Bernoulli **9**(1), 2003, 97–135

Wavelet analysis of conservative cascades

SIDNEY RESNICK[1*], GENNADY SAMORODNITSKY[1**], ANNA GILBERT[2*] and WALTER WILLINGER[2**]

[1]*School of Operations Research and Industrial Engineering and Department of Statistical Science, Cornell University, Ithaca NY 14853, USA. E-mail: *sid@orie.cornell.edu; **gennady@orie.cornell.edu*
[2]*AT&T Labs – Research, 180 Park Avenue, Florham Park NJ 07932, USA. E-mail: *agilbert@research.att.com; **walter@research.att.com*

A conservative cascade is an iterative process that fragments a given set into smaller and smaller pieces according to a rule which preserves the total mass of the initial set at each stage of the construction almost surely and not just in expectation. Motivated by the importance of conservative cascades in analysing multifractal behaviour of measured Internet traffic traces, we consider wavelet-based statistical techniques for inference about the cascade generator, the random mechanism determining the redistribution of the set's mass at each iteration. We provide two estimators of the structure function, one asymptotically biased and one not, and prove consistency and asymptotic normality in a range of values of the argument of the structure function less than a critical value. Simulation experiments illustrate the asymptotic properties of these estimators for values of the argument both below and above the critical value. Beyond the critical value, the estimators are shown not to be asymptotically consistent.

Keywords: asymptotic normality; conservative cascade; consistency; generator; Internet; MKP condition; martingale; multifractals; networks; protocols; structure function; wavelets

1. Introduction

A *multiplicative cascade* is an iterative process that fragments a given set into smaller and smaller pieces according to some geometric rule and, at the same time, distributes the total mass of the given set according to another rule. The limiting object generated by such a procedure generally gives rise to a singular measure or *multifractal* – a mathematical construct that is able to capture the highly irregular and intermittent behaviour associated with many naturally occurring phenomena, such as fully developed turbulence (see Kolmogorov 1941; Mandelbrot 1974; Frisch and Parisi 1985; Meneveau and Srinivasan 1987; and references therein), spatial rainfall (Gupta and Waymire 1993), the movements of stock prices (Mandelbrot 1998) and Internet traffic dynamics (Riedi and Lévy Véhel 1997; Feldmann *et al*. 1998).

The *generator* of a cascade determines the redistribution of the set's total mass at every iteration; it can be deterministic or random. Cascade processes with the property that the generator preserves the total mass of the initial set at each stage of the construction almost surely and not just in expectation are called *conservative cascades* and are the main focus of this paper. Originally introduced by Mandelbrot (1990) (also in the turbulence context),

1350–7265

conservative cascades have recently been considered in Feldmann *et al*. (1998) for use in describing the observed highly irregular small-time scaling behaviour of measured Internet traffic traces. In particular, Feldmann *et al*. (1998) build on empirical evidence that measured Internet traffic is consistent with multifractal behaviour by illustrating that 'data networks appear to act as conservative cascades!'. They demonstrate that multiplicative and measure-preserving structure becomes most apparent when analysing measured Internet traces at a particular layer within the well-defined protocol hierarchy of today's Internet Protocol (IP) based networks, namely the Transport Control Protocol (TCP) layer and at the level of individual TCP connections, and that this structure can often be recovered at the aggregate level (i.e., when considering the superposition process consisting of all IP packets generated by all active TCP connections) and tends to cause aggregate Internet traffic to exhibit multifractal-like behaviour. Well short of providing a physical explanation for the all-important networking question of why packets within individual TCP connections conform to a conservative cascade, the work of Feldmann *et al*. (1998) is empirical in nature and relies on a number of heuristics for inferring multifractal behaviour from traces of measured Internet traffic. However, to provide a more solid statistical basis for empirical studies of multifractal phenomena, progress in the area of statistical inference for multiplicatively generated multifractals is crucial.

In this paper we contribute to the effort of providing rigorous techniques for multifractal analysis by investigating wavelet-based estimators for conservative cascades (i.e., for the class of multifractal processes generated by conservative cascades) and studying their large-sample properties. In this context, the special appeal of relying on wavelet-based inference techniques lies in the wavelet's natural abilities to detect and analyse various scaling-related properties of an underlying signal or time series. Moreover, in contrast to time-domain-based methods for investigating multifractal scaling behaviour, wavelet-based techniques lend themselves in a natural way to a systematic investigation of certain underlying non-stationarity features in the data (see, for example, Abry and Veitch 1998), and we will allude to this ability towards the end of this paper.

Irrespective of the method used, the inference problem for conservative cascades consists of deducing from a single realization of the cascade process the distribution of the cascade generator that was presumably used to generate the sample or signal at hand. Intuitively, the generator's distribution can be inferred from the degree of variability and intermittency exhibited locally in time by the signal under consideration. It can be expressed mathematically in terms of the local Hölder exponents which in turn characterise the singularity behaviour of a signal locally in time. Moreover, since the local Hölder exponent at a point in time t_0 describes the local scaling behaviour of the signal as we look at smaller and smaller neighbourhoods around t_0, a wavelet-based analysis that fully exploits the time- and scale-localization ability of wavelets proves convenient and is tailor-made for our purpose. On the one hand, we exploit here the fact that the singularity behaviour of a process can (under certain assumptions) be fully recovered by studying the singularity behaviour in the wavelet domain; that is, by investigating the (possibly) time-dependent scaling properties of the wavelet coefficients associated with the underlying process in the fine-time scale limit. On the other hand, using Haar wavelets, the discrete wavelet transform of a conservative cascade can be explicitly expressed in terms of the cascade's generator (see, for example, Gilbert *et al*. 1999) and hence provides a promising setting for relating the local scaling behaviour of the sample to the distribution of the underlying conservative cascade generator. In particular, we relate the distribution of the generator to an invariant of the cascade, namely the *structure function* or *modified cumulant generating function* – also known as the *Mandelbrot–Kahane–Peyrière* (*MKP*) *function* (Holley and Waymire 1992) – and study the statistical properties (i.e., asymptotic consistency, asymptotic normality, confidence intervals) of two wavelet-based estimators of this function.

Although the results in this paper have been largely motivated by our empirical investigations into the multifractal nature of measured Internet traffic (Feldmann *et al*. 1998; Gilbert *et al*. 1999), we have clearly benefited from the recent random cascade work of Ossiander and Waymire (2000). Compared to the conservative cascades considered in this paper, random cascades are multiplicative processes with generators that preserve the total mass of the initial set only in expectation and not almost surely. This apparently minor difference ensures independence within and across the different stages of a random cascade construction but gives rise to subtle dependencies inherent in conservative cascades. Ossiander and Waymire (2000) study the large-sample asymptotics of estimators that are defined in the time domain rather than in the wavelet domain and allow for a rigorous statistical analysis of the scaling behaviour exhibited by random cascades (for related work, see Troutman and Vecchia 1999). While the large-sample properties of the time-domain-based estimators considered in Ossiander and Waymire (2000) and of the wavelet-based estimators studied in this paper are very similar, their potential advantages, disadvantages and pitfalls when implementing and using them in practice require further studies. However, in combination, these different estimators provide a set of statistically rigorous techniques for multifractal analysis of highly irregular and intermittent data that are assumed to be generated by certain types of multiplicative processes or cascades.

The rest of the paper is organized as follows. Sections 2–4 contain the basic facts about conservative cascades, their wavelet transforms, and some related quantities that are needed later in the paper. Section 5 discusses the critical constants, and Section 6 is concerned with certain martingales and leads into Section 7 where subcritical asymptotics (that is, asymptotics for values of the argument below the critical value) and strong consistency of our two wavelet-based estimators are established. (Subcritical) asymptotic normality of the estimators is explained and illustrated with some simulated data in Section 8, and Section 9 deals with some supercritical asymptotics when the value of the argument exceeds the critical value. The values of the estimators at large values of the argument of the structure function are uninformative and misleading, thus providing some practical guidance for properly interpreting the plots associated with the estimation procedure. We conclude in Section 10 with some interesting observations and open problems.

2. The conservative cascade

We now summarize the basic facts about the conservative cascade.

