOREM 1 (Relation between Rényi and HP dimensions.) *It holds that*

$$d^{\pm}_{\mu}(q;\Psi) = d^{\pm}_{\mu}(q;\Gamma), \quad q > -1, \tag{10}$$

$$d^{\pm}_{\mu}(q;\Psi) \le d^{\pm}_{\mu}(q;\Gamma), \quad q \le -1. \tag{11}$$

We have equality in (11) *for measures that are weakly diametrically regular in the* $(q, 1)$*th Rényi sense.*

Remark 1 (Choice of norm.) In the proof of Theorem 1 we use the max norm. However, since any two norms in $\mathbb{R}^p$ are strongly equivalent, there exist numbers $0 < a_1 \leq a_2$ such that

$$a_1 \|x\| \leq \|x\|_I \leq a_2 \|x\|$$

for all $x \in \mathbb{R}^p$, where $\|\cdot\|$ is the max norm and $\|\cdot\|_I$ any other norm. For instance, the Euclidean norm has $a_1 = 1$ and $a_2 = \sqrt{p}$. It follows that[6]

$$\mu_{h/a_2}(x) \leq \mu_h^I(x) \leq \mu_{h/a_1}(x)$$

and $d_\mu^{I\pm}(q; \Psi) = d_\mu^{\pm}(q; \Psi)$. In Proposition 2, only the last condition of (9) has to be changed to $\lceil Aa_2/a_1 \rceil \leq m$.

Remark 2 Since the original version of this paper Frigyesi and Hössjer (1996) was written, results related to Theorem 1 have appeared. Pesin (1997) proves (10) when $q > 0$ and, for diametrically regular measures, even when $q > -1$. Pesin, as well as Guerin (1998), also consider other functionals than $\Lambda_\mu(\cdot\,; \Gamma_h)$ and $\Lambda_\mu(\cdot\,; \Psi_h)$, all of which give equivalent generalized dimensions when $q > 0$.

For computational purposes, it is convenient to treat sequences of r-adic grids for some $r \in \{2, 3, 4, \ldots\}$, with cubes of the form

$$(i_1 r^{-m}, (i_1 + 1) r^{-m}] \times \cdots \times (i_p r^{-m}, (i_p + 1) r^{-m}].$$

The following result gives conditions which allow restriction to r-adic grids, when computing HP-dimensions:

COROLLARY 1 *For* $q > -1$,

$$\begin{aligned} d_\mu^+(q; \Psi) &= \limsup_{m \to \infty} \frac{\log(\Lambda_\mu(q; \Gamma_{r^{-m}}))}{\log r^{-m}} \\ d_\mu^-(q; \Psi) &= \liminf_{m \to \infty} \frac{\log(\Lambda_\mu(q; \Gamma_{r^{-m}}))}{\log r^{-m}}. \end{aligned} \tag{12}$$

[6]Superscript I indicates that $\|\cdot\|_I$ is being used.

If further

$$\lim_{m\to\infty} \frac{\log\left(\sum_{\gamma\in\Gamma_{cr^{-m}}} \mu(\gamma)^{1+q} / \sum_{\gamma\in\Gamma_{r^{-m}}} \mu(\gamma)^{1+q}\right)}{\log r^{-m}} = 0, \tag{13}$$

for some function $c = c(r^{-m}) > 0$ *such that* $\log c/m \to 0$ *as* $m \to \infty$, *then* (12) *holds for* $q \le -1$ *as well.*

We illustrate, with three examples, how to use r-adic grids for computing HP dimensions. Although the examples are well known, Corollary 1 allows us to find the qth generalized HP dimension for any $q \in \mathbb{R}$. The Rényi grids $\Gamma_{r^{-m}}$ are just introduced to facilitate computations.

Example 1 (Uniform Measure on the Cantor Set.) Let μ be the uniform Cantor measure on $[0, 1]$ and $\{\lambda_i\}_{i=1}^{3^m}$ the intervals of $\Gamma_{3^{-m}}$ that intersect $(0, 1]$. Then, using (7), we get

$$\Lambda_\mu(q;\Gamma_{3^{-m}}) = \begin{cases} \sum_{i=1}^{3^m} \mu(\gamma_i)^{q+1} = 2^m(2^{-m})^{q+1} = (3^{-m})^{q\log 2/\log 3}, & q \neq 0, \\ \sum_{i=1}^{3^m} \mu(\gamma_i)\log(\mu(\gamma_i)) = \log(2^{-m}), & q = 0. \end{cases} \tag{14}$$

Since $\Gamma_{c3^{-m}} = \Gamma_{3^{-m}}$ for small enough c, Corollary 1 implies, $d_\mu(q;\Psi) = \log 2/\log 3$, $\forall q \in \mathbb{R}$. The uniform measure on the Cantor set is thus maximum uniform.

Example 2 (Generalized Cantor Distributions (GCD).) Consider a vector $w = (v_1, \ldots, v_{r^p})$, $v_i \ge 0$ with $\sum_{i=1}^{r^p} v_i = 1$. Define μ_w on $[0, 1]^p$ by recursively choosing cubes in $\Gamma_{r^{-1}}, \Gamma_{r^{-2}}, \ldots$ according to the $\text{Mult}(1; v_1, \ldots, v_{r^p})$ multinomial distribution. By varying w we get a large class of measures, the Generalized Cantor Distributions. For $p=1$, $r=3$, $w=(1/2, 0, 1/2)$ we get the classic Cantor Distribution on $[0, 1]$. The case $p=1$, $r=2$ and $w=(\rho, 1-\rho)$, is investigated in Cutler (1993), Example 3.3.16. In general, it is easily shown that

$$\lim_{m\to\infty} \frac{\log(\Lambda_\mu(q;\Gamma_{r^{-m}}))}{\log r^{-m}} = \begin{cases} -(1/q)\log\left(\sum_{i=1}^{r^p} v_i^{q+1}\right)/\log r, & q \neq 0 \\ \left(-\sum_{i=1}^{r^p} v_i \log v_i\right)/\log r, & q = 0. \end{cases} \tag{15}$$

In the chapter containing the proofs, we show that (13) holds. Thus $d_\mu(q;\Psi)$ is given by the RHS of (15) for all $q \in \mathbb{R}$.

Assuming $p=1$, $r=2$, $w=(\rho, 1-\rho)$, with $0<\rho<1/2$, there exists an interval γ belonging to $(0,1] \cap \Gamma_{2^{-m}}$ such that $\mu(\gamma)=(1-\rho)\rho^{m-1}$ and $\gamma' \in U(\gamma)$ with $\mu(\gamma')=(1-\rho)^{m-1}\rho$. Thus, with $h=2^{-m-1}$ and x the midpoint of γ we have

$$\mu_h(x) < c\mu_{3h}(x), \quad \text{with } c = \left(\frac{\rho}{1-\rho}\right)^{m-2}.$$

Since c can be made arbitrarily small by letting $m \to \infty$ and $\text{supp}(\mu)=[0,1]$, we have found that μ is not DR even if we restrict consideration in (6) to points $x \in \text{supp}(\mu)$.

Example 3 (Absolutely Continuous Measures with Nontrivial Dimension Spectrum.) Consider the following density: $\partial\mu/\partial\lambda(x) = (1+\beta)x^\beta$, $x \in [0,1]$, for some constant $\beta \in (-1,\infty)$. Partitioning $(0,1]$ into cells $\gamma_i=((i-1)r^{-m}, ir^{-m}]$ for $i=1,\ldots,r^m$, we get

$$\mu(\gamma_i) = \int_{\gamma_i} (1+\beta)x^\beta dx = r^{-m(\beta+1)}(i^{\beta+1} - (i-1)^{\beta+1}).$$

Clearly, $\Gamma_{cr^{-m}} = \Gamma_{r^{-m}}$ for small enough c, so $d_\mu(q;\Psi)$ can be obtained from (12). Notice first that

$$\begin{aligned}\sum_{i=1}^{r^m} \mu(\gamma_i)^{q+1} &= r^{-m(q+1)(\beta+1)} \sum_{i=1}^{r^m} (i^{\beta+1} - (i-1)^{\beta+1})^{q+1} \\ &:= r^{-m(q+1)(\beta+1)} S\end{aligned}$$

for any $q \in \mathbb{R}$. This gives

$$S \underset{\cap}{\cup} \sum_{i=1}^{r^m} (i^\beta)^{q+1} \underset{\cap}{\cup} \begin{cases} 1, & \beta(q+1) < -1 \\ \log r^m, & \beta(q+1) = -1 \\ r^{m(\beta(q+1)+1)}, & \beta(q+1) > -1. \end{cases}$$

where $a_m \underset{\cap}{\cup} b_m$ means that both a_m/b_m and b_m/a_m are bounded sequences as $m \to \infty$. We thus have

$$d_\mu(q;\Psi) = \lim_{m\to\infty} \frac{(q+1)(\beta+1)\log r^{-m} + \log S}{q \log r^{-m}}, \quad q \neq 0,$$

and the case $q=0$ is handled similarly. Depending on the value of β, we get three cases (*cf.* also Beck, 1990)

$$d_\mu(q;\Psi) = \begin{cases} (q+1)(\beta+1)/q, & q \geq -(1/\beta)-1, \\ 1, & q \leq -(1/\beta)-1, \end{cases} \quad -1<\beta<0,$$

$$d_\mu(q;\Psi) = 1, \ \beta = 0,$$

and

$$d_\mu(q;\Psi) = \begin{cases} (q+1)(\beta+1)/q, & q \leq -(1/\beta)-1, \\ 1, & q \geq -(1/\beta)-1, \end{cases} \quad \beta>0.$$

We thus have a minimum uniform absolutely continuous measure.

From Example 3 we find that the dimension spectrum may be non-trivial, *i.e.*, non-maximum uniform, for absolutely continuous distributions. In Cutler (1993), it is proved that absolutely continuous probability measures that are bounded from above and from zero are maximum uniform with $d_\mu(q) \equiv p$.

Are there any measures for which the Rényi and HP dimensions are different when $q \leq -1$? The answer is yes, for certain measures that are not weakly DR in the $(q,1)$th Rényi sense. We will exemplify two such measures on $\mathbb{R}^p$, one continuous and one discrete:

Example 4 We will construct a continuous measure on $[0,1]$ related to the GCD measures of Example 2 which is not $\mathrm{DR}^{\text{Rényi}}_{(-1,1)}$. Given $r \in \{2,3,\dots\}$, a rapidly increasing sequence $m_k \to \infty$ and another sequence $c_k \searrow 0$ with $\log c_k / \log r^{-m_k} = \xi$ and $0<\xi<1$, we proceed as follows: First choose cube of $\Gamma_{r^{-m_1}}$ according to $\mathrm{Mult}(1;w_1)$, where $w_1 = \mathrm{const}\cdot(1, c_1^{-1}, \dots, c_1^{-r^{m_1}+1})$ and const assures that w_1 is a probability vector. Then proceed recursively, so that given an interval $\gamma \in \Gamma_{r^{-m_{k-1}}}$ subintervals $\Gamma_{r^{-m_k}} \ni \gamma' \subset \gamma$ are chosen according to a $\mathrm{Mult}(1;w_k)$ distribution, where $w_k = \mathrm{const}\cdot(1, c_k^{-1}, \dots, c_k^{-r^{m_k-m_{k-1}}+1})$.

With

$$\Gamma_h^+ := \{\gamma \in \Gamma_h; \mu(\gamma) > 0\}, \tag{16}$$

we notice that $N_h^+ := \sharp\Gamma_h^+ = \lceil 1/h \rceil$, so that $d_\mu(-1;\Gamma) = \lim_{h\to 0} (\log N_h^+ / \log h^{-1}) = 1$.

Given $\gamma \in \Gamma_{r^{-m_k}}$ we let right (γ) denote the interval in $\Gamma_{r^{-m_k}}$ immediately to the right of γ. Then introduce $(k \geq 2)$

$$\Gamma'_{r^{-m_k}} = \{\gamma \in \Gamma_{r^{-m_k}} \cap (0,1]; \gamma \text{ and right } (\gamma) \text{ belong to different intervals of } \Gamma_{r^{-m_{k-1}}}\}.$$

Assume $\gamma \notin \Gamma'_{r^{-m_k}}$ and let $h_k = 2r^{-m_k}$. Since $B(x, h_k) \supseteq \gamma \cup \text{right}(\gamma)$ if $x \in \gamma$ and $\mu(\gamma) = c_k \mu(\text{right}(\gamma))$, we get

$$\int_\gamma \frac{d\mu(x)}{\mu_{h_k}(x)} \leq (1 + c_k^{-1})^{-1} \int_\gamma \frac{d\mu(x)}{\mu(\gamma)} = (1 + c_k^{-1})^{-1}.$$

Let $\|\mu_h\|_q = (\int \mu_h(x)^q d\mu(x))^{1/q}$ for $q \neq 0$. Then

$$\begin{aligned} \|\mu_{h_k}\|_{-1}^{-1} &= \left(\sum_{\gamma \in \Gamma'_{r^{-k}}} + \sum_{\gamma \notin \Gamma'_{r^{-k}}} \right) \int_\gamma \frac{d\mu(x)}{\mu(B(x, h_k))} \\ &\leq \sharp \Gamma'_{r^{-m_k}} + N^+_{r^{-m_k}} (1 + c_k^{-1})^{-1} \leq r^{m_{k-1}} + r^{m_k} c_k \sim r^{m_k} c_k, \end{aligned}$$

with $a_k \sim b_k$ meaning that $a_k / b_k \to 1$ as $k \to \infty$, and the last step follows by our definition of $\{m_k\}$ and $\{c_k\}$. Thus

$$\begin{aligned} d_\mu^-(-1; \Psi) &= \liminf_{h \to 0} \frac{\log \|\mu_h\|_{-1}}{\log h} \leq \liminf_{k \to \infty} \frac{\log \|\mu_{h_k}\|_{-1}^{-1}}{\log h_k^{-1}} \\ &\leq \liminf_{k \to \infty} \frac{\log (r^{m_k} c_k)}{\log (r^{m_k}/2)} = 1 - \xi < 1. \end{aligned}$$

In conclusion, we have verified that $d_\mu^-(-1; \Psi) < d_\mu(-1; \Gamma)$. According to Theorem 1, μ cannot be a $\mathrm{DR}^{\text{Rényi}}_{(-1,1)}$ measure. This can also be seen directly, since $\Gamma^1_{c_k r^{-m_k}} \subset \Gamma'_{r^{-m_k}}$.