Consider the binary tree. Nodes of the tree at depth l will be indicated by

$(j_1, \ldots, j_l) \in \{0, 1\}^l$. Alternatively, we consider successive subdivisions of the unit interval [0, 1]. After subdividing l times we have equal subintervals of length 2^{-l} indicated by

$$I(j_1, \ldots, j_l) = \left[\sum_{k=1}^{l} \frac{j_k}{2^k}, \sum_{k=1}^{l} \frac{j_k}{2^k} + \frac{1}{2^l}\right), \qquad (j_1, \ldots, j_l) \in \{0, 1\}^l. \tag{2.1}$$

An infinite path through the tree is denoted by

$$\mathbf{j} = (j_1, j_2, \ldots) \in \{0, 1\}^\infty,$$

and the first l entries of $\mathbf{j}$ are denoted by

$$\mathbf{j}|l = (j_1, \ldots, j_l).$$

We will sometimes write, when convenient,

$$\mathbf{j}|l, j_{l+1} = (j_1, \ldots, j_l, j_{l+1}).$$

The *conservative cascade* is a random measure on the Borel subsets of [0, 1] which may be constructed in the following manner. Suppose we are given a random variable W, called the *cascade generator*, which has range [0, 1] and which is symmetric about $\frac{1}{2}$ so that $W \stackrel{d}{=} 1 - W$. The symmetry implies that $E(W) = \frac{1}{2}$. We assume the random variable is not almost surely equal to $\frac{1}{2}$. There is a family of identically distributed random variables

$$\{W(\mathbf{j}|l), \mathbf{j} \in \{0, 1\}^\infty, l \geqslant 1\},$$

each of which is identically distributed as W. These random variables satisfy the *conservative property*

$$W(\mathbf{j}|l, 1) = 1 - W(\mathbf{j}|l, 0). \tag{2.2}$$

Random variables associated with different depths of the tree are independent, and random variables of the same depth which have different antecedents in the tree are likewise independent. Dependence of random variables having the same depth is expressed by (2.2). The conservative cascade is the random measure μ_∞ defined by

$$\mu_\infty(I(\mathbf{j}|l)) = \prod_{i=1}^{l} W(\mathbf{j}|i). \tag{2.3}$$

Note the conservative property entails that

$$\mu_\infty(I(\mathbf{j}|l, 0)) + \mu_\infty(I(\mathbf{j}|l, 1)) = \mu_\infty(I(\mathbf{j}|l)), \tag{2.4}$$

so that the weight of two offspring equals the weight of the parent. This implies

$$\sum_{\mathbf{j}|l} \mu_\infty(I(\mathbf{j}|l)) = 1. \tag{2.5}$$

3. Wavelet coefficients

We compute the wavelet transform

$$d_{-l,n} = \int_0^1 \psi_{-l,n}(x)\mu_\infty(\mathrm{d}x), \qquad n = 0, \ldots, 2^l - 1;\ l \geqslant 1, \tag{3.1}$$

using the Haar wavelets

$$\psi_{-l,n}(x) := \begin{cases} 2^{l/2}, & \text{if } 2n/2^{l+1} \leqslant x < (2n+1)/2^{l+1}, \\ -2^{l/2}, & \text{if } (2n+1)/2^{l+1} \leqslant x < (2n+2)/2^{l+1}. \end{cases} \tag{3.2}$$

We have, by examining where the Haar wavelet is constant, that

$$d_{-l,n} = 2^{l/2}\left(\mu_\infty\left(\left[\frac{2n}{2^{l+1}}, \frac{2n+1}{2^{l+1}}\right)\right) - \mu_\infty\left(\left[\frac{2n+1}{2^{l+1}}, \frac{2n+2}{2^{l+1}}\right)\right)\right).$$

Now suppose that $\sum_{k=1}^{l} j_k/2^k = n/2^l$. Then we have from (2.3),

$$\begin{aligned} d_{-l,n} &= 2^{l/2}\left[\prod_{i=1}^{l} W(\mathbf{j}|i)W(\mathbf{j}|l, 0) - \prod_{i=1}^{l} W(\mathbf{j}|i)W(\mathbf{j}|l, 1)\right] \\ &= 2^{l/2}\prod_{i=1}^{l} W(\mathbf{j}|i)[W(\mathbf{j}|l, 0) - W(\mathbf{j}|l, 1)]; \end{aligned}$$

using the conservative property (2.2), this is

$$d_{-l,n} = 2^{l/2}\prod_{i=1}^{l} W(\mathbf{j}|i)[2W\mathbf{j}|l, 0) - 1], \tag{3.3}$$

for $n = 0, 1, \ldots, 2^l - 1$. Sometimes where convenient, we will also write

$$d_{-l,n} = d(-l, \mathbf{j}|l) = 2^{l/2}\prod_{i=1}^{l} W(\mathbf{j}|i)[2W\mathbf{j}|l, 0) - 1]. \tag{3.4}$$

4. Notation

Before continuing the analysis, in this section we present some notation for ease of reference. We seek to estimate the distribution of the cascade generator W, and this will be accomplished if we estimate

$$c(q) := 2\mathrm{E}(W^q), \qquad q > 0; \tag{4.1}$$

equivalently, we could estimate the *structure function*

$$\tau(q) = 1 + \log_2 \mathrm{E}(W^q) = \log_2 c(q). \tag{4.2}$$

The structure function will be estimated using estimators constructed from the process

$$Z(q,\, l) = \sum_{\mathbf{j}|l} \prod_{i=1}^{l} W(\mathbf{j}|i)^q |2W(\mathbf{j}|l,\, 0) - 1|^q, \tag{4.3}$$

noting from (3.3) that

$$Z(q,\, l) = \frac{1}{2^{ql/2}} \sum_{n=0}^{2^l-1} |d_{-l,n}|^q. \tag{4.4}$$

Our analysis rests on the process $M(q,\, l)$, which we will show to be a martingale and which is defined as

$$M(q,\, l) = \frac{1}{c(q)^l} \sum_{\mathbf{j}|l} \prod_{i=1}^{l} W(\mathbf{j}|i)^q, \qquad q > 0,\ l \geqslant 1; \tag{4.5}$$

note that the normalization makes

$$\mathrm{E}(M(q,\, l)) = 1.$$

Where no confusion can result, we sometimes, for convenience, write $c = c(q)$. The following constant functions are also needed:

$$b(q) = \mathrm{E}|2W - 1|^q, \tag{4.6}$$

$$a(q) = \frac{c(2q)}{c^2(q)} = \frac{\mathrm{E}(W^{2q})}{2(\mathrm{E}(W^q))^2}, \tag{4.7}$$

$$a_r(q) = \frac{c(rq)}{c^r(q)} = \frac{2^{1-r}\mathrm{E}(W^{rq})}{(\mathrm{E}(W^q))^r}. \tag{4.8}$$

Note that $a(q) = a_2(q)$. Finally, we need three variances:

$$\sigma_1^2(q) := \frac{1}{c^2} \operatorname{var}(W^q + (1 - W)^q), \tag{4.9}$$

$$\sigma_2^2(q) := \frac{1}{b^2} \operatorname{var}(|2W - 1|^q), \tag{4.10}$$

$$\sigma_3^2(q) := \frac{1}{b^2} \operatorname{var}\left(\frac{W_1^q}{c}|2W_2 - 1|^q + \frac{(1 - W_1)^q}{c}|2W_3 - 1|^q - |2W_1 - 1|^q\right), \tag{4.11}$$

where W_i, $i = 1, 2, 3$ are independent and identically distributed (i.i.d.), having the distribution of the cascade generator.

It is convenient to define $W = \mathrm{e}^{-Y}$ so that the Laplace transform of Y is

$$\phi(q) := \mathrm{E}\mathrm{e}^{-qY} = \mathrm{E}(W^q) \tag{4.12}$$

and

$$a_r(q) = \frac{2^{1-r}\phi(rq)}{\phi^r(q)}.$$

5. Critical constants

We now define the quantity

$$q_* := \sup\{q > 0 : a(q) < 1\} \tag{5.1}$$

so that for $q < q_*$ we have $a(q) < 1$. It will turn out that when $q < q_*$, the sequence $\{M(q, l), l \geqslant 1\}$ is an L_2-bounded and uniformly integrable martingale, and this is the easiest case to analyse. It is always the case that $q_* \geqslant 1$, which follows from the fact that

$$a(1) = \frac{\mathrm{E}(W^2)}{2(\mathrm{E}(W))^2} = \frac{\mathrm{E}(W^2)}{2(\frac{1}{2})^2} = 2\mathrm{E}(W^2)$$

so that

$$a(1) = 2\big(\mathrm{var}(W) + \tfrac{1}{4}\big) = 2\mathrm{E}\big(W - \tfrac{1}{2}\big)^2 + \tfrac{1}{2} \leqslant 2 \cdot \big|1 - \tfrac{1}{2}\big|^2 + \tfrac{1}{2} = 1.$$

Let W be the cascade generator and define

$$X_q = \frac{W^q}{\mathrm{E}W^q}, \qquad q > 0,$$

so that $\mathrm{E}X_q = 1$. The Mandelbrot–Kahane–Peyrière (MKP) condition (see Kahane and Peyrière, 1976) is satisfied for q if

$$\mathrm{E}(X_q \log_2 X_q) < 1 \tag{5.2}$$

if and only if

$$\frac{q}{\mathrm{E}(W^q)}\mathrm{E}(W^q \log W) - \log \mathrm{E}(W^q) < \log 2 \tag{5.3}$$

if and only if

$$q(\log \phi)'(q) - \log \phi(q) < \log 2. \tag{5.4}$$

Define

$$\Lambda^* := \{q : \mathrm{E}(X_q \log_2 X_q) < 1\}.$$

Then Λ^* is an interval and we define the second critical constant

$$q^* := \sup \Lambda^*. \tag{5.5}$$

Why is q^* considered a critical quantity? It turns out that the martingale $\{M(q, l), l \geqslant 1\}$ converges as $l \to \infty$ to

$$M(q, \infty) = \begin{cases} 0, & \text{if } q \geqslant q^*, \\ \text{something non-degenerate}, & \text{if } q < q^*. \end{cases}$$

The associated martingale is uninformative asymptotically when $q > q^*$.

The two critical constants are related numerically by the inequality

$$\max(1, q^*/2) \leqslant q_* \leqslant q^*. \tag{5.6}$$

To see why the inequality $q_* \leqslant q^*$ in (5.6) is true, it suffices to show that if $q > 0$ satisfies $a(q) < 1$, then the MKP condition is satisfied for this q. However, if $a(q) < 1$, then

$$\begin{aligned}\log 2 > \log\phi(2q) - 2\log\phi(q) &= \log\phi(2q) - \log\phi(q) - \log\phi(q) \\ &= \int_q^{2q} (\log\phi)'(s)\mathrm{d}s - \log\phi(q),\end{aligned}$$

and since $\log\phi$ is convex, $(\log\phi)'$ is increasing and the foregoing is bounded below by

$$(\log\phi)'(q)\int_q^{2q} \mathrm{d}s - \log\phi(q) = q(\log\phi)'(q) - \log\phi(q).$$

The conclusion that

$$\log 2 > q(\log\phi)'(q) - \log\phi(q)$$

is equivalent to the MKP condition holding by (5.4).

On the other hand, suppose that $q_* < \infty$. Since $a(q_*) = 1$, we have, in the same way as above,

$$\begin{aligned}\log 2 &= \log\phi(2q_*) - 2\log\phi(q_*) = 2(\log\phi(2q_*) - \log\phi(q_*)) - \log\phi(2q_*) \\ &= 2\int_{q_*}^{2q_*} (\log\phi)'(s)\mathrm{d}s - \log\phi(2q_*) \\ &\leqslant 2(\log\phi)'(2q_*)\int_{q_*}^{2q_*} \mathrm{d}s - \log\phi(2q_*) \\ &= 2q_*(\log\phi)'(2q_*) - \log\phi(2q_*).\end{aligned}$$

Therefore, the MKP condition does not hold for $2q_*$, and so $q^* \leqslant 2q_*$.

Example 1. Suppose W is uniformly distributed on [0, 1]. In this case $\mathrm{E}(W^q) = 1/(1+q)$ and so

$$a(q) = \frac{1}{2}\left(1 + \frac{q^2}{2q+1}\right)$$

and

$$q_* = 1 + \sqrt{2} \approx 2.4.$$

Likewise, q^* satisfies the equation

$$\log(1+q) - \frac{q}{1+q} = \log 2$$

and so $q^* \approx 3.311$.

Example 2. Suppose, more generally, that W has the beta distribution with mean $\frac{1}{2}$ (i.e., the shape parameters α and β are equal). Then

$$\mathrm{E}(W^q) = \frac{\Gamma(2\alpha)\Gamma(\alpha+q)}{\Gamma(\alpha)\Gamma(2\alpha+q)} \tag{5.7}$$

and q_* satisfies

$$4^{-\alpha}\sqrt{\pi}\Gamma(\alpha+2q)\Gamma(2\alpha+q)^2 - \Gamma(\alpha+\tfrac{1}{2})\Gamma(2\alpha+2q)\Gamma(\alpha+q)^2 = 0.$$

Example 3. Suppose W has the two-point distribution concentrating mass $\frac{1}{2}$ at p, $1-p$ for some $0 \leqslant p < \frac{1}{2}$. Then

$$\mathrm{E}(W^q) = \tfrac{1}{2}(p^q + (1-p)^q) \tag{5.8}$$

and

$$a(q) = 1 - \frac{2p^q(1-p)^q}{(p^q+(1-p)^q)^2}.$$

Note in this case that $a(q)\uparrow 1$ as $q\uparrow\infty$ so $q_* = q^* = \infty$.

Example 4. If W does not have a two-point distribution but nevertheless has an atom of size $p_1 < \frac{1}{2}$ at 1 (and hence by symmetry there is an atom of the same size at 0) we have $q^* < \infty$ (and, hence, also $q_* < \infty$). To see this, we express condition (5.3), when $q > 1$, in the equivalent form

$$\frac{\mathrm{E}(W^q \log_2 W^q)}{\mathrm{E}(W^q)\log_2(2\mathrm{E}(W^q))} > 1. \tag{5.9}$$

Note that if W does not have a two-point distribution, then for $q > 1$, we have $P[W^q < W] > 0$ and $\mathrm{E}(W^q) < \mathrm{E}(W) = \frac{1}{2}$, so $\log_2(2E(w^q)) < 0$, which explains the sign reversal in (5.9) compared with (5.3).

By the dominated convergence theorem the numerator in (5.9) converges to 0 as $q \to \infty$, while the denominator converges to $P[W=1]\log_2(2P[W=1]) \neq 0$. Hence, (5.9) fails for large q.

Based on the experience of Examples 3 and 4, it is natural to wonder how common it can be that $q^* = \infty$. This is discussed in the next proposition.

Proposition 5.1. *Unless W has a two-point distribution, it must be the case that $q^* < \infty$.*

Proof. Because of Example 3, we may assume that W does *not* have atoms at 0 and 1. Let $p \in [0, \frac{1}{2})$ be the leftmost point of the support of the distribution of W. Then $1-p$ is the rightmost point of the support of the distribution of W. For $0 < \rho < 1$, we have

$$\theta(\rho) := P[W \geqslant \rho(1-p)] > 0.$$

Since the distribution of W is not a two-point distribution, $\lim_{\rho\to 1}\theta(\rho) < \frac{1}{2}$. Thus, we can find and fix a value of $0 < \rho < 1$ such that

$$\theta(\rho) < \tfrac{1}{2}, \qquad 0 < \rho < 1.$$

For this value of ρ, it is convenient to set

$$\delta(\rho) := \delta = \rho(1 - p).$$

Apply Jensen's inequality with the convex function $g(x) = x \log x$, $x > 0$, to obtain

$$\begin{aligned}
\mathrm{E}\left(\frac{W^q}{c(q)} \log_2\left(\frac{W^q}{c(q)}\right) \cdot 1_{[W \geqslant \delta]}\right) &= \theta(\rho)\mathrm{E}\left(\frac{W^q}{c(q)} \log_2 \frac{W^q}{c(q)} \middle| W \geqslant \delta\right) \\
&\geqslant \theta(\rho)\mathrm{E}\left(\frac{W^q}{c(q)} \middle| W \geqslant \delta\right) \log_2 \mathrm{E}\left(\frac{W^q}{c(q)} \middle| W \geqslant \delta\right) \\
&= \mathrm{E}\left(\frac{W^q}{2\mathrm{E}(W^q)} 1_{[W \geqslant \delta]}\right) \cdot \log_2 \mathrm{E}\left(\frac{W^q}{2\mathrm{E}(W^q)} \middle| W \geqslant \delta\right). \qquad (5.10)
\end{aligned}$$

Also we have

$$\lim_{q\to\infty} \mathrm{E}\left(\frac{W^q}{c(q)} 1_{[W \geqslant \delta]}\right) = \tfrac{1}{2}. \qquad (5.11)$$

To verify (5.11), note that

$$\tfrac{1}{2} = \mathrm{E}\left(\frac{W^q}{c(q)}\right) = \mathrm{E}\left(\frac{W^q}{c(q)} 1_{[W \geqslant \delta]}\right) + \mathrm{E}\left(\frac{W^q}{c(q)} 1_{[W < \delta]}\right),$$

and

$$0 \leqslant \lim_{q\to\infty} \frac{\mathrm{E}(W^q 1_{[W<\delta]})}{\mathrm{E}(W^q 1_{[W \geqslant \delta]})} \leqslant \lim_{q\to\infty} \frac{1}{P[W \geqslant \delta]} \mathrm{E}\left(\frac{W}{\delta}\right)^q 1_{[W/\delta < 1]} = 0,$$

by dominated convergence. Thus from (5.10) and (5.11), we conclude that

$$\liminf_{q\to\infty} \mathrm{E}\left(\frac{W^q}{c(q)} \log_2 \frac{W^q}{c(q)} \cdot 1_{[W \geqslant \delta]}\right) \geqslant \frac{1}{2} \cdot \log_2\left(\frac{1}{2\theta(\rho)}\right) =: h > 0, \qquad (5.12)$$

because $\theta(\rho) < \tfrac{1}{2}$.

We also claim that

$$\lim_{q\to\infty} \mathrm{E}\left(\left|\frac{W^q}{c(q)} \log \frac{W^q}{c(q)}\right| 1_{[W<\delta]}\right) = 0. \qquad (5.13)$$

To verify (5.13), note the expectation is the same as

$$\mathrm{E}\left(\left|\frac{W^q}{c(q)} \log \frac{W^q}{c(q)}\right| 1_{[W^q/c(q) < \delta^q/c(q)]}\right).$$

Provided that

$$\lim_{q\to\infty} \frac{\delta^q}{c(q)} = 0, \tag{5.14}$$

we obtain for any $\epsilon < \mathrm{e}^{-1}$, by the monotonicity of $|x \log x|$ in $(0, \mathrm{e}^{-1})$, that the expectation is bounded by $|\epsilon \log_2 \epsilon|$ for q so large that $\delta^q / c(q) < \epsilon$. So it remains to check (5.14), or equivalently to check

$$\lim_{q\to\infty} \mathrm{E}\left(\frac{W^q}{\delta^q}\right) = 0.$$

However, by Fatou's lemma

$$\liminf_{q\to\infty} \mathrm{E}\left(\frac{W^q}{\delta^q}\right) \geqslant \mathrm{E}\left(\liminf_{q\to\infty}\left(\frac{W}{\delta}\right)^q 1_{[W\geqslant\delta]}\right) = \infty,$$

since $P[W > \delta] > 0$ (otherwise, the definition of p would be contradicted).