Example 5 Let us now construct a discrete measure on $[-1, 0]$ which is not $\mathrm{DR}^{\text{Rényi}}_{(q,1)}$ for $q < -1$. We put

$$\mu = \sum_{k=1}^\infty p_k \delta_{-4^{-k}} + \sum_{k=1}^\infty s_k \delta_{-4^{-k} + 8^{-k}},$$

where δ_x is a one point distribution at x, $p_k = 2^{-k-1}$ and $s_k = 3k^{-2}\pi^{-2}$ so that $\mu[-1/4, 0] = 1$. Note that $-4^{-1}, \ldots, -4^{-m}$, $-4^{-1} + 8^{-1}, \ldots, -4^{-m+1} + 8^{-m+1}$ all lie in different $\gamma \in \Gamma_{4^{-m}}$, whereas all remaining atoms of μ belong to the same interval $\gamma = (-4^{-m}, 0]$.

Thus, since $q < -1$,

$$\sum_{\gamma \in \Gamma_{4^{-m}}} \mu(\gamma)^{1+q} = \sum_1^m p_k^{q+1} + \sum_1^{m-1} s_k^{q+1} + \left(\sum_{k=m+1}^{\infty} p_k + \sum_{k=m}^{\infty} s_k \right)^{q+1}$$
$$\underset{\cap}{\cup} 2^{-m(q+1)} + m^{-2q-1} + m^{-q-1} \underset{\cap}{\cup} 2^{-m(q+1)},$$

giving $d_\mu^+(q;\Gamma) \geq \lim_{m\to\infty}(\log \Lambda_\mu(q;\Gamma_{4^{-m}})/\log 4^{-m}) = (q+1)/(2q)$. Turning to $d_\mu(q;\Psi)$, we notice that

$$\int \mu_{4^{-m}}(x)^q d\mu(x) = \sum_{k=1}^{\infty} p_k \mu_{4^{-m}}(-4^{-k})^q + \sum_{k=1}^{\infty} s_k \mu_{4^{-m}}(-4^{-k} + 8^{-k})^q,$$

where

$$\mu_{4^{-m}}(-4^{-k}) = \begin{cases} p_k, & k = 1, \ldots, [(m-1)/2], \\ p_k + s_k, & k = [(m-1)/2] + 1, \ldots, m-1, \\ \sum_{l=m}^{\infty}(p_l + s_l), & k \geq m, \end{cases}$$

and

$$\mu_{4^{-m}}(-4^{-k} + 8^{-k}) = \begin{cases} s_k, & k = 1, \ldots, [(m-1)/2], \\ p_k + s_k, & k = [(m-1)/2] + 1, \ldots, m-1, \\ \sum_{l=m}^{\infty}(p_l + s_l), & k \geq m. \end{cases}$$

combining the last three displays we find

$$\begin{aligned} \int \mu_{4^{-m}}(x)^q d\mu(x) = & \sum_1^{[(m-1)/2]} p_k^{q+1} + \sum_1^{[(m-1)/2]} s_k^{q+1} \\ & + \sum_{k=[(m-1)/2]+1}^{m-1} (s_k + p_k)^{q+1} \\ & + \left(\sum_{k=m}^{\infty} (s_k + p_k) \right)^{q+1} \underset{\cap}{\cup} 2^{-m(q+1)/2}. \end{aligned}$$

Now, use (38) to deduce $d_\mu(q;\Psi) = (q+1)/(4q)$. Thus we have $d_\mu(q;\Psi) < d_\mu^+(q;\Gamma)$. In view of Theorem 1, this implies that μ is not a $\mathrm{DR}_{(q,1)}^{\text{Rényi}}$ measure.

The last two examples can be summarized as follows:

THEOREM 2 *It is possible to construct measures μ such that*

$$d_\mu^-(-1;\Psi) < d_\mu(-1;\Gamma), \tag{17}$$

$$d_\mu(q;\Psi) < d_\mu^+(q;\Gamma), \quad q < -1. \tag{18}$$

We close this section by considering the box dimension:

DEFINITION 6 The *box dimension* of a set A is defined as

$$\Delta(A) = \lim_{m\to\infty} \frac{\log N_h(A)}{\log h^{-1}},$$

whenever the limit exists. Here $N_h(A)$ is the number of elements from Γ_h needed to cover A.

Using Theorem 1, we may find measures for which the box dimension of supp(μ) differs from the $q = -1$ HP dimension.

COROLLARY 2 *For any measure μ we have*

$$\Delta(supp(\mu)) = d_\mu(-1;\Gamma), \tag{19}$$

and in particular, we can find a measure μ such that

$$d_\mu^-(-1;\Psi) \le \Delta(supp(\mu)).$$

3. A NON-GRID RÉNYI TYPE FUNCTIONAL

We now define a new functional, that will be further discussed in the next section in connection with density estimation, but it is also of independent interest. It has the form

$$\Lambda'_\mu(q;\Psi_h) = \begin{cases} \exp(h^{-p}\int(\int\Psi_h(x,y)d\mu(y))\log(\int\Psi_h(x,y)d\mu(y))dx), & q=0, \\ h^{-p/q}(\int(\int\Psi_h(x,y)d\mu(y))^{1+q}dx)^{1/q}, & q\neq 0. \end{cases}$$

$$= \begin{cases} \exp(h^{-p}\int\mu_h(x)\log\mu_h(x)dx), & q=0, \\ h^{-p/q}(\int\mu_h(x)^{1+q}dx)^{1/q}, & q\neq 0. \end{cases}$$

Now $\Lambda'_\mu(q;\Psi_h)$ is closely related to the Rényi functional $\Lambda_\mu(q;\Gamma_{2h})$. In fact,

$$\Lambda_\mu(q;\Gamma_{2h}) = \Lambda'_\mu(q;\Gamma_{2h}) = \begin{cases} \exp((2h)^{-p}\int \mu_h(y(x))\log\mu_h(y(x))dx), & q=0, \\ (2h)^{-p/q}(\int \mu_h(y(x))^{1+q}dx)^{1/q}, & q\neq 0, \end{cases}$$

where $y(x)$ is the midpoint of the unique cube $\gamma(x)\in\Gamma_{2h}$ such that $x\in\gamma(x)$. Except for the unimportant scaling factor 2^{-p}, $\Lambda_\mu(q;\Gamma_{2h})$ differs from $\Lambda'_\mu(q;\Psi_h)$ through the replacement of x by $y(x)$ in the integrals. Thus, $\Lambda'_\mu(q;\Psi_h)$ can be viewed as a non-grid version of the Rényi functional $\Lambda_\mu(q;\Gamma_{2h})$. Moreover, $\Lambda'_\mu(q;\Psi_h)$ has links to the HP-functional, since we obtain $\Lambda'_\mu(q;\Psi_h)$ by replacing $d\mu(x)$ with a smoothed version $h^{-p}\mu_h(x)dx$ in the definition of $\Lambda_\mu(q;\Psi_h)$. Guerin (1998, Section 1.2) also considers $\Lambda'_\mu(q;\Psi_h)$ when $q>0$.

The fractal dimension corresponding to $\Lambda'_\mu(q;\Psi_h)$ is

$$d'_\mu(q;\Psi) = \lim_{h\to 0}\frac{\log\Lambda'_\mu(q;\Psi_h)}{\log h},$$

if the limit exists. Similarly, we define $d'^{\pm}_\mu(q;\Psi)$ as the lim inf or lim sup of the RHS. The relation between this dimension concept and the Rényi and HP dimension is treated in the following theorem:

THEOREM 3 *For any probability measure μ on $\mathcal{B}(\mathcal{R}^p)$, the relation between $d'^{\pm}_\mu(q;\Psi)$ and the Rényi dimension is such that*

$$d'^{\pm}_\mu(q;\Psi) = d^{\pm}_\mu(q;\Gamma), \quad q\geq -1, \tag{20}$$

whereas for $q<-1$, there exist measures with $d'^{+}_\mu(q;\Psi) < d^{+}_\mu(q;\Gamma)$ as well as those with $d'^{\pm}_\mu(q;\Psi) > d^{\pm}_\mu(q;\Gamma)$.

The dimension $d'^{\pm}_\mu(q;\Psi)$ is further related to the HP dimension as

$$d'^{\pm}_\mu(q;\Psi) \geq d^{\pm}_\mu(q;\Psi), \quad q<-1. \tag{21}$$

and there exist measures with strict inequality in (21).

Combining Corollary 2 and Theorem 3 (with $q=-1$), we get the following alternative characterization of box dimension:

PROPOSITION 3 (Another Characterization of box dimension.) *The box dimension of a measure μ is obtained through*

$$\Delta(supp(\mu)) = p - \lim_{h\to 0} \frac{log\lambda(supp(\mu)^h)^{-1}}{logh^{-1}},$$

as soon as the RHS limit exists and $A^h = \{y \in \mathbb{R}^p; inf_{x\in A}\|x-y\| \le h\}$.

This result, originally due to Minkowski, is well known (*cf. e.g.*, Falconer, 1997, p. 20). To illustrate its connection with Corollary 2 and Theorem 3, we give a separate proof of Proposition 3 in Section 7.

4. ESTIMATION OF FUNCTIONALS

Define the empirical distribution based on $\{X_i\}_{i=1}^n$;

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \delta X_i.$$

As estimators of the Rényi and HP functionals we take $\Lambda_{\hat{\mu}}(q;\Gamma_h)$ and

$$\hat{\Lambda}(q;\Psi_h) = \begin{cases} \exp\left(\frac{1}{n}\sum_{i=1}^n \log\left(\frac{1}{n-1}\sum_{j;j\neq i}\Psi_h(X_i,X_j)\right)\right), & q=0, \\ \left(\frac{1}{n}\sum_{i=1}^n \left(\frac{1}{n-1}\sum_{j;j\neq i}\Psi_h(X_i,X_j)\right)^q\right)^{1/q}, & q\neq 0 \end{cases} \tag{22}$$

respectively. The first estimator is simply a 'plug-in' version of the Rényi type functional (replacing μ with $\hat{\mu}$). The second estimator replaces μ with $\hat{\mu}$ in the outer integral appearing in the definition of Λ_μ whereas in the inner integral, μ is replaced with the ith leave-out measure[7]

$$\hat{\mu}^{-i} = \frac{1}{n-1}\sum_{j;j\neq i} \delta X_j, \quad i = 1,\ldots,n.$$

[7]Alternatively, we could use $\Lambda_{\hat{\mu}}(q;\Psi_h)$ for estimating the HP functional, since $\hat{\Lambda}(q;\Psi_h)$ and $\Lambda_{\hat{\mu}}(q;\Psi_h)$ are asymptotically equivalent. However, the version given in (22) more naturally generalizes the sample correlation integral, as will be seen from the examples.

The following result can be stated concerning almost sure (a.s.) convergence of the two estimators of the Rényi and HP functionals:

THEOREM 4 *Suppose $\{X_i\}$ is a stationary and ergodic stochastic process with marginal distribution μ. If $q \leq -1$ we also require μ to have compact support.*

Then,

$$\sup_{0<h<\infty} |\Lambda_{\hat{\mu}}(q;\Gamma_h) - \Lambda_\mu(q;\Lambda_h)| \xrightarrow{a.s.} 0, \tag{23}$$

$$\sup_{0<h<\infty} |\hat{\Lambda}(q;\Psi_h) - \Lambda_\mu(q;\Psi_h)| \xrightarrow{a.s.} 0 \tag{24}$$

as $n \to \infty$.

Example 6 (Sample Correlation Integral.) Serinko (1996) has proved (24) when $q=1$. Notice that

$$\hat{\Lambda}(1;\Psi_h) = \binom{n}{2}^{-1} \sum_{i<j}\sum I(\| X_i - X_j \| \leq h) := \hat{C}(h).$$

This quantity, frequently referred to as the *sample correlation integral*, is a well known estimator of the correlation integral $C(h)$, defined in Section 1. It was introduced by Grassberger and Procaccia (1983). Thus, the family of estimators $\hat{\Lambda}(q;\Psi_h)$, $q \in \mathbb{R}$, incorporates $\hat{C}(h)$ as a special case.

Another generalization of the sample correlation integral is

$$\binom{n}{q+1}^{-1} \sum_{i_1<\dots<i_{q+1}}\sum I(\| X_{i_k} - X_{i_l} \| \leq h; 1 \leq k < l \leq q+1),$$

which is an estimator of $\Lambda_\mu(q;\Psi_h)$, $q = 1,2,3,\dots$. This is a U-statistic of order $q+1$, and was proposed by Grassberger and Procaccia (1983) and further considered by Pesin (1993).

Example 7 (Box Counting.) When $q=-1$, the estimate of the Rényi type functional reduces to

$$\Lambda_{\hat{\mu}}(-1;\Gamma_h) = \sharp\{\gamma \in \Gamma_h; \hat{\mu}(\gamma) > 0\}^{-1}$$

frequently referred to as *box counting*.

Let us now consider estimation of the functional $\Lambda'_{\mu}(q;\Psi_h)$ introduced in Section 3. As an estimator, we take

$$\Lambda'_{\hat{\mu}}(q;\Psi_h) = \begin{cases} \exp(h^{-p}\int \hat{\mu}_h(x)\log\hat{\mu}_h(x)dx), & q=0, \\ h^{-p/q}(\int \hat{\mu}_h(x)^{1+q}dx)^{1/q}, & q\neq 0, \end{cases}$$

with $\hat{\mu}_h(x) = \hat{\mu}(B(x,h))$. Thus $\Lambda'_{\hat{\mu}}(q;\Psi_h)$ is obtained by replacing μ with the empirical measure $\hat{\mu}$ in the definition of Λ'_{μ}.