Our conclusion from (5.12) and (5.13) is that, for all large q,

$$\mathrm{E}\left(\frac{W^q}{c(q)}\log_2\left(\frac{W^q}{c(q)}\right)\right) = \mathrm{E}\left(\frac{W^q}{c(q)}\log_2\left(\frac{W^q}{c(q)}\right)\right)1_{[W\geqslant\delta]} + \mathrm{E}\left(\frac{W^q}{c(q)}\log_2\left(\frac{W^q}{c(q)}\right)\right)1_{[W<\delta]} > 0.$$

Thus

$$\mathrm{E}\frac{W^q}{2EW^q}\left(\log_2\frac{W^q}{\mathrm{E}(W^q)} - \log_2 2\right) = \frac{1}{2}\mathrm{E}X_q \log_2 X_q - \frac{1}{2} > 0,$$

and so the MKP condition (5.3) fails for all large q. Therefore, $q^* < \infty$. □

We need the following properties of the function $a_r(q)$.

Proposition 5.2.

(i) *For any fixed* $r > 1$, *the function* $a_r(q)$ (*and therefore* $a(q)$) *is strictly increasing in* $q > 0$.

(ii) *For any fixed* $q > 0$, *the function* $\log a_r(q)$ *is strictly convex in the region* $r > 1$.

(iii) *If* q *satisfies the MKP condition, then*

$$\frac{\mathrm{d}}{\mathrm{d}r}\log a_r(q)\bigg|_{r=1} < 0,$$

and there exists $r_0 \in (1, 2)$ *such that*

$$a_{r_0}(q) < a_1(q) = 1. \tag{5.15}$$

(iv) *If the MKP condition fails for* q *and the inequalities in* (5.2)–(5.4) *are reversed to become strictly greater than, we have*

$$\frac{\mathrm{d}}{\mathrm{d}r} \log a_r(q)\bigg|_{r=1} > 0, \tag{5.16}$$

and there exists $0 < r_1 < 1$ *and* $a_{r_1}(q) < a_1(q) = 1$.

Proof. (i) Recall the definition of ϕ from (4.12). For fixed $r > 1$, if we differentiate with respect to q, we get

$$\left(\frac{\phi(rq)}{\phi^r(q)}\right)' = \frac{\phi^r(q) r \phi'(rq) - \phi(rq) r \phi^{r-1}(q)\phi'(q)}{\phi^{2r}(q)}.$$

This is positive if and only if

$$\phi(q)\phi'(rq) > \phi(rq)\phi'(q)$$

or

$$\frac{\phi'(rq)}{\phi(rq)} > \frac{\phi'(q)}{\phi(q)}.$$

Since $r > 1$, it suffices to show ϕ'/ϕ is strictly increasing, which is true if its derivative is strictly positive. The derivative is

$$\frac{\phi(q)\phi''(q) - (\phi'(q))^2}{\phi^2(q)},$$

which is strictly positive if and only if

$$\phi(q)\phi''(q) > (\phi'(q))^2, \tag{5.17}$$

that is, if and only if

$$\mathrm{E}\big(\mathrm{e}^{-qY}\big)\mathrm{E}\big(Y^2\mathrm{e}^{-qY}\big) > \Big(\mathrm{E}\Big(Y\mathrm{e}^{-q/2Y} \cdot \mathrm{e}^{-q/2Y}\Big)\Big)^2$$

which follows from the Cauchy–Schwarz inequality.

(ii) Fix $q > 0$ and check that

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}(\log a_r(q)) = \frac{q^2}{\phi^2(rq)}[\phi''(rq)\phi(rq) - (\phi'(rq))^2],$$

which is positive by (5.17).

(iii) For fixed $q > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}r} \log a_r(q) = q(\log\phi)'(qr) - \log 2 - \log\phi(q),$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}r}\log a_r(q)\bigg|_{r=1} = \left(\frac{1}{a_r(q)}\frac{\mathrm{d}}{\mathrm{d}r}a_r(q)\right)\bigg|_{r=1} = q(\log\phi)'(q) - \log\phi(q) - \log 2 < 0.$$

Since $a_1(q) = 1$, we have

$$\frac{\mathrm{d}}{\mathrm{d}r}a_r(q)\bigg|_{r=1} < 0, \qquad a_1(q) = 1.$$

Hence there exists $r_0 \in (1, 2)$ such that

$$a_{r_0}(q) < a_1(q) = 1.$$

(iv) if

$$\frac{\mathrm{d}}{\mathrm{d}r}\log a_r(q)\bigg|_{r=1} = \left(\frac{1}{a_r(q)}\frac{\mathrm{d}}{\mathrm{d}r}a_r(q)\right)\bigg|_{r=1} = q(\log\phi)'(q) - \log\phi(q) - \log 2 > 0,$$

then since $\log a_1(q) = 0$, there exists $r_1 < 1$ such that

$$\log a_{r_1}(q) < 0 \text{ or } a_{r_1}(q) < 1.$$

□

6. The associated martingale

In this section we study the properties of the process $\{M(q, l), l \geqslant 1\}$ defined in (4.5) for each fixed $q > 0$. We define the increasing family of σ-fields

$$\mathcal{F}_l := \sigma\{W(\mathbf{j}|l), \mathbf{j}|l \in \{0, 1\}^l\}$$

generated by the weights up to and including depth l.

Proposition 6.1. *For each $q > 0$, the family*

$$\{(M(q, l), \mathcal{F}_l), l \geqslant 1\}$$

is a non-negative martingale with constant mean 1 *such that $M(q, l)$ converges almost surely to a limiting random variable $M(q, \infty)$:*

$$M(q, l) \stackrel{\text{a.s.}}{\to} M(q, \infty), \qquad \mathrm{E}(M(q, \infty) \leqslant 1.$$

If the MKP condition fails for q, then

$$P[M(q, \infty) = 0] = 1,$$

and if q satisfies the MKP condition, then $\mathrm{E}(M(q, \infty)) = 1$ *so that*

$$P[M(q, \infty) > 0] = 1.$$

Proof. The martingale property is easily established:

$$\begin{aligned} \mathrm{E}(M(q, l+1)|\mathcal{F}_l) &= \sum_{\mathbf{j}|l, j_{l+1}} \mathrm{E}\left(\frac{\prod_{i=1}^{l} W^q(\mathbf{j}|i)}{c^l} \frac{W^q(\mathbf{j}|l, j_{l+1})}{c} |\mathcal{F}_l\right) \\ &= \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c^l} \sum_{j_{l+1}} \mathrm{E}\left(\frac{W^q(\mathbf{j}|l, j_{l+1})}{c}\right) \\ &= M(q, l) 2 \cdot \frac{\mathrm{E}(W^q)}{c} = M(q, l). \end{aligned}$$

By the martingale convergence theorem (Neveu 1975; Resnick 1998) a non-negative martingale always converges almost surely. The last statements follow by the methods of Kahane and Peyrière (1976). See also Propositions 6.2 and 6.3 below. □

Example 5. Recall the example of the two-point distribution of Example 3 in Section 4. In this case we have $M(q, l) = 1$ for all $q > 0$ and $l \geqslant 1$. For verifying this, the key observation is that

$$W^q + (1 - W)^q = p^q + (1 - p)^q. \tag{6.1}$$

Recall (5.8) and then observe, for $l > 1$, that

$$\begin{aligned} M(q, l) = \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \left(\frac{W^q(\mathbf{j}|i)}{c}\right) &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c}\right) \sum_{j_l} \frac{W^q(\mathbf{j}|l-1, j_l)}{c} \\ &= \sum_{j|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c}\right) \frac{W^q(\mathbf{j}|l-1, 0) + W^q(\mathbf{j}|l-1, 1)}{c}, \end{aligned}$$

and since $W(\mathbf{j}|l-1, 1) = 1 - W(\mathbf{j}|l-1, 0)$ we apply (6.1) to obtain

$$\begin{aligned} &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c}\right) \frac{p^q + (1-p)^q}{c} \\ &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c}\right) = M(q, l-1). \end{aligned}$$

One can easily see that $M(q, 1) = 1$ and the assertion is shown. □

Define $M(q, 0) = 1$ and let the martingale differences be

$$d(q, l) := M(q, l) - M(q, l-1), \qquad l \geqslant 1.$$

For $l > 1$ we have from the definition of $M(q, l)$ that

$$\begin{aligned} d(q, l) &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c} \right) \left[\frac{W^q(\mathbf{j}|l-1, 0) + (1 - W(\mathbf{j}|l-1, 0))^q}{c} - 1 \right] \\ &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c} \right) [\xi(\mathbf{j}|l)]. \end{aligned} \tag{6.2}$$

We can now easily see that $\mathrm{E}(d(q, l)|\mathcal{F}_{l-1}) = 0$. For the conditional variance, note that $\mathrm{E}(\xi(\mathbf{j}|l)) = 0$ and recall expression (4.9),

$$\sigma_1^2(q) = \mathrm{var}(\xi(\mathbf{j}|l) = \frac{1}{c^2} \mathrm{var}(W^q + (1 - W)^q).$$

So the conditional variance of $d(q, l)$ is

$$\begin{aligned} \mathrm{E}(d^2(q, l)|\mathcal{F}_{l-1} &= \mathrm{E}\left(\left(\sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \frac{W^q(\mathbf{j}|i)}{c} \xi(\mathbf{j}|l) \right)^2 |\mathcal{F}_{l-1} \right) \\ &= \sum_{\substack{\mathbf{j}|l-1 \\ \mathbf{p}|l-1}} \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c} \right) \prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{p}|i)}{c} \right) \mathrm{E}(\xi(\mathbf{j}|l) \xi(\mathbf{p}|l)). \end{aligned}$$

Since $\xi(\mathbf{p}|l)) \perp \xi(\mathbf{j}|l))$ if $\mathbf{p}|l \neq \mathbf{j}|l$, we have

$$\begin{aligned} \mathrm{E}(d^2(q, l)|\mathcal{F}_{l-1}) &= \sum_{\mathbf{j}|l-1} \left(\prod_{i=1}^{l-1} \left(\frac{W^q(\mathbf{j}|i)}{c} \right) \right)^2 \sigma_1^2(q) \\ &= \sum_{\mathbf{j}|l-1} \prod_{i=1}^{l-1} \left(\frac{W^{2q}(\mathbf{j}|i)}{c(2q)} \right) a^{l-1}(q) \sigma_1^2(q) \\ &= M(2q, l-1) a^{l-1}(q) \sigma_1^2(q). \end{aligned}$$

Thus the conditional variance of $M(q, l)$ is

$$\sum_{i=1}^{l} \mathrm{E}(d^2(q, i)|\mathcal{F}_{i-1}) = \sum_{i=1}^{l} M(2q, i-1) a^{i-1}(q) \sigma_1^2(q). \tag{6.3}$$

Furthermore,

$$\mathrm{E}(d^2(q, l)) = \mathrm{E}(\mathrm{E}(d^2(q, l)|\mathcal{F}_{l-1})) = \mathrm{E}M(2q, l-1)a^{l-1}(q)\sigma_1^2(q) = a^{l-1}(q)\sigma_1^2(q)$$

and thus

$$\mathrm{var}(M(q, l)) = \sum_{i=1}^{l} \mathrm{E}(d^2(q, i)) = \sigma_1^2(q)\sum_{i=1}^{l} a^{i-1}(q). \tag{6.4}$$

This leads to the following facts.

Proposition 6.2. If $q < q_*$ *so that* $a(q) < 1$, *the martingale* $\{(M(q, l), \mathcal{F}_l), l \geq 0\}$ *is* L_2*-bounded and hence uniformly integrable. It follows that*

$$\mathrm{E}(M(q, \infty) = 1, \qquad M(q, l) = \mathrm{E}(M(q, \infty)|\mathcal{F}_l), \tag{6.5}$$

and $M(q, l) \to M(q, \infty)$ *almost surely and in* L_2. *Moreover, if* $q_* \leq q < q^*$, *then the martingale* $\{(M(q, l), \mathcal{F}_l), l \geq 0\}$ *is* L_p*-bounded for some* $1 < p < 2$ *and hence still uniformly integrable,* (6.5) *still holds and* $M(q, l) \to M(q, \infty)$ *almost surely and in* L_p.

Remark. The proof will show that when $q_* \leq q < q^*$, we may take $p = r_0$, where r_0 is given in Proposition 5.2(iii); see (5.15).

Proof. Suppose first that $q < q_*$. We have from (6.4) that

$$\begin{aligned}\sup_{l\geq 0} \mathrm{E}(M(q, l) - 1)^2 &= \sup_{l\geq 0} \mathrm{var}(M(q, l)) \\ &= \lim_{l\to\infty}\uparrow \sum_{i=1}^{l} \mathrm{E}(d^2(q, i) \\ &= \sum_{i=1}^{\infty} a^{i-1}(q)\sigma_1^2(q) < \infty.\end{aligned}$$

The rest follows from standard martingale theory (see, for example, Neveu 1975, p. 68; Resnick 1998).

Now let $q < q^*$ and we consider uniform integrability without L_2 boundedness. Suppose $1 < p \leq 2$, and for two paths $\mathbf{j}_1$ and $\mathbf{j}_2$ denote by $m_{\mathbf{j}_1,\mathbf{j}_2}$ the largest $i \leq l$ such that $\mathbf{j}_1|i = \mathbf{j}_2|i$. We have

$$
\begin{aligned}
\mathrm{E}(M(q,l))^p &= \frac{1}{c(q)^{lp}}\mathrm{E}\left(\sum_{\mathbf{j}_1|l}\sum_{\mathbf{j}_2|l}\prod_{i=1}^{l}W^q(\mathbf{j}_1|i)W^q(\mathbf{j}_2|i)\right)^{p/2} \\
&\leqslant \frac{1}{c(q)^{lp}}\sum_{k=0}^{l}\mathrm{E}\left(\sum_{\substack{\mathbf{j}_1|l\\ \mathbf{j}_2|l\\ m_{\mathbf{j}_1,\mathbf{j}_2}=k}}\prod_{i=1}^{l}W^q(\mathbf{j}_1|i)W^q(\mathbf{j}_2|i)\right)^{p/2} \\
&\leqslant \frac{1}{c^{lp}(q)}\sum_{k=0}^{l}\mathrm{E}\left(\sum_{\mathbf{j}|k}\prod_{i=1}^{k}W^{2q}(\mathbf{j}|i)\right. \\
&\quad \left.\cdot\sum_{\substack{j^{(1)}_{k+1},\ldots,j^{(1)}_{l}\\ j^{(2)}_{k+1},\ldots,j^{(2)}_{l}\\ j^{(1)}_{k+1}\neq j^{(2)}_{k+1}}}\prod_{i=k+1}^{l}W^q(j_1,\ldots,j_k,j^{(1)}_{k+1},\ldots,j^{(1)}_{l}|i)W^q(j_1,\ldots,j_k,j^{(2)}_{k+1},\ldots,j^{(2)}_{l}|i)\right)^{p/2} \\
&\leqslant \frac{1}{c^{lp}(q)}\sum_{k=0}^{l}\sum_{\mathbf{j}|k}\mathrm{E}(W^{pq})^k \\
&\quad \cdot\mathrm{E}\left(\sum_{\substack{j^{(1)}_{k+1},\ldots,j^{(1)}_{l}\\ j^{(2)}_{k+1},\ldots,j^{(2)}_{l}\\ j^{(1)}_{k+1}\neq j^{(2)}_{k+1}}}\prod_{i=k+1}^{l}W^q(j_1,\ldots,j_k,j^{(1)}_{k+1},\ldots,j^{(1)}_{l}|i)W^q(j_1,\ldots,j_k,j^{(2)}_{k+1},\ldots,j^{(2)}_{l}|i)\right)^{p/2} \\
&\leqslant \sum_{k=0}^{l}\left(\frac{c(pq)}{c^p(q)}\right)^k\frac{1}{c^{(l-k)p}(q)} \\
&\quad \cdot\left(\mathrm{E}\sum_{\substack{j^{(1)}_{k+1},\ldots,j^{(1)}_{l}\\ j^{(2)}_{k+1},\ldots,j^{(2)}_{l}\\ j^{(1)}_{k+1}\neq j^{(2)}_{k+1}}}\prod_{i=k+1}^{l}W^q(j_1,\ldots,j_k,j^{(1)}_{k+1},\ldots,j^{(1)}_{l}|i)W^q(j_1,\ldots,j_k,j^{(2)}_{k+1},\ldots,j^{(2)}_{l}|i)\right)^{p/2} \\
&= \sum_{k=0}^{l}a_p(q)^k\frac{1}{c^{(l-k)p}(q)}\left(\mathrm{E}(W(1-W))^q(\mathrm{E}W^q)^{2(l-k-1)}2^{2(l-k-1)}\cdot 2\right)^{p/2} \\
&= \left((\mathrm{E}(W(1-W))^q)^{p/2}\frac{2^{p/2}}{c(q)}\right)\sum_{k=0}^{l}a_p(q)^k.
\end{aligned}
$$

Here a product over the empty set is equal to 1. By Proposition 5.2(iii) there is a $p \in (1, 2)$ such that $a_p(q) < 1$. For this p the martingale $\{(M(q, l), \mathcal{F}_l), l \geqslant 0\}$ is L_p-bounded, and the rest follows, once again, from standard martingale theory. □

The distribution of $M(q, \infty)$ satisfies a simple recursion which can be used to derive additional information.