When investigating the asymptotic properties of $\Lambda'_{\hat{\mu}}(q;\Psi_h)$, we will need an extra regularity condition when $q < -1$. For all $h > 0$ such that $\mu_h(x) > 0$ we assume

$$\begin{aligned} &E(\hat{\mu}_h(x)^{1+q}) \leq C\mu_h(x)^{q+1} \text{ for some constant } C > 1, \\ &\lim_{n\to\infty} E(\hat{\mu}_h(x)^{1+q}) = \mu_h(x)^{q+1}. \end{aligned} \tag{25}$$

Now (25) is a rather mild restriction on $\{X_i\}$, satisfied (for instance) by all i.i.d. processes, since then $n\hat{\mu}_h(x)$ has a binomial distribution $(n\hat{\mu}_h(x) \in \mathrm{Bin}(n, \mu_h(x)))$.

We now prove the analogue of Theorem 4; uniform consistency of $\Lambda'_{\hat{\mu}}(q;\Psi_h)$ w.r.t. h:

THEOREM 5 *Suppose $\{X_i\}$ is a stationary and ergodic stochastic process with marginal distribution μ. If $q < -1$, we also impose* (25) *and assume that μ has compact support. Then*

$$\sup_{0<h<\infty} |\Lambda'_{\hat{\mu}}(q;\Psi_h) - \Lambda'_{\mu}(q;\Psi_h)| \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty, \tag{26}$$

for any $q \in \mathbb{R}$.

Example 8 (Volume of Overlapping Boxes.) Notice that

$$\Lambda'_{\hat{\mu}}(-1;\Psi_h) = \lambda(\cup_{i=1}^n B(X_i,h))^{-1}.$$

The union $\cup_{i=1}^n B(X_i,h)$ of balls of radius h is called a *Minkowski sausage*. Its volume is an estimate of $\lambda(\{x;\mu_h(x) > 0\})$, which is closely related to $\lambda(\mathrm{supp}(\mu)^h)^{-1}$. In view of Proposition 3, box dimension can be expressed in terms of $\lambda(\mathrm{supp}(\mu)^h)$. Box dimension is frequently estimated using box counting, where the number of disjoint boxes over a grid is of interest. The advantage of the Minkowski sausage over box counting is that it requires no choice of grid.

Example 9 (The Henon Attractor.) The Henon mapping is a dynamical system $\{X_i = (X_{i1}, X_{i2})\}$, known to have a fractal invariant measure μ. The mapping is defined through $(X_{i+1,1}, X_{i+1,2}) = (X_{i2} + 1 - 1.4X_{i1}^2, 0.3X_{i1})$. Figure 1 shows a plot of $\cup_{i=1}^n B(X_i, h)$ for different values of h, for the Henon mapping. These plots give successively finer approximations of the attractor supp(μ).

Example 10 (Estimating the Correlation Dimension.) Let $K_{\rm rec}(x) = I(\|x\| \leq 1)$ denote a rectangular kernel function. After some computations, one finds

$$\Lambda'_{\hat{\mu}}(1;\Psi_h) = \frac{1}{n^2}\sum_{i,j=1}^{n}\sum K\left(\frac{X_i - X_j}{h}\right) = \frac{2^p}{n} + \frac{2}{n^2}\sum_{i<j}\sum K\left(\frac{X_i - X_j}{h}\right),$$

with $K(x) = (K_{\rm rec} * K_{\rm rec})(x) = \lambda(B(0,1) \cap B(x,1))$. When $p = 1$, we get the triangular kernel $K(x) = (2 - |x|)_+$. Thus $\Lambda'_{\hat{\mu}}(1;\Psi_h)$ differs from the sample correlation integral $\hat{C}(h)$ in two ways. First, $K_{\rm rec}$ has been replaced by K. Secondly, there is a contribution $2^p/n$ from the diagonal terms, which is asymptotically negligible if $h \gg n^{-1/p}$.

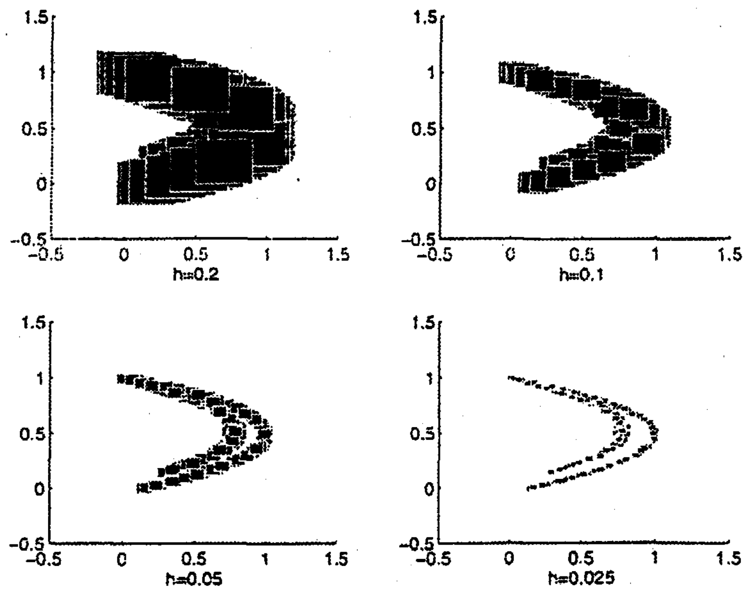

FIGURE 1 The set $\cup_{i=1}^n B(X_i, h)$ for the Henon mapping and different values of h.

5. ESTIMATION OF FRACTAL DIMENSIONS

The existence of the qth generalized HP-dimension $d_\mu(q;\Psi_h)$, can be expressed as

$$\Lambda_\mu(q;\Psi_h) = a(h)h^{d_\mu(q;\Psi_h)}, \quad \frac{\log a(h)}{\log h} \to 0 \text{ as } h \to 0+, \tag{27}$$

or, if we take logarithms,

$$\log \Lambda_\mu(q;\Psi_h) = \log a(h) + d_\mu(q;\Psi_h)\log h.$$

A natural estimator $\hat{d}(q;\Psi)$ of $d_\mu(q;\Psi)$ is therefore the slope of a least squares regression line, computed from 'data' $(\log h_j, \log \Lambda_{\hat{\mu}}(q;\Psi_{h_j}))_{j=1}^{m}$. Here $0 < h_1 < \cdots < h_m$ are some fixed real numbers.

The quality of $\hat{d}(q;\Psi)$ depends not only on the sample size, but also on how much the 'intercept term' $\log a(h)$ varies with h. The oscillating behaviour of $\log a(h)$ has been named *lacunarity* by Mandelbrot (1982). The most regular case is referred to as *exact scaling*, and it occurs when

$$a(h) \equiv a \quad \text{for some } h \leq h_0. \tag{28}$$

As a direct consequence of Theorem 4, we have the following result:

COROLLARY 3 (Consistency of Dimension Estimates.) *Suppose $\{X_i\}$ is a stationary and ergodic stochastic process with marginal distribution μ. If $q \leq -1$, we require μ to have compact support. Suppose further, that $d_\mu(q;\Psi)$ $(d_\mu(q;\Gamma))$ exists, with the exact scaling property* (28) *fulfilled. Consider the least squares estimator $\hat{d}(q;\Psi)$ $(\hat{d}(q;\Gamma))$ based on fixed numbers $0 < h_1 < \cdots < h_m$, that are independent of n and $h_m \leq h_0$. Then $\hat{d}(q;\Psi)$ $(\hat{d}(q;\Gamma))$ is a strongly consistent estimator of $d_\mu(q;\Psi)$ $(d_\mu(q;\Gamma))$.*

Corollary 3 is quite weak, since exact scaling is in general *not* satisfied. In fact, Serinko (1996) has shown that the least squares estimators of the correlation dimension is typically inconsistent when exact scaling fails and $h_1, \ldots, h_m$ are fixed (independent of n). Consistent estimators can be obtained by letting $h_1, \ldots, h_m$ tend to zero with increasing n, *cf.* Cutler (1991) and Serinko (1994). In our setting, this entails a sharpening of Theorem 4, taking *simultaneous* limits $h \to 0$ and $n \to \infty$. Such a result requires further mixing conditions on the process $\{X_i\}$, and is an interesting topic for future research.

Example 11 (The Henon Attractor, revisited.) In Figure 2, the estimated dimension spectrum $\hat{d}(\cdot, \Psi)$ is displayed based on a sequence $\{X_i\}_{i=1}^n$ from the Henon mapping. We used $n = 2^{12}$ for the dashed line and $n = 2^{13}$ for the solid line. By property (4), the Rény dimension is monotonic, and thus any non-monotonic region is essentially bogus. In this example however, the *larger* sample size gives a non-monotonic estimate. The bandwidths were $h = 0.0125, 0.006$. As a comparison, Cutler (1993) reports 1.28 when $q = -1$ and 1.21 when $q = 1$. Because of the fairly moderate sample size, we do not expect the estimates to be accurate for negative q. To be able to characterize the more rarified regions of the measure we need large samples. The difficulty of estimating fractal dimensions when q is negative (especially $q < -1$) has been addressed by Roberts (1996).

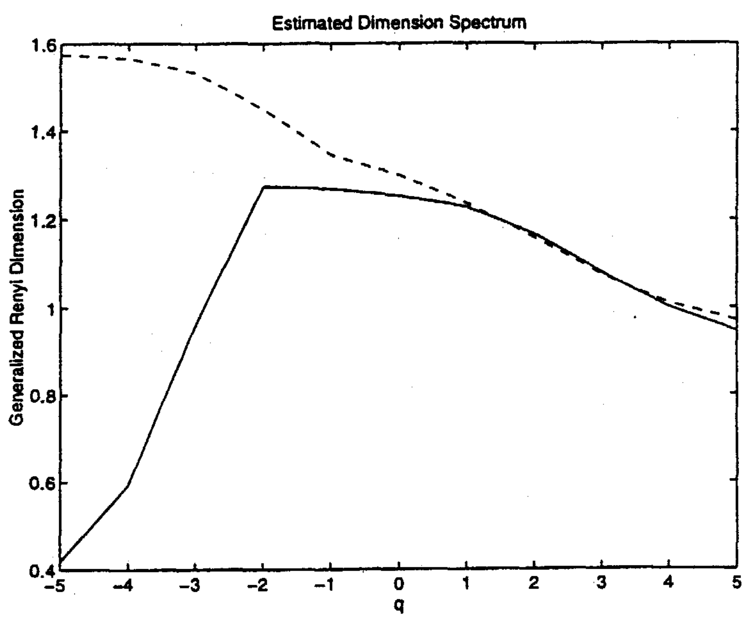

FIGURE 2 The estimated dimension spectrum for the Henon mapping, for two different sample sizes.

Equipped with Theorem 5, it is easy to establish a dimension estimation result for $\tilde{d}'(\cdot,\Psi)$ corresponding to Corollary 4. We refrain from this and give an example instead:

Example 12 (Estimating Dimension Spectrum for a GCD.) The estimated dimension spectrum $\tilde{d}'(\cdot,\Psi)$, for a GCD with parameters $p=1$, $r=2$ and $w=(\rho,1-\rho)=(1/5,4/5)$ is shown in Figure 3. The estimates are marked as circles. The solid line indicates the theoretical dimension spectrum given by $d_\mu(q;\Psi)=-1/q\ \log(\rho^{q+1}+(1-\rho)^{q+1})/\log 2$. In the log–log plot, we used $n=2^{15}$ and $h=0.025$, 0.0125. As in Example 11, we notice how difficult it is to estimate the dimension for small q, since there are few data points in the more rarified regions of the measure. The estimates may be improved by choosing a larger n.[8]

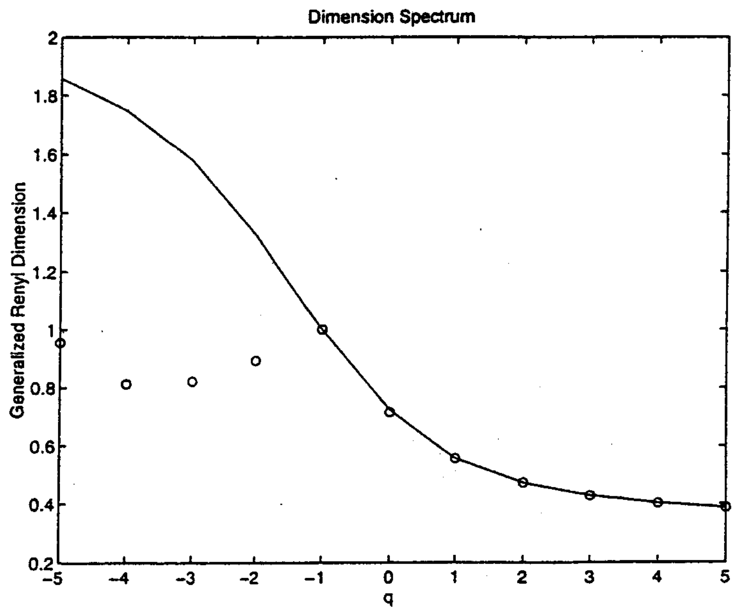

FIGURE 3 The true and estimated dimension spectrum $\tilde{d}'(\cdot,\Psi)$ for a GCD (solid line and circles respectively).

[8]To be precise, we have not established $d'^{\pm}_{\mu}(\cdot,\Psi)$ for the GCD distributions. In view of Theorem 3, we know that $d^{-}_{\mu}(\cdot,\Psi)\geq -1/q\ \log(\rho^{q+1}+(1-\rho)^{q+1})/\log 2$.

6. CONNECTIONS TO DENSITY ESTIMATION

Suppose $H:[0,\infty)\to[0,\infty)$ is a convex function with $H(0)=0$, and put $\tilde{H}(x)=H(x)/x$. Thus $\tilde{H}(x)$ is monotone and continuous, and we also require

$$\lim_{x\to\infty}\tilde{H}(x)=\infty. \tag{29}$$

For any probability density function g on $\mathbb{R}^p$, define the functional

$$\begin{aligned}J(g;\tilde{H}) &= \tilde{H}^{-1}\left(\int H(g(x))dx\right)=\tilde{H}^{-1}\left(\int \tilde{H}(g(x))g(x)dx\right)\\ &= \tilde{H}^{-1}(E(\tilde{H}(g(Y)))),\end{aligned}$$

where Y is a random variable with density $f_Y=g$. Thus $J(g;\tilde{H})$ measures how large $g(Y)$ is in a weighted sense. For instance, if $H(x)=x^2$, we get $J(g;\tilde{H})=E(g(Y))$.

Consider now the kernel density estimator $\hat{f}$ defined in (5), with a bandwidth $h=h_n$ tending to zero with increasing sample size. It was proved in Frigyesi and Hössjer (1998), (*cf.* also Frigyesi, 1994) under some mild regularity conditions on H, $\{X_i\}$, $\{h_n\}$ and K, that

$$\begin{aligned}&J(\hat{f};\tilde{H}) \xrightarrow{\text{a.s.}} \infty, && \text{if } \mu \text{ has a singular part,}\\ &J(\hat{f};\tilde{H}) \xrightarrow{p} J(f;\tilde{H}), && \text{if } d\mu=f\,d\lambda.\end{aligned} \tag{30}$$

This was used to device a test for descriminating probability measures having a singular part from absolutely continuous ones with $J(f;\tilde{H})\le C<\infty$.[9] The first part of (30) is a consequence of (29). When μ has a singular part, $\hat{f}$ will have a spikes on a set with small Lebesgue measure, and these are 'magnified' by $\tilde{H}$, causing $\int H(\hat{f})dx$ to diverge.

Consider in particular the functions

$$H_q(x)=\begin{cases}x^{1+q}, & q\neq 0,\\ x\log x, & q=0.\end{cases} \Leftrightarrow \tilde{H}_q(x)=\begin{cases}x^q, & q\neq 0,\\ \log x, & q=0.\end{cases}$$

[9]This subset of all absolutely continuous probability measures can be made arbitrarily large, by choosing a $\tilde{H}$ that increases slowly enough to infinity.

Then H_q is convex for $q \geq 0$, but we will allow any value of q in the present context. Assume Y is a random variable with density g. Then

$$J(g;\tilde{H}_q) = \begin{cases} (\int g(x)^{1+q}dx)^{1/q}, & q \neq 0, \\ \exp(\int g(x)\log(g(x))dx), & q = 0. \end{cases}$$
$$= \begin{cases} \lambda(\{x; g(x) > 0\})^{-1}, & q = -1, \\ (Eg(Y)^q)^{1/q}, & q \in (-1,0) \cup (0,\infty), \\ \exp(E\log g(Y)), & q = 0. \end{cases} \tag{31}$$

Now $J(\cdot, \tilde{H}_q)$ is intimately related to $\Lambda'_\mu(q;\Psi_h)$. Introduce

$$f_h(x) = E\hat{f}(x) = h^{-p}\int K((y-x)/h)d\mu(y)$$

as a smoothed version of f. If h does not depend on n we get, by the law of large numbers for ergodic processes (*cf. e.g.*, Billingsley, 1965), $\hat{f}(x) \xrightarrow{\text{a.s.}} f_h(x)$, and moreover,

$$\begin{aligned} J(f_h;\tilde{H}_q) &= h^{-p}\Lambda'_\mu(q;K_h), \\ J(\hat{f};\tilde{H}_q) &= h^{-p}\Lambda'_{\hat{\mu}}(q;K_h), \end{aligned} \tag{32}$$

with $K_h(x,y) = K((y-x)/h)$. We will allow the following class of kernels K:

Definition 7 Let $\mathcal{K}$ be the class of kernels K that are non-negative, bounded, have compact support and are bounded away from zero in a neighbourhood of $(0,\ldots,0)$.[10]

Notice that $\mathcal{K}$ includes many of the standard kernels used in univariate density estimation, *e.g.*, the rectangular $(I(|x| \leq 1)/2)$, triangular, Epanechnikov and biweight kernels. For any $K \in \mathcal{K}$, we can find nonnegative reals c_1, c_1, M_1 and M_2 such that

$$M_1 I(\|x\| \leq c_1) \leq K(x) \leq M_2 I(\|x\| \leq c_2).$$

[10] Even though $\int K(x)dx = 1$ is imposed for density estimation, this restriction is not essential when estimating fractal dimensions. Thus we do not impose $\int K(x)dx = 1$ for elements of $\mathcal{K}$.

Using this, it is not difficult to prove that

$$d'^{\pm}_{\mu}(q;K) = d'^{\pm}_{\mu}(q;\Psi),$$
$$d^{\pm}_{\mu}(q;K) = d^{\pm}_{\mu}(q;\Psi),$$

with K representing the sequence of kernels $\{K_h\}$. In particular,

$$J(\hat{f};\tilde{H}_q) \approx h^{d'_\mu(q;\Psi)-p}$$

for large n and small h, when $d'_\mu(q;\Psi)$ exists. Thus, $d'_\mu(q;\Psi)$ determines the rate at which $J(\hat{f};\tilde{H}_q)$ tends to infinity.

Example 13 (GCD, revisited.) We computed a log–log plot of $J(\hat{f}_n;\tilde{H}_1)$ *versus* h^{-1} in Figure 4, using the same GCD distribution as considered in Example 2 and a rectangular kernel. From the central linear region, in which the slope 0.44 approximates $1 - d'_\mu(1;\Psi)$, the dimension was estimated to 0.56.

Example 14 (The Henon Attractor, revisited.) Using Proposition 3, (31) and (32), we plotted $\log J(\hat{f}_n;\tilde{H}_{-1})$ *versus* $-\log h$ in Figure 5, for

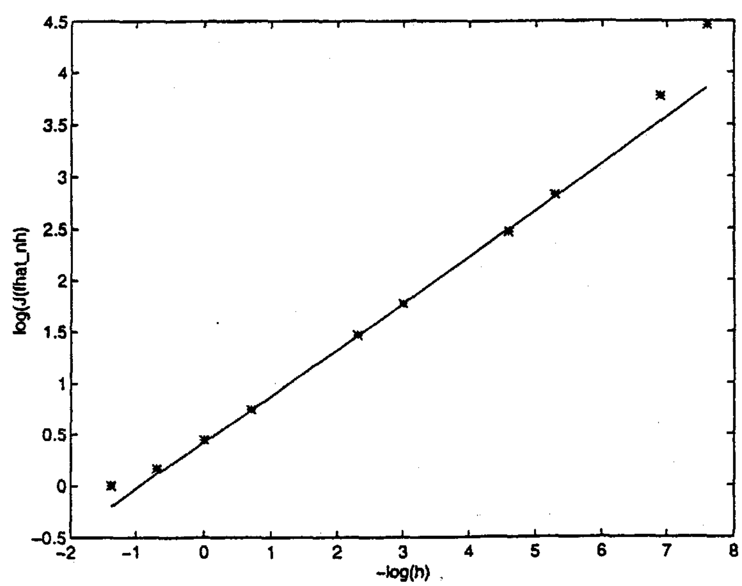

FIGURE 4 The log–log plot of $J(\hat{f};\tilde{H}_1)$ *versus* h^{-1}, for a GCD.

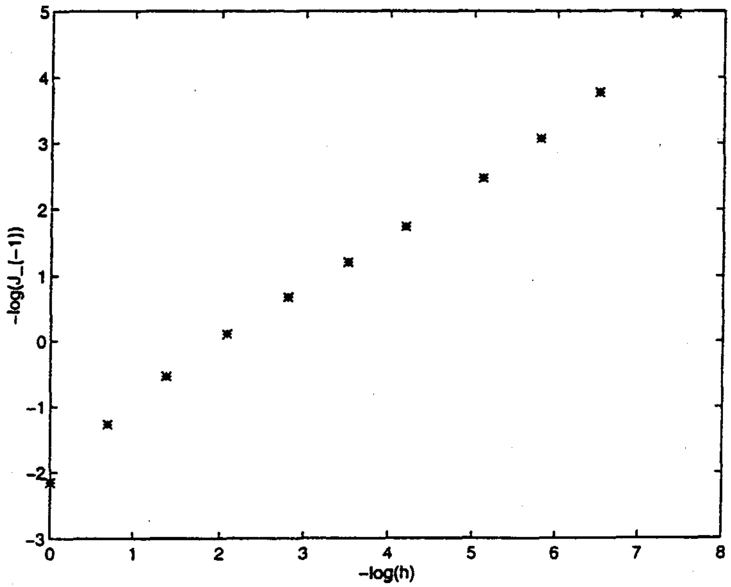

FIGURE 5 A plot of $\log J(\hat{f};\tilde{H}_{-1})$ *versus* $-\log h$, for the Henon mapping. The slope in the central region was 0.77, giving $\Delta(\text{supp}(\mu)) \approx 2-0.77=1.23$.

a rectangular kernel. We obtained $\Delta(\text{supp}(\mu)) \approx 1.23$. In Cutler (1993), this quantity is estimated to 1.28, and in Falconer (1990), the value 1.26 is reported. The number of sample points was $n=8192$, and the sequence of bandwidths $h=1$, 0.5, 0.25, 0.125, 0.06, 0.03, 0.015, 0.006, 0.003, 0.0015, 0.0006, of which 0.125, 0.06, 0.03, 0.015 were used in the regression.

We close this section with a remark concerning the histogram estimator. Suppose $\lambda(x)$ is the unique element of Γ_h such that $x \in \lambda(x)$. Then, the histogram density estimator at x, is defined as

$$\tilde{f}(x) = h^{-p}\int \Gamma_h(x,y)\,d\hat{\mu}(y) = \frac{\sharp\{i; X_i \in \lambda(x)\}}{nh^p},$$

and it is closely related to the Rényi type functional. In fact, one has

$$J(\tilde{f};\tilde{H}_q) = h^{-p}\Lambda'_{\hat{\mu}}(q;\Gamma_h) = h^{-p}\Lambda_{\hat{\mu}}(q;\Gamma_h),$$

which means that $J(\tilde{f};\tilde{H}_q) \approx h^{d_\mu(q;\Gamma)-p}$ for small h and large n, provided $d_\mu(q;\Gamma)$ exists.

7. PROOFS

Proof of Proposition 1 Let $\Omega = \mathrm{supp}(\mu)$ and fix $A > 1$. Pick an arbitrary $h > 0$. Then any $x \notin \Omega$ with $\mathrm{dist}(x, \Omega) := \inf\{\|x - y\|; y \in \Omega\} = \sqrt{A}h$ has $\mu_h(x) = 0$ and $\mu_{Ah}(x) > 0$ and thus (6) fails. ■

Proof of Lemma 1 Let us first verify that any probability measure is weakly diametrically regular in the (q, m)th Rényi sense when $-1 < q < 0$. In fact, we will establish the following, stronger, result: Given any $0 < c < (2m+1)^{-p}$ and $q > -1$, there exists a positive constant $C_1 = C_1(p, q, c, m)$ such that

$$\sum_{\gamma \in \Gamma_{ch}^m} \mu(\gamma)^{1+q} \geq C_1 \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q}, \tag{33}$$

with $C_1(p, q, \cdot, m)$ bounded in a neighbourhood of 0. In order to prove this, we define a map $\pi : \Gamma_h^+ \to \Gamma_{ch}^m$ (*cf.* (16) for the definition of Γ_h^+) with the following properties. Given $\gamma_0 \in \Gamma_h^+$, we form a chain $\gamma_0, \gamma_1, \ldots, \gamma_N$ with $0 \leq N < \infty$. If $\gamma_i \in \Gamma_{ch}^m$, terminate the chain, and put $N = i$. If γ_i is not in Γ_{ch}^m let $\gamma_{i+1} = \mathrm{argmax}_{\gamma' \in U_m(\gamma_i)} \mu(\gamma')$ (In case several of the $\mu(\gamma')$ are equal, choose for uniqueness the γ' whose center coordinates has the greatest first coordinate, then the greatest second coordinate and so forth). Finally put $\pi(\gamma_0) = \gamma_N$, and let $L(\gamma_0) = N$ be the number of edges in the chain. (Tightness of μ assures that $N < \infty$.) Writing

$$\sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q} = \sum_{\gamma \in \Gamma_{ch}^m} \sum_{\gamma' \in \pi^{-1}(\gamma)} \mu(\gamma')^{1+q},$$

it suffices to show that $\mu(\gamma)^{1+q} \geq C_1 \sum_{\gamma' \in \pi^{-1}(\gamma)} \mu(\gamma')^{1+q}$ for all $\gamma \in \Gamma_{ch}^m$. Notice that $\mu(\pi(\gamma)) \geq ((2m+1)^p c)^{-L(\gamma)} \mu(\gamma)$ and $\|\pi(\gamma) - \gamma\| \leq mL(\gamma)$, with $\|\pi(\gamma) - \gamma\|$ as defined in Section 2. Thus

$$\begin{aligned}
\sum_{\gamma' \in \pi^{-1}(\gamma)} \mu(\gamma')^{1+q} &= \sum_{l=0}^{\infty} \sum_{\substack{\gamma' \in \pi^{-1}(\gamma) \\ \|\gamma' - \gamma\| = l}} \mu(\gamma')^{1+q} \\
&\leq \sum_{l=0}^{\infty} \sharp\{\gamma'; \|\gamma' - \gamma\| = l\} \\
&\qquad (((2m+1)^p c)^{l/m} \mu(\gamma))^{1+q} \\
&\leq \left(\sum_{l=0}^{\infty} N_l ((2m+1)^p c)^{l(1+q)/m} \right) \mu(\gamma)^{1+q},
\end{aligned}$$

with $N_0=1$ and $N_l=(2l+1)^p-(2l-1)^p$ for $l\geq 1$. Formula (33) follows if we choose $C_1 = (\sum_{l=0}^{\infty} N_l((2m+1)^p c)^{l(1+q)/m})^{-1}$, and clearly $C_1(p,q,\cdot,m)$ is bounded for small c.