Proposition 6.3. *Suppose* $\{M(q, \infty), M_1(q, \infty), M_2(q, \infty)\}$ are *i.i.d. with the same distribution as* $M(q, \infty)$*, the martingale limit. Let* W *have the distribution of the cascade generator and suppose* W *and* $\{M(q, \infty), M_1(q, \infty), M_2(q, \infty)\}$ *are independent. Then*

$$M(q, \infty) \stackrel{d}{=} W^q \frac{M_1(q, \infty)}{c(q)} + (1 - W)^q \frac{M_2(q, \infty)}{c(q)} \tag{6.6}$$

and, for any $q > 0$,

$$P[M(q, \infty) = 0] = 0 \text{ or } 1, \tag{6.7}$$

so that $\mathrm{E}(M(q, \infty)) = 1$ *implies* $P[M(q, \infty) = 0] = 0$.

Proof. We write

$$\begin{aligned} M(q, \infty) &= \lim_{l\to\infty} \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c^l} \\ &= \lim_{l\to\infty} \left(\sum_{j_2,\ldots,j_l} \prod_{i=1}^{l} \frac{W^q(0, j_2, \ldots, j_i)}{c^l} + \sum_{j_2,\ldots,j_l} \prod_{i=1}^{l} \frac{W^q(1, j_2, \ldots, j_i)}{c^l} \right) \\ &= \lim_{l\to\infty} \left(W^q(0) \sum_{j_2,\ldots,j_l} \prod_{i=2}^{l} \frac{W^q(0, j_2, \ldots, j_i)}{c^l} + (1 - W(0))^q \sum_{j_2,\ldots,j_l} \prod_{i=2}^{l} \frac{W^q(1, j_2, \ldots, j_i)}{c^l} \right) \\ &\stackrel{d}{=} W^q(0) \frac{M_1(q, \infty)}{c} + (1 - W(0))^q \frac{M_2(q, \infty)}{c}. \end{aligned}$$

We now verify (6.7). Define

$$p_0 = [M(q, \infty) = 0],$$

$$p_W(0) = P[W = 0] = P[W = 1].$$

Then, since $c(q) \neq 0$,

$$\begin{aligned} p_0 = [M(q, \infty) = 0] &= P[W^q M_1(q, \infty) + (1 - W)^q M_2(q, \infty) = 0] \\ &= P[A, W = 0] + P[A, W = 1] + P[A, 0 < W < 1]. \end{aligned}$$

where $A = W^q M_1(q, \infty) + (1 - W)^q M_2(q, \infty) = 0$. From this we conclude that

$$p_0 = 2p_W(0)p_0 + (1 - 2p_W(0))p_0^2,$$

so that

$$p_0(1 - 2p_W(0)) = p_0^2(1 - 2p_W(0)).$$

If $0 < p_W(0) < \frac{1}{2}$, then $p_0 = p_0^2$ and $p_0 = 0$ or 1. If $p_W(0) = \frac{1}{2}$, then $P[W = 0] = P[W = 1] = \frac{1}{2}$ and W has a two-point distribution, and hence from Example 5 we know $M(q, l) = 1$ which implies $M(q, \infty) = 1$. □

7. Estimation: subcritical consistency

We propose two estimators of the structure function which depend on scaled summed powers of the wavelet coefficients $\{Z(q, l), l \geqslant 1\}$. These are

$$\hat{\tau}_1(q) = \hat{\tau}_1(q, l) = \frac{\log_2 Z(q, l)}{l} = \frac{\log_2 \sum_{n=0}^{2^l-1} |d_{-l,n}|^q - ql/2}{l}, \tag{7.1}$$

$$\hat{\tau}_2(q) = \hat{\tau}_2(q, l) = \log_2\left(\frac{Z(q, l+1)}{Z(q, l)}\right) = \log_2\left(\frac{\sum_{n=0}^{2^{l+1}-1} |d_{-(l+1),n}|^q}{2^{q/2}\sum_{n=0}^{2^l-1} |d_{-l,n}|^q}\right). \tag{7.2}$$

Analysis depends on showing that scaled versions of $Z(q, l)$ are well approximated by the martingale, and this is discussed next. Recall notational definitions (4.1), (4.2), (4.3) and (4.6).

Proposition 7.1. *For $q > 0$,*

$$\frac{Z(q, l)}{c^l b} - M(q, l) \xrightarrow{P} 0.$$

If $q \neq q^*$ the convergence is almost sure, and if $q < q_*$ the convergence is in L_2. Thus

$$\frac{Z(q, l)}{c^l b} \to M(q, \infty) \tag{7.3}$$

in the appropriate sense, depending on the case.

Proof. Begin by writing

$$\frac{Z(q, l)}{c^l b} - M(q, l) = \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c}\left[\frac{|2W(\mathbf{j}|l, 0) - 1|^q}{b} - 1\right]$$

$$= \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c}\xi(\mathbf{j}|l, 0), \tag{7.4}$$

where $\xi(\mathbf{j}|l, 0) \perp \xi(\mathbf{p}|l, 0)$ if $\mathbf{j}|l \neq \mathbf{p}|l$. Also $\mathrm{E}\xi(\mathbf{j}|l, 0) = 0$ and recall expression (4.10),

$$\sigma_2^2(q) := \mathrm{E}\xi^2(\mathbf{j}|l, 0) = \frac{1}{b^2}\,\mathrm{var}(|2W-1|^q).$$

If $q < q_*$, so $a(q) < 1$, then, similarly to the calculations leading to (6.3) and (6.4), we find

$$\begin{aligned}\mathrm{E}\left(\frac{Z(q, l)}{c^l b} - M(q, l)\right)^2 &= \sum_{\mathbf{j}|l}\sum_{\mathbf{p}|l}\mathrm{E}\left(\prod_{i=1}^{l}\frac{W^q(\mathbf{j}|i)}{c}\prod_{i=1}^{l}\frac{W^q(\mathbf{p}|i)}{c}\xi(\mathbf{j}|l, 0)\xi(\mathbf{p}|l, 0)\right)\\ &= \sigma_2^2(q)\frac{(2\mathrm{E}W^{2q})^l}{c^{2l}(q)} = \sigma_2^2(q)a^l(q)\\ &\to 0\end{aligned}$$

as $l\to\infty$ since $a(q) < 1$. This shows the L_2 convergence.

For $q > 0$, the same method shows

$$\begin{aligned}\mathrm{E}\left(\left(\frac{Z(q, l)}{c^l b} - M(q, l)\right)^2 |\mathcal{F}_l\right) &= \sigma_2^2(q)M(2q, l)a^l(q)\\ &= \sigma_2^2(q)\sum_{\mathbf{j}|l}\prod_{i=1}^{l}\frac{W^{2q}(\mathbf{j}|i)}{c^2(q)} =: \sigma_2^2(q)V(q, l), \qquad (7.5)\end{aligned}$$

and we need to show $V(q, l)\to 0$ almost surely as $l\to\infty$. If the MKP condition fails, then $M(q, l)\to 0$ as $l\to\infty$ and

$$V(q, l) \leqslant (M(q, l))^2 \to 0.$$

If the MKP condition holds, then from Proposition 5.2(iii) there exists $r_0\in(1, 2)$ such that $a_{r_0}(q) < 1$, and for $p = r_0/2 \in (1/2, 1)$ we have by the triangle inequality

$$0 \leqslant V(q, l)^p \leqslant \sum_{\mathbf{j}|l}\prod_{i=1}^{l}\frac{W^{2pq}(\mathbf{j}|i)}{c^{2p}(q)} = (a_{r_0}(q))^l M(r_0 q, l) \overset{\text{a.s.}}{\to} 0, \qquad (7.6)$$

as $l\to\infty$, since $M(r_0 q, \infty) < \infty$ almost surely.

So in all cases $V(q, l)\to 0$. For any $\delta > 0$, $\epsilon > 0$ we have

$$P\left[\left|\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right| > \epsilon|\mathcal{F}_l\right] = P\left[\left|\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right| > \epsilon|\mathcal{F}_l\right] 1_{[V(q,l)\sigma_2^2(q)>\delta]}$$

$$+ P\left[\left|\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right| > \epsilon|\mathcal{F}_l\right] 1_{[V(q,l)\sigma_2^2(q)\leqslant\delta]}$$

$$\leqslant 1_{[V(q,l)\sigma_2^2(q)>\delta]} + \epsilon^{-2}\mathrm{E}\left(\left(\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right)^2 |\mathcal{F}_l\right)$$

$$\times 1_{[V(q,l)\sigma_2^2(q)\leqslant\delta]}$$

$$= 1_{[V(q,l)\sigma_2^2(q)>\delta]} + \epsilon^{-2} V(q,\,l)\sigma_2^2(q) 1_{[V(q,l)\sigma_2^2(q)\leqslant\delta]}$$

$$\leqslant 1_{[V(q,l)\sigma_2^2(q)>\delta]} + \frac{\delta^2}{\epsilon^2}.$$

Take expectations and use $V(q,\,l) \overset{\text{a.s.}}{\to} 0$ and the arbitrariness of δ to conclude that

$$\frac{Z(q,\,l)}{c^l b} - M(q,\,l) \overset{P}{\to} 0,$$

as $l \to \infty$.