Let us now verify that any measure is weakly diametrically regular also in the (q,A)th sense when $-1<q<0$. Conditioning on which cube X belongs to we have

$$E(\mu_h(X)^q) = \sum_{\gamma\in\Gamma_h} E(\mu_h(X)^q | X\in\gamma)\mu(\gamma). \tag{34}$$

Let $\gamma(x)$ be the unique cube in Γ_h such that $x\in\gamma(x)$. Then $\gamma(x)\subseteq B(x,h)$. Since $q<0$, this implies

$$\sum_{\gamma\in\Gamma_h} \mu(\gamma)^{1+q} \geq E(\mu_h(X)^q) \geq E(\mu_h(X)^q I(X\in\Omega_{ch}^A)). \tag{35}$$

Suppose now that $\gamma\backslash\Omega_{ch}^A$ is nonempty and that $x\in\gamma\backslash\Omega_{ch}^A$. Then $\gamma\subset B(x,h)$. If m is chosen so large that $\lceil A\rceil\leq m$ we also have $B(x,Ah)\subset\cup_{\gamma'\in U_m(\gamma)}\gamma'$. Thus

$$\mu(\gamma)\leq\mu_h(x)\leq c\mu_{Ah}(x)\leq c\sum_{\gamma'\in U_m(\gamma)}\mu(\gamma'),$$

i.e., $\gamma\notin\Gamma_{ch}^m$ whenever $\gamma\backslash\Omega_{ch}^A$ is nonempty. But this in turn implies

$$\begin{aligned} E(\mu_h(X)^q I(X\in\Omega_{ch}^A)) &\geq \sum_{\gamma\in\Gamma_{ch}^m}\int_\gamma \mu_h(x)^q d\mu(x) \\ &\geq \sum_{\gamma\in\Gamma_{ch}^m}\int_\gamma \Big(\sum_{\gamma'\in U_m(\gamma)}\mu(\gamma')\Big)^q d\mu(x) \\ &\geq c^{-q}\sum_{\gamma\in\Gamma_{ch}^m}\mu(\gamma)^{1+q}. \end{aligned} \tag{36}$$

To complete the proof, take logarithms of (35) and (36), divide by $\log h$, let $h\to 0$ and use the fact (just proved) that any measure μ is $\mathrm{DR}_{(q,m)}^{\text{Rényi}}$. ■

Proof of Proposition 2 The first part of (9) follows easily from the fact that $\Omega_{ch}^A=\mathbb{R}^p$ for DR measures and sufficiently small $h>0$, provided $c>0$ is chosen as in Definition 3. The second part follows similarly by noticing that for DR measures we have $\Gamma_{ch}=\Gamma_h^+$ for some $c>0$ and all $h>0$ small enough. (Here Γ_h^+ is defined as in (16).) The third and fourth parts follow immediately from the facts that $\Omega_{ch}^{A_2} \subset \Omega_{ch}^{A_1}$ and $\Gamma_{ch}^{m_2} \subset \Gamma_{ch}^{m_1}$ respectively. Finally, the last part of (9) follows from (35) and (36) (recall that $\lceil A \rceil \leq m$ was assumed in the proof of Lemma 1), since the argument leading to these equations is equally valid for $q \leq -1$ as for $-1 < q < 0$. ■

Before proving Theorem 1, we will establish two lemmas:

LEMMA 2 *Let $x_i \geq 0$ and $r > 0$. Then* $(\sum_{i=1}^N x_i)^r \leq C_2(N, r) \sum_{i=1}^N x_i^r$.

Proof Put $C_2(N, r) = 1$ and $C_2(N, r) = N^{r-1}$ when $0 < r < 1$ and $r \geq 1$ respectively. ■

LEMMA 3 *Assume $0 < c < 3^{-p}$, and $\max_{\gamma \in \Gamma_h} \mu(\gamma) \leq \varepsilon < 1$. Then there exists $0 < C_3(p, \varepsilon, c) < 1$ such that*

$$\sum_{\gamma \in \Gamma_{ch}} \mu(\gamma) log(\mu(\gamma)) \leq C_3(p, \varepsilon, c) \sum_{\gamma \in \Gamma_h} \mu(\gamma) log(\mu(\gamma)),$$

and $\lim_{c \to 0} C_3(p, \varepsilon, c) = 1$.

Proof The idea is in principle the same as in Lemma 1 (with $m = 1$). Let $f(x) = x \log x$. It suffices to show that for each $\gamma \in \Gamma_{ch}$ we have

$$f(\mu(\gamma)) \leq C_3(p, \varepsilon, c) \sum_{\gamma' \in \pi^{-1}(\gamma)} f(\mu(\gamma')),$$

with C_3 chosen appropriately. Assume $0 < x \leq \varepsilon$ and $0 < \delta \leq 1$. We then get

$$f(\delta x) = \delta\left(1 + \frac{\log \delta}{\log x}\right) f(x) \geq \delta\left(1 + \frac{\log \delta}{\log x}\right) f(x) \geq C(\varepsilon) \delta^{1/2} f(x),$$

with $C(\varepsilon) = \sup_{0 < \delta < 1} \delta^{1/2}(1 + \log \delta / \log \varepsilon)$. From this it follows

$$\begin{aligned} \sum_{\gamma' \in \pi^{-1}(\gamma)} f(\mu(\gamma')) &= f(\mu(\gamma)) + \sum_{l=1}^{\infty} \sum_{\substack{\gamma' \in \pi^{-1}(\gamma) \\ \|\gamma' - \gamma\| = l}} f(\mu(\gamma')) \\ &\geq f(\mu(\gamma)) + C(\varepsilon) f(\mu(\gamma)) \sum_{l=1}^{\infty} N_l (3^p c)^{l/2}, \end{aligned}$$

with N_l as in the proof of Lemma 1 and $\mu(\gamma')/\mu(\gamma) \leq (3^p c)^l$. We may thus put $C_3(p, \varepsilon, c) = (1 + C(\varepsilon) \sum_{l=1}^{\infty} N_l (3^p c)^{l/2})^{-1}$, and it follows that $\lim_{c \to 0} C_3(p, \varepsilon, c) = 1$. ■

Proof of Theorem 1 Consider first $q > 0$. Let $\gamma(x)$ be the unique cube in Γ_h such that $x \in \gamma(x)$. Then $\gamma(x) \subseteq B(x, h) \subseteq \cup_{\gamma' \in U(\gamma(x))} \gamma'$. As a consequence, for any $\gamma \in \Gamma_h$,

$$\mu(\gamma)^q \leq \mu_h(X)^q | X \in \gamma \leq \Big(\sum_{\gamma' \in U(\gamma)} \mu(\gamma') \Big)^q . \tag{37}$$

Use (34) and (37) to obtain

$$\begin{aligned} \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q} \leq E\mu_h(X)^q &\leq \sum_{\gamma \in \Gamma_h} \mu(\gamma) \Big(\sum_{\gamma' \in U(\gamma)} \mu(\gamma') \Big)^q \\ &\leq \sum_{\gamma \in \Gamma_h} \Big(\sum_{\gamma' \in U(\gamma)} \mu(\gamma') \Big)^{1+q} \\ &\leq C_2(3^p, 1+q) \sum_{\gamma \in \Gamma_h} \sum_{\gamma' \in U(\gamma)} \mu(\gamma')^{1+q} \\ &\leq 3^p C_2(3^p, 1+q) \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q}, \end{aligned}$$

where the last inequality follows from Lemma 2 and the fact that $\sharp U(\gamma) = 3^p$. Letting $h \to 0$ we obtain (10).

When $q < 0$ we fix $0 < c = c(h) < 3^{-p}$ as in the proof of Lemma 1 (with $m = 1$). Then

$$\begin{aligned} \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{1+q} \geq E\mu_h(X)^q &\geq \sum_{\gamma \in \Gamma_h} \mu(\gamma) \Big(\sum_{\gamma' \in U(\gamma)} \mu(\gamma') \Big)^q \\ &\geq \sum_{\gamma \in \Gamma_{ch}} \mu(\gamma) \Big(\sum_{\gamma' \in U(\gamma)} \mu(\gamma') \Big)^q \\ &\geq c^{-q} \sum_{\gamma \in \Gamma_{ch}} \mu(\gamma)^{1+q}. \end{aligned}$$

The first inequality proves (11). What remains to show for $q < 0$ is, in view of Lemma 1, the identity $d_\mu^\pm(q; \Psi) = d_\mu^\pm(q; \Gamma)$ for any

$\mathrm{DR}^{\text{Rényi}}_{(q,1)}$ measure. But the last displayed inequalities imply (if $h<1$)

$$\begin{aligned}\frac{\log\left(E\mu_h(X)^q\right)^{1/q}}{\log h} \geq & \frac{\log c^{-1}}{\log h} \\ & + \frac{\log\left(\sum_{\gamma\in\Gamma_{ch}}\mu(\gamma)^{q+1}/\sum_{\gamma\in\Gamma_h}\mu(\gamma)^{q+1}\right)^{1/q}}{\log h} \\ & + \frac{\log\left(\sum_{\gamma\in\Gamma_h}\mu(\gamma)^{q+1}\right)^{1/q}}{\log h},\end{aligned}$$

and this yields (10) when $h\to 0$ since by assumption $\log c/\log h\to 0$.

It remains to consider the case $q=0$. Write

$$E\log\mu_h(X) = \sum_{\gamma\in\Gamma_h}\mu(\gamma)E(\log\mu_h(X)|X\in\gamma)$$

and assume without loss of generality that μ is not a one point distribution. Then there exists $0<\varepsilon<1$ such that $\max\mu(\gamma)\leq\varepsilon$ for all h small enough and $\gamma\in\Gamma_h$. This argument and Lemma 3 yield

$$\begin{aligned}\sum_{\gamma\in\Gamma_h}\mu(\gamma)\log(\mu(\gamma)) \leq E\log\mu_h(X) & \leq \sum_{\gamma\in\Gamma_h}\mu(\gamma)\log\Bigg(\sum_{\gamma'\in U(\gamma)}\mu(\gamma')\Bigg) \\ & \leq \sum_{\gamma\in\Gamma_{ch}}\mu(\gamma)\log\Bigg(\sum_{\gamma'\in U(\gamma)}\mu(\gamma')\Bigg) \\ & \leq \sum_{\gamma\in\Gamma_{ch}}\mu(\gamma)\log(c^{-1}\mu(\gamma)) \\ & = \log c^{-1}\sum_{\gamma\in\Gamma_{ch}}\mu(\gamma) + \sum_{\gamma\in\Gamma_{ch}}\mu(\gamma)\log(\mu(\gamma)) \\ & \leq \log c^{-1} + C_3(p,\varepsilon,c)\sum_{\gamma\in\Gamma_h}\mu(\gamma)\log(\mu(\gamma))\end{aligned}$$

as $\lim_{c\to 0}C_3(p,\varepsilon,c)=1$, (10) will follow by letting $h\to 0$ and $c\to 0$ so that $\log c/\log h\to 0$. ■

Proof of Corollary 1 Suppose we can prove

$$d_\mu(q;\Psi) = \lim_{m\to\infty}\frac{\Lambda_\mu(q;\Psi_{r^{-m}})}{\log r^{-m}}, \tag{38}$$

and similarly for $d^{\pm}_{\mu}(q;\Psi)$, with lim replaced by lim inf or lim sup. Then the rest of the corollary follows as in the proof of Theorem 1, taking $h=r^{-m}$. To prove (38), assume $r^{-m-1}<h\leq r^{-m}$. If $q>0$, we get

$$\frac{m}{m+1}\frac{\log\left(E\mu_{r^{-m}}(X)^q\right)^{1/q}}{\log r^{-m}}\leq\frac{\log\left(E\mu_h(X)^q\right)^{1/q}}{\log h}$$
$$\leq\frac{m+1}{m}\frac{\log\left(E\mu_{r^{-m-1}}(X)^q\right)^{1/q}}{\log r^{-m-1}}.$$

Let $h\to 0$ and $m\to\infty$ simultaneously, so that $r^{-m-1}<h\leq r^{-m}$ is fulfilled. Then (38) follows. The case $q\leq 0$ is similar. ∎

Proof of (13) from Example 2 For ease of notation, let $h_1=r^{-m}$ and $h_2=r^{-m+m_0}$, with m_0 a fixed positive integer, not depending on m. It suffices to establish (13) for some constant $c=c(h)>0$. This we will do by showing that for small enough c,

$$K\leq\log\frac{\sum_{\gamma\in\Gamma_{ch_1}}\mu(\gamma)^{1+q}}{\sum_{\gamma\in\Gamma_{h_1}}\mu(\gamma)^{1+q}}\leq 0,$$

with K independent of m. Write $\sum_{\gamma\in\Gamma_{h_1}}\mu(\gamma)^{1+q}=\sum_{\tilde\gamma\in\Gamma_{h_2}}\sum_{\gamma\subset\tilde\gamma}\mu(\gamma)^{1+q}$, and note that for any $\tilde\gamma\in\Gamma_{h_2}$ with $\mu(\tilde\gamma)>0$,

$$\max_{\substack{\gamma,\gamma'\subset\tilde\gamma\\ \mu(\gamma),\mu(\tilde\gamma)>0}}\frac{\mu(\gamma)}{\mu(\gamma')}=\left(\frac{\bar\nu}{\underline{\nu}}\right)^{m_0},$$

where $\bar\nu=\{\max\nu_i;\nu_i>0\}$ and $\underline{\nu}=\{\min\nu_i;\nu_i>0\}$. Now write $\Gamma_{h_1}=\Gamma'_{h_1}\cup\Gamma''_{h_1}$, with $\Gamma'_{h_1}=\{\gamma\in\Gamma_{h_1};\mu(\gamma)>0,\gamma\subset\tilde\gamma\in\Gamma_{h_2},U(\gamma)\subset\tilde\gamma\}$, and $\Gamma''_{h_1}=\Gamma_{h_1}\backslash\Gamma'_{h_1}$. For any $0<c<(1+(\bar\nu/\underline{\nu})^{m_0}(3^p-1))^{-1}$, $\Gamma_{ch_1}\supseteq\Gamma'_{h_1}$, whence

$$\sum_{\gamma\in\Gamma_{ch_1}}\mu(\gamma)^{1+q}\geq\sum_{\gamma\in\Gamma'_{h_1}}\mu(\gamma)^{1+q}=\sum_{\tilde\gamma\in\Gamma_{h_2}}\sum_{\substack{\gamma\in\Gamma'_{h_1}\\ \gamma\subset\tilde\gamma}}\mu(\gamma)^{1+q}.$$

Thus

$$\log\frac{\sum_{\tilde\gamma\in\Gamma_{h_2}}\sum_{\gamma\in\Gamma'_{h_1},\,\gamma\subset\tilde\gamma}\mu(\gamma)^{1+q}}{\sum_{\tilde\gamma\in\Gamma_{h_2}}\sum_{\gamma\subset\tilde\gamma}\mu(\gamma)^{1+q}}\leq\log\frac{\sum_{\gamma\in\Gamma_{ch_1}}\mu(\gamma)^{1+q}}{\sum_{\gamma\in\Gamma_{h_1}}\mu(\gamma)^{1+q}}\leq 0.$$

This sum is independent of m as $\sum_{\gamma\in\Gamma'_{h};\gamma\subset\tilde{\gamma}}\mu(\gamma)^{1+q}/\sum_{\gamma\subset\tilde{\gamma}}\mu(\gamma)^{1+q}$ does not depend on m and $\tilde{\gamma}$ as soon as $\mu(\tilde{\gamma})>0$. It only depends on m_0 that is kept constant, and large enough for the quotient to be positive. ∎

The following proof is due to Colleen Cutler (personal communication):

Proof of Corollary 2 Let N_h be the number of cubes $\gamma\in\Gamma_h$ that intersect supp(μ). Write $N_h=N_h^+ + N_h^0$, where N_h^0 consists of sets with measure zero and $N_h^+ = \#\Gamma_h^+$. Every set in N_h^0 must have at least one neighbour in N_h^+. But any γ_h in $\mathbb{R}^p$ has at most 3^p-1 neighbours. It follows that $N_h^0\leq(3^p-1)N_h^+$ so that N_h^+ and N_h grow at the same rate. This proves (19), and the second part then follows from (17). ∎

Proof of Theorem 3 We first prove (20) when $q\in[-1,0)\cup(0,\infty)$. Let $\gamma(x)$ be the unique element of Γ_h such that $x\in\gamma(x)$. Since $\gamma(x)\subset B(x,h)\subset\cup_{\gamma'\in U(\gamma(x))}\gamma'$, it follows that $\mu(\gamma(x))^{1+q}\leq\mu_h(x)^{1+q}\leq(\sum_{\gamma';\gamma'\in U(\gamma(x))})^{1+q}$. Integrating w.r.t. x, we get

$$\sum_{\gamma\in\Gamma_h}\mu(\gamma)^{1+q}\leq h^{-p}\int\mu_h(x)^{1+q}dx\leq\sum_{\gamma\in\Gamma_h}\left(\sum_{\gamma';\gamma'\in U(\gamma)}\mu(\gamma)\right)^{1+q}.$$

As in the proof of Theorem 1, we may use Lemma 1 to deduce that the RHS and LHS are of the same order, and this proves (20). Assume next $q=0$. Define the measure $\bar{\mu}_h$ through $\bar{\mu}_h(A)=(2h)^{-p}\int_A\mu_h(x)dx$. Then

$$\begin{aligned}\bar{\mu}_h(\gamma)&=(2h)^{-p}\int_\gamma\mu_h(x)dx=\int_\gamma\int(2h)^{-p}I(\|y-x\|\leq h)d\mu(y)dx\\&=\sum_{\gamma'\in U(\gamma)}\int_{y\in\gamma'}\int_{x\in\gamma}(2h)^{-p}I(\|y-x\|\leq h)dxd\mu(y)\\&:=\sum_{\gamma'\in U(\gamma)}w(\gamma,\gamma')\mu(\gamma').\end{aligned}\tag{39}$$

From the definition of $w(\cdot,\cdot)$, one has (since $\int I(\|x\|\leq1)dx=2^p$)

$$w(\gamma,\gamma')\geq0,\quad\sum_{\gamma\in U(\gamma')}w(\gamma,\gamma')=1.\tag{40}$$

Now $\mu_h(x) \le \mu(U(\gamma(x)))$. Thus

$$h^{-p} \int \mu_h(x) \log(\mu_h(x))dx \le 2^p \sum_{\gamma \in \Gamma_h} \bar{\mu}_h(\gamma) \log(U(\gamma)).$$

Use (39)–(40) to obtain

$$\begin{aligned}\sum_{\gamma \in \Gamma_h} \bar{\mu}_h(\gamma) \log(U(\gamma)) &= \sum_{\gamma \in \Gamma_h} \Big(\sum_{\gamma' \in U(\gamma)} w(\gamma, \gamma')\mu(\gamma') \Big) \log(U(\gamma)) \\ &= \sum_{\gamma' \in \Gamma_h} \mu(\gamma') \sum_{\gamma \in U(\gamma')} w(\gamma, \gamma') \log(U(\gamma)) \\ &\le \sum_{\gamma' \in \Gamma_h} \mu(\gamma') \log(U_2(\gamma')),\end{aligned}$$

where $U_2(\gamma') = \{\gamma \in \Gamma_h; \|\gamma' - \gamma\| \le 2\}$, and $\|\cdot\|$ is defined as in the proof of Lemma 1. Now (making use of Lemma 3, as in the proof of Theorem 1) $\sum_{\gamma' \in \Gamma_h} \mu(\gamma') \log(U_2(\gamma'))$ is of the same order as $\sum_{\gamma \in \Gamma_h} \mu(\gamma') \log(\gamma')$. Thus, the last two displays imply, letting $h \to 0$, $d_\mu'^{\pm}(0, \Psi) \ge d_\mu^{\pm}(0, \Gamma)$. For the reversed inequality, notice that

$$\begin{aligned}&h^{-p} \int \mu_h(x) \log(\mu_h(x))dx \\ &\quad \ge 2^p \sum_{\gamma \in \Gamma_h} \bar{\mu}_h(\gamma) \log(\bar{\mu}_h(\gamma)) \\ &\quad = 2^p \sum_{\gamma' \in \Gamma_h} \mu(\gamma') \sum_{\gamma \in U(\gamma')} w(\gamma, \gamma') \log \bar{\mu}_h(\gamma) \\ &\quad \ge 2^p \sum_{\gamma' \in \Gamma_h} \mu(\gamma') \sum_{\gamma \in U(\gamma')} w(\gamma, \gamma') \log(w(\gamma, \gamma')\mu(\gamma')) \\ &\quad \ge 2^p \log 3^{-p} + 2^p \sum_{\gamma' \in \Gamma_h} \mu(\gamma') \log(\mu(\gamma')),\end{aligned}$$

where we used Jensen's inequality and the convexity of $x \log x$ in the first and last step. This proves $d_\mu'^{\pm}(0, \Psi) \le d_\mu^{\pm}(0, \Gamma)$.

Next, we turn to (21). Since $q < -1$, we have

$$\begin{aligned}h^{-p} \int \mu_h(x)^{1+q} dx &= h^{-p} \int \int I(\|y - x\| \le h) d\mu(y) \mu_h(x)^q dx \\ &= \int \int h^{-p} I(\|y - x\| \le h) \mu_h(x)^q dx d\mu(y) \\ &\ge 2^p \int \mu_{2h}(y)^q d\mu(y).\end{aligned}$$

In the last step, we used the fact that if $d\mu(y) > 0$ and $x \in B(y, h)$, then $0 < \mu_h(x) \le \mu_{2h}(y)$. Letting $h \to 0$, we obtain (21).

To find a measure μ with $d'_\mu(q; \Psi) > d_\mu(q; \Psi) = d_\mu(q; \Gamma)$ when $q < -1$, we consider the absolutely continuous measures from Example 3. Suppose $\beta > 0$ is chosen so that $\beta(q+1) < -1$. We know already that $d_\mu(q; \Psi) = (q+1)(\beta+1)/q$, and it is not difficult to show that $d_\mu(q; \Gamma)$ exists and has the same value. On the other hand, $\int \mu_h(x)^{1+q} dx \ge \int_{-h}^{h} (x+h)^{(\beta+1)(q+1)} dx = \infty$, *i.e.*, $d'_\mu(q; \Psi) = \infty$.

Finally, we construct a measure μ on $[0, 1]$ with $d'^{+}_\mu(q; \Psi) < d^{+}_\mu(q; \Gamma)$ when $q < -1$. Define

$$\mu(-\infty, x] = \frac{1}{3}\theta(x) + \sum_{k=1}^{\infty} p_k \theta(x - 2^{-k}) + \sum_{k=1}^{\infty} s_k G\left(\frac{x - 2^{-k}}{2^{-k-2}}\right),$$

where $p_k, s_k \ge 0$, $\sum p_k = \sum s_k = 1/3$, $\theta(x) = I(x \ge 0)$ and

$$G(x) = \begin{cases} 0, & x < 0, \\ x^{1+\beta}/2, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

Let further $-1 < \beta < 0$ and $-1 < (\beta+1)(q+1) < 0$, and put $\tilde{\mu}(-\infty, x] = p\theta(x) + sG(x/\delta)$. We then get $\int_{-\infty}^{\infty} \tilde{\mu}_h(x)^{1+q} dx = O(s^{1+q}\delta\,(h/\delta)^{1+q})$, with $\tilde{\mu}_h(x) = \tilde{\mu}(B(x, h))$, assuming that $h = O(\delta)$. Notice that we can make this last estimate independent of p, by our choice of G. Let $2^{-m} \le h < 2^{-m+1}$ and put $\delta_k = 2^{-k-2}$. We then get

$$\begin{aligned} \int_{-\infty}^{\infty} \mu_h(x)^{1+q} dx &= O\left(\sum_{k=1}^{m} s_k^{1+q} \delta_k \left(\frac{h}{\delta_k}\right)^{1+q}\right) \\ &= O\left(h^{1+q} \sum_{k=1}^{\infty} s_k^{1+q} 2^{kq}\right) = O(h^{1+q}), \end{aligned}$$

where the last step follows if $s_k \searrow 0$ slowly enough. It follows that $d'^{+}_\mu(q; \Psi) \le 1$. Now, if $2^{-k} \in \gamma$ and $\gamma \in \Gamma_{2^{-m}}$, it follows that $\mu(\gamma) = p_k$ for $k \le m-2$. Thus $\sum_{\gamma \in \Gamma_{2^{-m}}} \mu(\gamma)^{1+q} \ge \sum_{k=1}^{m-2} p_k^{1+q} \ge p_{m-2}^{1+q}$. If now $p_k \searrow 0$ fast enough, so that $\log p_k / k \to -\infty$ we have $d^{+}_\mu(q; \Gamma) = \infty$. ∎

Proof of Theorem 4 Notice first that

$$\Lambda_{\hat{\mu}}(q; \Gamma_h) = \begin{cases} \left(\sum_{\gamma \in \Gamma_h} \hat{\mu}(\gamma)^{1+q}\right)^{1/q}, & q \ne 0, \\ \exp\left(\sum_{\gamma \in \Gamma_h} \hat{\mu}(\gamma) \log(\hat{\mu}(\gamma))\right), & q = 0. \end{cases} \tag{41}$$

Clearly, $\hat{\mu}(\gamma) \xrightarrow{\text{a.s.}} \mu(\gamma)$, $\forall \gamma \in \Gamma_h$, by the strong law of large numbers for ergodic sequences. We will need a stronger version, uniform in h and γ. Put $\hat{\mu}_h(x) = \hat{\mu}(B(x,h))$, and $(-\infty, x] = (-\infty, x_1] \times \cdots \times (-\infty, x_p]$, if $x = (x_1, \ldots, x_p)$. Then

$$\begin{aligned}\sup_{h>0;\gamma\in\Gamma_h} |\hat{\mu}(\lambda) - \mu(\lambda)| &\leq \sup_x |\hat{\mu}_{h/2}(x) - \mu_{h/2}(x)| \\ &\leq 2^p \sup_x |\hat{\mu}(-\infty, x] - \mu(-\infty, x]| := \delta \xrightarrow{\text{a.s.}} 0.\end{aligned} \tag{42}$$

We used the fact that $\hat{\mu}_{h/2}(x) - \mu_{h/2}(x)$ is a linear combination of $\hat{\mu}(-\infty, y] - \mu(-\infty, y]$, with y ranging over the 2^p corners of $B(x, h/2)$. The last step in (42) is a Glivenko–Cantelli type result for ergodic empirical processes.

Let us start introducing some notation needed for the proof of (23). Order the atoms of μ (if there are any) according to their probabilities, *i.e.*, $\mu(\{x_1\}) \geq \mu(\{x_2\}) \geq \cdots$. Given $\varepsilon > 0$, we choose $N = N(\varepsilon)$ as the smallest number such that $\sum_{i=N+1}^{\infty} \mu(\{x_i\})^{1\wedge(1+q)} \leq \varepsilon$. (If μ is continuous, then $N = 0$.) Subdivide the cubes of Γ_h into two groups according to $\Gamma'_h = \cup_{i=1}^{N} \gamma(x_i)$, $\Gamma''_h = \Gamma_h \backslash \Gamma'_h$, with $\gamma(x_i)$ the element of Γ_h containing x_i. Further, let $M = M(\varepsilon, h)$ be the smallest number satisfying $\sum_{\gamma\in\Gamma_h;\gamma\subset([-M,M]^p)^c} \mu(\gamma)^{1\wedge(1+q)} \leq \varepsilon$, and put

$$\Gamma'''_h = \Gamma'''_h(M) = \{\gamma \in \Gamma_h; \gamma \subset ([-M, M]^p)^c\}.$$

If $q > 0$, we split Γ_h into two subsets, according to whether $\gamma \in \Gamma'''_h$ or not;

$$\begin{aligned}&\left| \sum_{\gamma\in\Gamma_h} \hat{\mu}(\gamma)^{q+1} - \sum_{\gamma\in\Gamma_h} \mu(\gamma)^{q+1} \right| \\ &\quad \leq \left(2\left\lceil \frac{M}{h} \right\rceil\right)^p \max_{\gamma\notin\Gamma'''_h} |\hat{\mu}(\gamma)^{1+q} - \mu(\gamma)^{1+q}| \\ &\qquad + \sum_{\gamma\in\Gamma'''_h} (\hat{\mu}(\gamma)^{q+1} + \mu(\gamma)^{q+1}) \\ &\quad \leq \left(2\left\lceil \frac{M}{h} \right\rceil\right)^p (1 - (1-\delta)^{q+1}) + \hat{\mu}\left(\bigcup_{\gamma\in\Gamma'''_h} \gamma\right) \max_{\gamma\in\Gamma'''_h} \hat{\mu}(\gamma)^q \\ &\qquad + \mu\left(\bigcup_{\gamma\in\Gamma'''_h} \gamma\right) \max_{\gamma\in\Gamma'''_h} \mu(\gamma)^q \\ &\quad \leq \left(2\left\lceil \frac{M}{h} \right\rceil\right)^p (1 - (1-\delta)^{q+1}) + (\varepsilon + \delta)^{1+q} + \varepsilon^{1+q},\end{aligned} \tag{43}$$

using that fact that $\sum_{\gamma \in \Gamma_h'''} \mu(\gamma) \leq \varepsilon$. By definition of M, $\mu(((-M, M)^p)^c) \geq \varepsilon$, so $M(\varepsilon, \cdot)$ is a bounded function, and the RHS of (43) is small *simultaneously* for all $h \geq h_0 > 0$ (where h_0 will be chosen below), if ε and δ are small.