For almost sure convergence, when $q < q^*$, we obtain from (7.6) that

$$V(q,\,l) \leqslant (a_{r_0}(q)^{1/p})^l M^{1/p}(r_0 q,\,l)$$

and so $\sum_l V(q,\,l) < \infty$ almost surely. Thus, for any $\epsilon > 0$,

$$\sum_l P\left[\left|\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right| > \epsilon|\mathcal{F}_l\right] \leqslant \epsilon^{-2} \sum \mathrm{E}\left(\left(\frac{Z(q,\,l)}{c^l b} - M(q,\,l)\right)^2 |\mathcal{F}_l\right)$$

$$= (\text{const.}) \sum_l V(q,\,l) < \infty,$$

and by a generalization of the Borel–Cantelli lemma (Neveu 1975, p. 152) we have

$$\frac{Z(q,\,l)}{c^l b} - M(q,\,l) \overset{\text{a.s.}}{\to} 0.$$

For $q > q^*$, we prove almost sure convergence from Proposition 5.2(iv) in a similar way. □

We use the comparison result in Proposition 7.1 to obtain consistent estimators of the structure function $\tau(q)$ in the *subcritical* case, by which we mean the case where the MKP condition holds.

Proposition 7.2. *Define* $\hat{\tau}_i(q)$ *for* $i = 1, 2$ *by* (7.1) *and* (7.2). *Provided* $q < q^*$, *so that the MKP condition holds, both estimators are almost surely consistent for* $\tau(q)$*:*

$$\hat{\tau}_i(q) \stackrel{a.s.}{\to} \tau(q), \qquad i = 1, 2,$$

as $l \to \infty$.

Proof. In (7.3), take logarithms to the base 2 to obtain

$$\log_2 Z(q, l) - l \log_2 c(q) - \log_2 b \to \log_2 M(q, \infty) \tag{7.7}$$

almost surely as $l \to \infty$. Divide through by l to get consistency of $\hat{\tau}_1(q)$. Note from (7.7) that

$$\log_2 Z(q, l+1) - \log_2 Z(q, l) - (l + 1 - l)\tau(q) \to 0$$

almost surely, which proves consistency of $\hat{\tau}_2(q)$. □

8. Subcritical asymptotic normality of estimators

In this section we discuss second-order properties of the estimators $\hat{\tau}_i(q)$, $i = 1, 2$, defined in (7.1) and (7.2). The asymptotic normality for $\hat{\tau}_1(q)$ requires a bias term which cannot be eliminated. This drawback is overcome by using $\hat{\tau}_2(q)$, whose definition in terms of differencing removes the bias term. However, take note of the suggestive remarks at the end of this section about mean square error.

For this section it is convenient to write $\mathrm{E}^{\mathcal{F}_l}$ and $P^{\mathcal{F}_l}$ for the conditional expectation and conditional probability with respect to the σ-field $\mathcal{F}_l$.

We first consider the asymptotic normality of $\hat{\tau}_1(q)$. Begin by writing

$$\frac{Z(q, l)}{c^l b} - M(q, l) = \sum_{\mathbf{j}|l} \prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c} \left[\frac{|2W(\mathbf{j}|l, 0) - 1|^q}{b} - 1 \right] \tag{8.1}$$

$$=: \sum_{\mathbf{j}|l} Z(\mathbf{j}|l), \tag{8.2}$$

where

$$\mathrm{E}^{\mathcal{F}_l}(Z(\mathbf{j}|l)) = 0$$

$$\mathrm{E}^{\mathcal{F}_l}(Z(\mathbf{j}|l))^2 = \left(\prod_{i=1}^{l} \frac{W^{2q}(\mathbf{j}|i)}{c(2q)} \right) a^l(q)\sigma_2^2(q),$$

and recall that $\sigma_2^2(q)$ is defined in (4.10). Therefore,

$$\sum_{\mathbf{j}|l} \mathrm{E}^{\mathcal{F}_l}(Z(\mathbf{j}|l))^2 = M(2q, l)a^l(q)\sigma_2^2(q). \tag{8.3}$$

Our strategy for the central limit theorem is to regard $Z(q, l)/c^l b - M(q, l)$ as a sum of random variables which are conditionally independent given $\mathcal{F}_l$ and then apply the Lyapunov condition (Resnick 1998) for asymptotic normality in a triangular array.

Proposition 8.1. *If $2q < q^*$, then, as $l \to \infty$,*

$$P^{\mathcal{F}_l}\left[\frac{Z(q,\,l)/c^l b - M(q,\,l)}{\sqrt{M(2q,\,l)a^l(q)\sigma_2^2(q)}} \leqslant x\right] \to P[N(0,\,1) \leqslant x] \quad a.s., \tag{8.4}$$

where $N(0, 1)$ is a standard normal random variable. Taking expectations in (8.4) *yields*

$$P\left[\frac{Z(q,\,l)/c^l b - M(q,\,l)}{\sqrt{M(2q,\,l)a^l(q)\sigma_2^2(q)}} \leqslant x\right] \to P[N(0,\,1) \leqslant x]. \tag{8.5}$$

Proof. By Proposition 5.2(iii) there exists $\delta > 0$ such that both $2q + \delta < q^*$ and

$$a_{1+\delta/2}(2q) < 1. \tag{8.6}$$

Asymptotic normality in (8.4) will be shown if we establish the Lyapunov condition

$$\frac{\sum_{\mathbf{j}|l} \mathrm{E}^{\mathcal{F}_l}|Z(\mathbf{j}|l)|^{2+\delta}}{(M(2q,\,l)a^l(q))^{(2+\delta)/2}} \to 0 \quad \text{a.s.}, \tag{8.7}$$

where the denominator comes from (8.3). The numerator on the left-hand side of (8.7) is bounded above by

$$\begin{aligned}
&\mathrm{E}^{\mathcal{F}_l}\sum_{\mathbf{j}|l}\left|\prod_{i=1}^{l}\frac{W^q(\mathbf{j}|i)}{c}\right|^{2+\delta}\left|\frac{|2W(\mathbf{j}|l,\,0)-1|^q}{b}-1\right|^{2+\delta}\\
&\quad= c_1\sum_{\mathbf{j}|l}\left(\prod_{i=1}^{l}\frac{W^{q(2+\delta)}(\mathbf{j}|i)}{c((2+\delta)q)}\right)\frac{c^l((2+\delta)q)}{c^{l(2+\delta)}(q)}\\
&\quad= c_1 M((2+\delta)q,\,l)(a_{2+\delta}(q))^l,
\end{aligned}$$

where

$$c_1 = \mathrm{E}\left|\frac{|2W(\mathbf{j}|l,\,0)-1|^q}{b}-1\right|^{2+\delta}.$$

So the ratio in (8.7), apart from constants, is bounded by

$$\frac{M((2+\delta)q,\,l)(a_{2+\delta}(q))^l}{M(2q,\,l)^{1+\delta/2}(a_2(q))^{(1+\delta/2)l}} \sim \frac{M((2+\delta)q,\,\infty)(a_{2+\delta}(q))^l}{M(2q,\,\infty)^{1+\delta/2}(a_2(q))^{(1+\delta/2)l}}.$$

Note that the two random variables $M((2+\delta)q, \infty)$ and $M(2q, \infty)$ are non-zero with probability 1 by Proposition 6.1. Check that

$$\frac{a_{2+\delta}(q)}{(a_2(q))^{1+\delta/2}} = a_{1+\delta/2}(2q) < 1.$$

So the Lyapunov ratio is asymptotic to a finite non-zero random variable times $(a_{1+\delta/2}(2q))^l$, where $a_{1+\delta/2}(2q) < 1$, and the result is proven. □

Remark 8.1. In the denominator of (8.5) we may replace $M(2q, l)$ by its limit $M(2q, \infty)$. This follows since almost surely $0 < M(2q, \infty) < \infty$ for $2q < q^*$ and thus

$$\left(\frac{Z(q, l)/c^l b - M(q, l)}{\sqrt{M(2q, l)a^l(q)\sigma_2^2(q)}}, \sqrt{\frac{M(2q, l)}{M(2q, \infty)}} \right) \Rightarrow (N(0, 1), 1)$$

by Billingsley (1968). The desired result is obtained by multiplying components.

Remark 8.2. Set

$$N_l := \frac{Z(q, l)/c^l b - M(q, l)}{\sqrt{M(2q, l)a^l(q)\sigma_2^2(q)}}. \tag{8.8}$$

Then in $\mathbb{R}^2$, as $l \to \infty$,

$$(N_l, N_{l+1}) \Rightarrow (N_1(0, 1), N_2(0, 1)),$$

where $N_i(0, 1)$, $i = 1, 2$, are i.d.d. standard normal random variables.