Suppose now h_0 is chosen as the largest number satisfying $\sup_{\gamma \in \Gamma_{h_0}''} \mu(\gamma) \leq 2\varepsilon$. Then, if $0 < h \leq h_0$, we get

$$\begin{aligned}
&\left| \sum_{\gamma \in \Gamma_h} \hat{\mu}(\gamma)^{q+1} - \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{q+1} \right| \\
&\qquad \leq \sum_{\gamma \in \Gamma_h'} |\hat{\mu}(\gamma)^{q+1} - \mu(\gamma)^{q+1}| + \sup_{\gamma \in \Gamma_h''} \hat{\mu}(\gamma)^q + \sup_{\gamma \in \Gamma_h''} \mu(\gamma)^q \\
&\qquad \leq N(1 - (1-\delta)^{q+1}) + (2\varepsilon + \delta)^q + (2\varepsilon)^q. \qquad (44)
\end{aligned}$$

Notice that (43) and (44) give upper bounds for $|\Lambda_{\hat{\mu}}(q; \Gamma_h)^q - \Lambda_\mu(q; \Gamma_h)^q|$. Since $\Lambda_\mu(q; \Gamma_h)^q \leq 1$ for all $h > 0$, we obtain (23) when $q > 0$ by first choosing ε small end then letting $n \to \infty$ ($\delta \xrightarrow{\text{a.s.}} 0$).

If $-1 < q < 0$ we have $\Lambda_\mu(q; \Gamma_h)^q \overset{\cup}{\cap} \Lambda_\mu(q; \Psi_h)^q$ and $\Lambda_\mu(q; \Gamma_h)^q \overset{\cup}{\cap} \Lambda(q; \Psi_h)^q$. This is proved as in Theorem 1, using Lemma 1. Further $\Lambda_\mu(q; \Psi_h)$ is a nondecreasing function of h. Thus, we get two cases,

(1) $\Lambda_\mu(q; \Gamma_h) = \Lambda_\mu(q; \Psi_h) = 0$, $\forall h \in (0, \infty)$,
(2) $\Lambda_\mu(q; \Gamma_h)$, $\Lambda_\mu(q; \Psi_h) > 0$, $\forall h \in (0, \infty)$.

We will show below that $\Lambda_{\hat{\mu}}(q; \Gamma_h)^q \overset{\cup}{\cap} \hat{\Lambda}(q; \Psi_h)^q$, so (23) will follow for Case 1 if we establish (24) for Case 1, and this will be done below. Turning to Case 2, we first assume that μ is discrete. Then, splitting Γ_h into Γ_h' and Γ_h'', as in (44), we get

$$\begin{aligned}
\left| \sum_{\gamma \in \Gamma_h} \hat{\mu}(\gamma)^{q+1} - \sum_{\gamma \in \Gamma_h} \mu(\gamma)^{q+1} \right| &\leq N\delta^{1+q} \\
&\quad + \sum_{i=N+1}^{\infty} (\hat{\mu}\{x_i\}^{1+q} + \mu\{x_i\}^{1+q}) \\
&\leq N\delta^{1+q} + \varepsilon + \sum_{i=N+1}^{\infty} \hat{\mu}\{x_i\}^{1+q}. \qquad (45)
\end{aligned}$$

Since $\hat{\mu}\{x_i\} \xrightarrow{\text{a.s.}} \mu\{x_i\}$, $\forall i$ and $E(\hat{\mu}\{x_i\}) = \mu\{x_i\}$, it follows from Fatou's Lemma and Jensen's inequality that

$$\begin{cases} \liminf_{n\to\infty} \sum_{i=N+1}^{\infty} \hat{\mu}\{x_i\}^{1+q} \geq \sum_{i=N+1}^{\infty} \mu\{x_i\}^{1+q} \text{ a.s.} \\ E(\sum_{i=N+1}^{\infty} \hat{\mu}\{x_i\}^{1+q}) \leq \sum_{i=N+1}^{\infty} \mu\{x_i\}^{1+q}, \end{cases}$$

so

$$\lim_{n\to\infty} \sum_{i=N+1}^{\infty} \hat{\mu}\{x_i\}^{1+q} = \sum_{i=N+1}^{\infty} \mu\{x_i\}^{1+q} \text{ a.s.} \tag{46}$$

Combining (45) and (46), we obtain (23) for Case 2 when μ is discrete. Next, we consider Case 2 when μ is not discrete, *i.e.*, the decomposition $\mu = \mu_d + \mu_c$ into a discrete and continuous part is such that $\mu_c(\mathbb{R}^p) > 0$. If $h \geq h_0$ (with h_0 chosen below), we get

$$\begin{aligned} \sum_{\gamma \in \Gamma_h'''(M(\varepsilon,h_0))} \mu(\gamma)^{1+q} &= \sum_{\gamma \in \Gamma_h'''(M(\varepsilon,h_0))} \left(\sum_{\substack{\gamma' \in \Gamma_{h_0}'''(M(\varepsilon,h_0)) \\ \gamma'\cap\gamma\neq\emptyset}} \mu(\gamma\cap\gamma') \right)^{1+q} \\ &\leq 2^p \sum_{\gamma' \in \Gamma_{h_0}'''(M(\varepsilon,h_0))} \mu(\gamma')^{1+q} \leq 2^p\varepsilon, \end{aligned} \tag{47}$$

using the concavity of $x \to x^{1+q}$ and the fact that there are at most 2^p cubes γ that intersect a fixed γ'. Similarly

$$\begin{aligned} \sum_{\gamma \in \Gamma_h'''(M(\varepsilon,h_0))} \hat{\mu}(\gamma)^{1+q} &\leq 2^p \sum_{\gamma' \in \Gamma_{h_0}'''(M(\varepsilon,h_0))} \hat{\mu}(\gamma')^{1+p} \\ &\xrightarrow{\text{a.s.}} 2^p \sum_{\gamma' \in \Gamma_{h_0}'''(M(\varepsilon,h_0))} \mu(\gamma')^{1+p} \leq 2^p\varepsilon, \end{aligned} \tag{48}$$

uniformly for $h \geq h_0$. The a.s. convergence in (48) is derived in the same way as in (46). Thus, from (47) and (48) we get

$$\begin{aligned} &\sup_{h\geq h_0} \limsup_{n\to\infty} \left| \sum_{\gamma\in\Gamma_h} (\hat{\mu}(\gamma)^{1+q} - \mu(\gamma)^{1+q}) \right| \\ &\quad \leq \sup_{h\geq h_0} \limsup_{n\to\infty} \left(2 \left\lceil \frac{M(\varepsilon,h_0)}{h} \right\rceil \right)^p \delta^{1+q} + 2^{p+1}\varepsilon = 2^{p+1}\varepsilon. \end{aligned} \tag{49}$$

Since we assume $\mu_c(\mathbb{R}^p) > 0$, it follows that $\beta(h) := \sup_{\gamma \in \Gamma_h} \mu_c(\gamma) > 0$, and moreover, $\beta(h) \to 0$ when $h \to 0$. Thus, using the concavity of x^{1+q}, we get

$$\Lambda_\mu(q;\Gamma_h) \leq \left(\sum_{\gamma \in \Gamma_h} \mu_c(\gamma)^{1+q} \right)^{1/q} \leq \left(\left[\frac{\mu_c(\mathbb{R}^p)}{\beta(h)} \right] \beta(h)^{1+q} \right)^{1/q}$$
$$\sim \mu_c(\mathbb{R}^p)^{1/q} \beta(h) \to 0 \text{ as } h \to 0.$$

But since $\Lambda_\mu(q;\Gamma_h) \underset{\cap}{\overset{\cup}{}} \Lambda_\mu(q;\Psi_h)$ and $\Lambda_{\hat{\mu}}(q;\Gamma_h) \underset{\cap}{\overset{\cup}{}} \hat{\Lambda}(q;\Psi_h)$, it will follow from the proof of (24) below that $\sup_{h \leq h_0} |\Lambda_{\hat{\mu}}(q;\Gamma_h) - \Lambda_\mu(q;\Gamma_h)|$ can be made arbitrarily small if h_0 is small enough and n large enough. In conjunction with (49), we obtain (23) for Case 2 when μ is not discrete.

The case $q = 0$ is analogous to $-1 < q < 0$. Instead of concavity of x^{1+q}, one uses the convexity of $x \log x$.

If $q \leq -1$, (23) is easy to prove if μ has finite support. We thus assume that μ has infinite and compact support. Then there exists $L > 0$ such that $\mathrm{supp}(\mu) \subset [-L, L]^p$. Introduce $\alpha(h) = \min\{\mu(\gamma); \gamma \in \Gamma_h, \mu(\gamma) > 0\}$, $\mathcal{H}_1 = \{h; \alpha(h) \leq \varepsilon\}$ and $\mathcal{H}_2 = (0,\infty) \backslash \mathcal{H}_1$. If $h \in \mathcal{H}_1$, then

$$\Lambda_{\hat{\mu}}(q;\Gamma_h) \leq \left(\min_{\gamma \in \Gamma_h} \hat{\mu}(\gamma)^{1+q} \right)^{1/q} \leq (\alpha(h) + \delta)^{(1+q)/q} \leq (\varepsilon + \delta)^{(1+q)/q},$$

and similarly $\Lambda_\mu(q;\Gamma_h) \leq \varepsilon^{(1+q)/q}$. Thus

$$|\Lambda_{\hat{\mu}}(q;\Gamma_h) - \Lambda_\mu(q;\Gamma_h)| \leq (\varepsilon + \delta)^{(1+q)/q} + \varepsilon^{(1+q)/q}. \tag{50}$$

If $h \in \mathcal{H}_2$, then

$$|\Lambda_{\hat{\mu}}(q;\Gamma_h)^q - \Lambda_\mu(q;\Gamma_h)^q| \leq \left(2 \left\lceil \frac{L}{h} \right\rceil \right)^p |(\varepsilon - \delta)^{q+1} - \varepsilon^{q+1}|. \tag{51}$$

Since μ has infinite support, $(0, h_0) \subset \mathcal{H}_1$ for some $h_0 > 0$, so the RHS of (51) is small when ε and δ are small. Further, $\Lambda_\mu^q(q;\Gamma_h) \geq 1$ for any $h \in \mathcal{H}_2$, so (51) implies that $|\Lambda_{\hat{\mu}}(q;\Gamma_h) - \Lambda_\mu(q;\Gamma_h)|$ is uniformly small for $h \in \mathcal{H}_2$. Thus (50) and (51) together imply (23).

Next, we turn to the HP functional. Since $\hat{\Lambda}(q;\Psi_h)$ and $\Lambda(q;\Psi_h)$ are monotone functions of h and $0 \leq \hat{\Lambda}(q;\Psi_h)$, $\Lambda(q;\Psi_h) \leq 1$ $\forall h \in (0,\infty)$, pointwise convergence

$$\hat{\Lambda}(q;\Psi_h) \xrightarrow{\text{a.s.}} \Lambda(q;\Psi_h) \tag{52}$$

will imply uniform convergence (24). Put $\hat{\mu}_h^{-i}(x) = \hat{\mu}^{-i}(B(x,h))$. Then

$$\hat{\Lambda}(q;\Psi_h) = \begin{cases} \exp((1/n)\sum_{i=1}^{n} \log \hat{\mu}_h^{-i}(X_i)), & q = 0 \\ ((1/n)\sum_{i=1}^{n} \hat{\mu}_h^{-i}(X_i)^q)^{1/q}, & q \neq 0, \end{cases}$$

and

$$\max_{1 \le i,j \le n} |\hat{\mu}_h^{-i}(X_j) - \hat{\mu}_h(X_j)| \le \frac{1}{n}. \tag{53}$$

If $q > 0$, (42) and (53) imply that for any $\xi > 0$,

$$\begin{aligned} \lim_{n\to\infty} \left(\frac{1}{n}\sum_{i=1}^{n}(\mu_h(X_i) - \xi)^q\right)^{1/q} &\le \lim_{n\to\infty} \hat{\Lambda}(q;\Psi_h) \\ &\le \lim_{n\to\infty} \left(\frac{1}{n}\sum_{i=1}^{n}(\mu_h(X_i) + \xi)^q\right)^{1/q} \end{aligned} \tag{54}$$

holds a.s. To obtain (52), we first use the ergodicity of $\{X_i\}$ and then let $\xi \to 0$.

If $q \le -1$, we assume that $\mathrm{supp}(\mu)$ is compact. Define $U_h = \{x; \mu_{h/2}(x) > 0\}$ and $\eta(h) = \inf\{\mu_h(x); x \in U_h\}$. Note that if $x \in U_h$ and $\|x' - x\| \le h/2$, then $\mu_{h/2}(x') \le \mu_h(x)$. The compact support of μ implies that U_h is totally bounded, so there exists an integer $\tilde{N} = \tilde{N}(h)$ and a chain $\{x_i\}_1^{\tilde{N}} \subset U_h$ such that $\sup_{x \in U_h} \min_{1 \le i \le \tilde{N}} \|x - x_i\| \le h/2$, and so $\eta(h) \ge \min_{1 \le i \le \tilde{N}} \mu_{h/2}(x_i) > 0$. Since $P(\{X_1, \ldots, X_n\} \subseteq U_h) = 1$, (54) holds for any $0 < \xi < \eta(h)$. Again, we prove (52) by first letting $n \to \infty$ and then $\xi \to 0$.

When $-1 < q < 0$, we only prove (52) when $\Lambda_\mu(q;h) > 0$ (the case $\Lambda_\mu(q;h) = 0$ is easier and omitted). Given $\varepsilon > 0$, choose $M = M(\varepsilon, h)$ as previously in the proof and put $\Omega = [-M-h, M+h]^p$. Similarly as for the case $q \le -1$, one proves that $\eta(h;\varepsilon) := \inf\{\mu_h(x); x \in U_h \cap \Omega\} > 0$. Thus, for any $0 < \xi < \eta(h;\varepsilon)$, one gets

$$\begin{aligned} &\lim_{n\to\infty} \left(\frac{1}{n}\sum_{X_i \in \Omega}(\mu_h(X_i) - \xi)^q + \frac{1}{n}\sum_{X_i \notin \Omega} \hat{\mu}_h^{-i}(X_i)^q\right)^{1/q} \\ &\quad \le \lim_{n\to\infty} \hat{\Lambda}(q;\Psi_h) \le \lim_{n\to\infty} \left(\frac{1}{n}\sum_{i=1}^{n}(\mu_h(X_i) + \xi)^q\right)^{1/q} \end{aligned} \tag{55}$$

almost surely, and in view of (53),

$$\begin{aligned}\frac{1}{n}\sum_{X_i\notin\Omega}\hat{\mu}_h^{-i}(X_i)^q &\le \frac{1}{n}\sum_{X_i\notin\Omega}(\hat{\mu}_h(X_i)-1/n)_+^q \le \frac{1}{n}\sum_{X_i\notin\Omega}(\hat{\mu}_h(X_i)/2)^q \\ &\le 2^{-q}\sum_{\gamma\in\Gamma_h'''}\hat{\mu}(\gamma)^{q+1} \xrightarrow{\text{a.s.}} 2^{-q}\sum_{\gamma\in\Gamma_h'''}\mu(\gamma)^{q+1}\le 2^{-q}\varepsilon, \qquad (56)\end{aligned}$$

with $\Gamma_h''' = \Gamma_h'''(M)$ defined as above, and the a.s. convergence in the last step is established as in (46). Combining (55) and (56), we obtain (52) by first using ergodicity, then letting $\xi\to 0$, and finally $\varepsilon\to 0$.

The case $q=0$ is similar to $-1<q<0$, so we only highlight the differences. Assume $\Lambda_\mu(0;\Psi_h)>0$ (which is the more difficult case). Then, if $0<\xi<\eta(h,\varepsilon)$,

$$\begin{aligned}&\lim_{n\to\infty}\left(\frac{1}{n}\sum_{X_i\in\Omega}\log(\mu_h(X_i)-\xi)+\frac{1}{n}\sum_{X_i\notin\Omega}\log(\hat{\mu}_h^{-i}(X_i))\right)\\ &\qquad\le \lim_{n\to\infty}\log\hat{\Lambda}(q;\Psi_h)\le \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\log(\mu_h(X_i)+\xi), \qquad (57)\end{aligned}$$

holds a.s., and

$$\begin{aligned}\frac{1}{n}\sum_{X_i\notin\Omega}\log(\hat{\mu}_h^{-i}(X_i)) &\ge \frac{1}{n}\sum_{X_i\notin\Omega}\log(\hat{\mu}_h(X_i)/2)\ge -\frac{\sharp\{X_i\notin\Omega\}\log 2}{n}\\ &\quad+\sum_{\gamma\in\Gamma_h'''}\hat{\mu}(\gamma)\log\hat{\mu}(\gamma)\xrightarrow{\text{a.s.}} -\mu(\Omega)\log 2\\ &\quad+\sum_{\gamma\in\Gamma_h'''}\mu(\gamma)\log\mu(\gamma), \qquad (58)\end{aligned}$$

and the RHS of (58) goes to 0 as $\varepsilon\to 0$. Equation (52) now follows from (57) and (58), by first using ergodicity of $\{X_i\}$ when $n\to\infty$, then letting $\xi\to 0$, and finally $\varepsilon\to 0$. ■

Proof of Theorem 5 We start proving that $\Lambda'_\mu(q;\Psi_h)$ and $\Lambda'_{\hat{\mu}}(q;\Psi_h)$ are non-decreasing and bounded functions of h. If $q\neq 0$,

$$\begin{aligned}\Lambda'_\mu(q;\Psi_h)^q &= h^{-p}\int\mu_h(x)^q\int_{B(x,h)}d\mu(y)dx\\ &=\int\int_{B(0,1)}\mu_h(y+vh)^q dv d\mu(y),\end{aligned}$$

using Fubini's Theorem and changing variables $x = y + vh$. For fixed y and v, the sets $B(y+vh, h) = y + hB(v, 1)$ increase with h, so $\mu_{h_1}(y + vh_1) \leq \mu_{h_2}(y + vh_2)$ if $h_1 < h_2$, and therefore $\Lambda'_\mu(q; \Psi_h)$ is a nondecreasing function of h. Since also $\mu_h(y+vh, h) \leq 1$, it follows that $\Lambda'_\mu(q; \Psi_h) \leq 2^{p/q}$ for all h. A completely analogous computation for $q = 0$ proves that $\Lambda'_\mu(0; \Psi_h)$ is a non-decreasing function of h, bounded by 1. Since μ can be replaced by $\hat{\mu}$ in all calculations, $\Lambda'_{\hat{\mu}}(q; \Psi_h)$ must be a non-decreasing and bounded function as well.

It remains to establish pointwise convergence $\Lambda'_{\hat{\mu}}(q; \Psi_h) \xrightarrow{\text{a.s.}} \Lambda'_\mu(q; \Psi_h)$. If $q > 0$, we let $M = M(\varepsilon)$ be the smallest number such that $\int_{\|x\| \geq M} d\mu(x) \leq \varepsilon$.

Put also $\Omega = [-M-h, M+h]^p$. Then

$$
\begin{aligned}
h^{-p} \int_{\Omega^c} \mu_h(x) dx &= h^{-p} \int_{\Omega^c} \int_{B(x,h)} d\mu(y) dx \\
&= \int \frac{|B(y,h) \cap \Omega^c|}{h^{-p}} d\mu(y) \\
&\leq 2^p \int_{\|y\| \geq M} d\mu(y) \leq 2^p \varepsilon,
\end{aligned}
$$

and this leads to

$$
\begin{aligned}
|\Lambda'_{\hat{\mu}}(q; \Psi_h)^q - \Lambda'_\mu(q; \Psi_h)^q| \leq h^{-p} &\int_\Omega |\hat{\mu}_h(x)^{1+q} - \mu_h(x)^{1+q}| dx \\
+ h^{-p} &\int_{\Omega^c} \mu_h(x)^{1+q} dx \\
+ h^{-p} &\int_{\Omega^c} \hat{\mu}_h(x)^{1+q} dx. \qquad (59)
\end{aligned}
$$

The first term on the RHS of (59) tends to 0 a.s. by dominated convergence, and the sum of the other two terms is upper bounded by

$$
\begin{aligned}
2^p(\mu(((-M,M)^p)^c) + \hat{\mu}(((-M,M)^p)^c)) &\xrightarrow{\text{a.s.}} 2^{p+1} \mu(((-M,M)^p)^c) \\
&\leq 2^{p+1} \varepsilon.
\end{aligned}
$$

Letting $\varepsilon \to 0$, we obtain pointwise convergence of $\Lambda'_{\hat{\mu}}(q; \Psi_h)$.

When $-1 < q < 0$, use $\hat{\mu}_h(x) \xrightarrow{\text{a.s.}} \mu_h(x)$ for all x and Fatou's Lemma to deduce $\liminf_{n \to \infty} \Lambda'_{\hat{\mu}}(q; \Psi_h)^q \geq \Lambda'_\mu(q; \Psi_h)^q$. On the other hand, the concavity of $x \to x^{1+q}$ and Jensen's Inequality implies $E(\hat{\mu}_h(x)^{1+q}) \leq \mu_h(x)^{1+q}$ for all x and whence $E\Lambda'_{\hat{\mu}}(q; \Psi_h)^q \leq \Lambda'_\mu(q; \Psi_h)^q$. Combining this with the a.s. lower bound for lim $\inf \Lambda'_{\hat{\mu}}(q;\Psi_h)^q$, we get pointwise convergence.

The case $q=0$ is competely analogous to $-1<q<0$, combining Fatou's Lemma and Jensen's Inequality (the convexity of $x \to x \log x$ is used).

Finally, when $q \leq -1$, Fatou's Lemma implies $\lim\inf_{n\to\infty} \Lambda'_{\hat{\mu}}(q;\Psi_h)^q \geq \Lambda'_{\mu}(q;\Psi_h)^q$, and this proves pointwise convergence when $\Lambda'_{\mu}(q;\Psi_h)^q = \infty$. If $\Lambda'_{\mu}(q;\Psi_h)^q < \infty$, the compact support of μ, (25) and dominated convergence imply $\lim_{n\to\infty} E\Lambda'_{\hat{\mu}}(q;\Psi_h)^q = \Lambda'_{\mu}(q;\Psi_h)^q$. Combining this with the a.s. lower bound for lim $\inf \Lambda'_{\hat{\mu}}(q;\Psi_h)^q$ we get pointwise convergence. ■

Proof of Proposition 3 Combining Corollary 2 and Theorem 3, we get $\Delta(\mathrm{supp}(\mu)) = \lim_{h\to 0}(\log \Lambda'(-1;\Psi_h)/\log h)$. But $\Lambda'(-1;\Psi_h) = h^p \lambda(\Omega_h)^{-1}$, where $\Omega_h = \{x; \mu_h(x) > 0\}$. The result follows since $\mathrm{supp}(\mu)^{h/2} \subset \Omega_h \subset \mathrm{supp}(\mu)^h$. ■

Proof of Corollary 4 Notice that

$$\hat{d}(q;\Psi) = d(q;\Psi) + \sum_{i=1}^{m} \log\ a(h_i)(x_i - \bar{x})/S_{xx} + \sum_{i=1}^{m} \varepsilon_i (x_i - \bar{x})/S_{xx},$$

where $x_i = \log h_i$, $\varepsilon_i = \log \Lambda_{\hat{\mu}}(q;\Psi_{h_i}) - \log \Lambda_{\mu}(q;\Psi_{h_i})$ and $S_{xx} = \sum_{i=1}^{m}(x_i - \bar{x})^2$. Since $a(h_i) \equiv a$ the bias term vanishes, and further $\varepsilon_i \xrightarrow{\text{a.s.}} 0$, $i = 1,\ldots,m$ by Theorem 4. The expansion for $\hat{d}(q;\Gamma)$ is completely analogous. ■

Acknowledgements

The authors are most grateful to suggestions made by two referees, which led us to reorganize the paper, improve notation and add several relevant references.

References

Beck, C. (1990) Upper and lower bounds on the Rényi Dimensions and the uniformity of multifractals, *Physica D*, 41, 67–78, North-Holland.

Billingsley, P. (1965) *Ergodic Theory and Information*, Wiley, New York.

Cheng, B. and Tong, H. (1992) On consistent nonparametric order determination and chaos, *J.R. Statist. Soc. B*, 54, 471–472.

Cutler, C. (1991) Some results on the behaviour and estimation of fractal dimension of distributions on attractors, *Journal of Statistical Physics*, **62**, 651–708.

Cutler, C. (1993) A review of the theory and estimation of fractal dimension. Dimension estimation and modèls, Volume I in *Nonlinear Time Series and Chaos*, Edited by Howell Tong, pp. 1–107. World Scientific.

Cutler, C. (1994) A theory of correlation dimension for stationary time series, *Phil. Trans. Roy. Soc. (Lond.)*, **A348**, 343–355.

Denker, M. and Keller, G. (1986) Rigorous statistical procedures for data from dynamical systems, *Journal of Statistical Physics*, **44**, 67–93.

Falconer, K. (1990) *Fractal Geometry*, John Wiley & Sons, Chichester.

Falconer, K. (1997) *Techniques in Fractal Geometry*, John Wiley & Sons, Chichester.

Federer, H. (1969) *Geometric Measure Theory*, Springer Verlag, Berlin.

Frigyesi, A. (1994) Testing Singularity and Estimating Hausdorff Dimensions Using Kernel Estimates. *Master's Thesis*. Technical report 1994:E6, Dept. of Mathematical Statistics, University of Lund.

Frigyesi, A. and Hössjer, O. (1996) Kernel estimates of dimension spectra for multifractal measures. *Technical report* 1996:14, Dept. of Mathematical Statistics, University of Lund.

Frigyesi, A. and Hössjer, O. (1998) A test for singularity, *Statistics and Probability Letters*, **40**, 215–226.

Grassberger, P. and Procaccia, I. (1983) Characterization of strange attractors, *Physics Review Letters*, **50**, 346–349.

Guerin, C.-A. (1998) Diffusion sur les Mesures Fractales, *Ph.D. Dissertation*, Université de Provence-aix-Marseille i U.F.R Sciences.

Hentschel, H. G. E. and Procaccia, I. (1983) The infinite number of fractals and strange attractors, *Physica D*, **8**, 435–444.

Mandelbrot, B. B. (1982) *The Fractal Geometry of Nature*. W. H. Freedman and Co., San Fransisco.

Muzy, J. F., Bacry, E. and Arneodo, A. (1994) The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos*, **4**(2) 245–302. World Scientific Publishing Company.

Pesin, Y. (1993) On rigorous definition of correlation dimension and generalized spectrum for dimensions, *Journal of Statistical Physics*, 7.

Pesin, Y. (1997) *Dimension Theory in Dynamical Systems*. Rigorous results and applications. University of Chicago Press, Chicago.

Roberts, A. J. (1996) Use the information dimension, not the Hausdorff, Preprint, Dept. of Mathematics and Computing, Univ. of Southern Queensland, Australia.

Rényi, A. (1970) *Probability Theory*. Akadémiai Kiadó, Budapest.

Serinko, R. J. (1994) A consistent approach to least squares estimation of correlation dimension in weak Bernoulli dynamical systems, *Annals of Applied Probability*, **4**, 1234–1254.

Serinko, R. J. (1996) Ergodic theorems arising in correlation dimension estimation, *Journal of Statistical Physics*, **85**, 25–40.

van de Water, W. and Scram, P. (1988) Generalized dimensions from near-neighbor information, *Physical Review A*, **37**(8), 3118–3125.

Wand, M. and Jones, M. (1995) *Kernel Smoothing*, Chapman and Hall, London.