To see this, write, for any $x, y \in \mathbb{R}$,

$$\begin{aligned} P[N_l \leqslant x, N_{l+1} \leqslant y] &= \mathrm{E}P^{\mathcal{F}_{l+1}}[N_l \leqslant x, N_{l+1} \leqslant y] \\ &= \mathrm{E}1_{[N_l \leqslant x]} P^{\mathcal{F}_{l+1}}[N_{l+1} \leqslant y]. \end{aligned}$$

By Proposition 8.1,

$$P^{\mathcal{F}_{l+1}}[N_{l+1} \leqslant y] = \Phi(y) + \epsilon_l(y) \quad \text{a.s.},$$

where $\Phi(y)$ is the standard normal cdf and where $\epsilon_l(y) \xrightarrow{L_1} 0$ and $|\epsilon_l(y)| \leqslant 2$. So

$$\begin{aligned} P[N_l \leqslant x, N_{l+1} \leqslant y] &= \mathrm{E}1_{[N_l \leqslant x]}(\Phi(y) + \epsilon_l(y)) \\ &= \mathrm{E}1_{[N_l \leqslant x]}\Phi(y) + o(1) \end{aligned}$$

from the dominated convergence theorem, and hence we obtain

$$P[N_l \leqslant x, N_{l+1} \leqslant y] \to \Phi(x)\Phi(y).$$

We now describe how this central limit behaviour transfers to $\hat{\tau}_1(q)$.

Corollary 8.1. *Under the assumptions in force in Proposition* 8.1, *we have*

$$\frac{(\hat{\tau}_1(q) - \tau(q)) - l^{-1}\log_2 bM(q, l)}{\sqrt{M(2q, \infty)a^l(q)\sigma_2^2(q)}/(l \log 2 \cdot M(q, l))} \Rightarrow N(0, 1). \tag{8.9}$$

Remark 8.3. The bias term $l^{-1}\log_2(bM(q, l))$ cannot be neglected.

Proof. For brevity, write

$$d(q) := M(2q, \infty)a^l(q)\sigma_2^2(q), \tag{8.10}$$

and using the notation of (8.8) we have

$$Z(q, l) = c^l b\Big(N_l\sqrt{d(q)} + M(q, l)\Big).$$

Since

$$\hat{\tau}_1(q) = \frac{1}{l}\log_2 Z(q, l),$$

we have

$$l\hat{\tau}_1(q) = l\log_2 c + \log_2 b + \log_2\Big(N_l\sqrt{d(q)} + M(q, l)\Big)$$

and thus

$$l(\hat{\tau}_1(q) - \tau(q)) = \log_2 bM(q, l) + \log_2\left(1 + \frac{N_l\sqrt{d(q)}}{M(q, l)}\right).$$

Since by (5.6) and assumption $2q < q^*$ we have $q < q_*$, we know that $d(q) \to 0$. Therefore, $N_l\sqrt{d(q)}/M(q, l) \to 0$, and the desired result follows by using the relation $\log(1 + x) \sim x$ for $x\downarrow 0$. □

The bias term in (8.9) is an unpleasant feature and thus we consider how to remove it by differencing. Consider the asymptotic normality of $\hat{\tau}_2(q)$. It is possible to proceed from Proposition 8.1, but it is simpler to proceed with a direct proof.

Proposition 8.2. *Suppose* $2q < q^*$. *Then*

$$\frac{\hat{\tau}_2(q) - \tau(q)}{\sqrt{M(2q, \infty)a^l(q)\sigma_3^2(q)}/(\log 2 \cdot M(q, \infty))} \Rightarrow N(0, 1), \tag{8.11}$$

where $\sigma_3^2(q)$ *is defined in* (4.11).

Proof. Begin by observing that

$$\begin{aligned}\frac{Z(q, l+1)}{c^{l+1}b} - \frac{Z(q, l)}{c^l b} &= \sum_{\mathbf{j}|l}\left(\prod_{i=1}^{l}\frac{W^q(\mathbf{j}|i)}{c}\right)\left[\frac{W^q(\mathbf{j}|l, 0)}{c}\frac{|2W(\mathbf{j}|l, 0, 0) - 1|^q}{b} + \frac{W^q(\mathbf{j}|l, 1)}{c}\frac{|2W(\mathbf{j}|l, 1, 0) - 1|^q}{b}\right.\\ &\qquad\left. - \frac{|2W(\mathbf{j}|l, 0) - 1|^q}{b}\right]\\ &=: \sum_{\mathbf{j}|l}\left(\prod_{i=1}^{l}\frac{W^q(\mathbf{j}|i)}{c}\right)H(\mathbf{j}|l),\end{aligned}$$

where we have set

$$H(\mathbf{j}|l) = \frac{W^q(\mathbf{j}|l, 0)}{c}\frac{|2W(\mathbf{j}|l, 0, 0) - 1|^q}{b} + \frac{W^q(\mathbf{j}|l, 1)}{c}\frac{|2W(\mathbf{j}|l, 1, 0) - 1|^q}{b} - \frac{|2W(\mathbf{j}|l, 0) - 1|^q}{b}. \tag{8.12}$$

We have that

$$\mathrm{E}^{\mathcal{F}_l} H(\mathbf{j}|l) = \frac{1}{2} + \frac{1}{2} - 1 = 0$$

and

$$\mathrm{E}^{\mathcal{F}_l} H^2(\mathbf{j}|l) = \sigma_3^2(q).$$

It follows that, conditionally on $\mathcal{F}_l$, we may treat

$$\frac{Z(q, l+1)}{c^{l+1}b} - \frac{Z(q, l)}{c^l b}$$

as a sum of i.i.d. random variables with (conditional) variance

$$\begin{aligned}
\mathrm{E}^{\mathcal{F}_l}\left(\frac{Z(q, l+1)}{c^{l+1}b} - \frac{Z(q, l)}{c^l b}\right)^2 &= \sum_{\mathbf{j}|l} \mathrm{E}^{\mathcal{F}_l}\left(\left(\prod_{i=1}^{l} \frac{W^q(\mathbf{j}|i)}{c(q)}\right) H(\mathbf{j}|l)\right)^2 \\
&= \sum_{\mathbf{j}|l}\left(\prod_{i=1}^{l} \frac{W^{2q}(\mathbf{j}|i)}{c(2q)}\right) a^l(q) EH^2(\mathbf{j}|l) \\
&= M(2q, l) a^l(q)\sigma_3^2(q).
\end{aligned}$$

As in the proof of Proposition 8.1, we may check the Lyapunov condition and conclude that

$$\frac{Z(q, l+1)/c^{l+1}b - Z(q, l)/c^l b}{\sqrt{M(2q, \infty)a^l(q)\sigma_3^2(q)}} \Rightarrow N(0, 1),$$

or equivalently

$$\frac{(c^{-1}Z(q, l+1)/Z(q, l) - 1)Z(q, l)/c^l b}{\sqrt{M(2q, \infty)a^l(q)\sigma_3^2(q)}} \Rightarrow N(0, 1),$$

and since $Z(q, l)/c^l b) \to M(q, \infty)$ we have

$$\frac{(c^{-1}Z(q, l+1)/Z(q, l) - 1)M(q, \infty)}{\sqrt{M(2q, \infty)a^l(q)\sigma_3^2(q)}} \Rightarrow N(0, 1). \tag{8.13}$$

Since

$$c^{-1}\frac{Z(q, l+1)}{Z(q, l)} - 1 \xrightarrow{P} 0,$$

it follows that

$$\hat{\tau}_2(q) - \tau(q) = \log_2\left(c^{-1}\frac{Z(q,\, l+1)}{Z(q,\, l)}\right)$$

$$= \frac{\log(1 + (c^{-1}Z(q,\, l+1)/Z(q,\, l) - 1))}{\log 2}$$

$$\sim \frac{c^{-1}Z(q,\, l+1)/Z(q,\, l) - 1}{\log 2}$$

in probability. Combine this with (8.13) to complete the proof. □

For statistical purposes, the result (8.11) contains unobservables so, as in Troutman and Vecchia (1999) and Ossiander and Waymire (2000), consideration needs to be given to replacing quantities which are not observed by observable estimators. We assume that the random measure μ_∞ is observed, or equivalently that the wavelet coefficients $\{d_{-l,n}\}$ are known. This means we have the quantities $\{Z(q,\, l)\}$.

Define the following observable quantity

$$D^2(q,\, l) = \sum_{\mathbf{j}|l}\prod_{i=1}^{l} W^{2q}(\mathbf{j}|i)\left[\frac{W^q(\mathbf{j}|l,\, 0)|2W(\mathbf{j}|l,\, 0,\, 0) - 1|^q}{Z(q,\, l+1)} + \frac{W^q(\mathbf{j}|l,\, 1)|2W(\mathbf{j}|l,\, 1,\, 0) - 1|^q}{Z(q,\, l+1)} - \frac{|2W(\mathbf{j}|l,\, 0) - 1|^q}{Z(q,\, l)}\right]^2$$

$$=: \sum_{\mathbf{j}|l}\prod_{i=1}^{l} W^{2q}(\mathbf{j}|i)V^2(\mathbf{j}|l)$$

$$=: \sum_{\mathbf{j}|l}\prod_{i=1}^{l} W^{2q}(\mathbf{j}|i)\left[\frac{A+B}{Z(q,\, l+1)} - \frac{C}{Z(q,\, l)}\right]^2. \tag{8.14}$$

Note that in this notation,

$$H(\mathbf{j}|l) = \frac{A+B}{cb} - \frac{C}{b},$$

where $H(\mathbf{j}|l)$ is defined in (8.12). Recall also that $\mathrm{E}H^2(\mathbf{j}|l) = \sigma_3^2(q)$. In terms of the wavelet coefficients, we have

$$D^2(q,\, l) = \sum_{\mathbf{j}|l}\left[\frac{|d(-l,\, (\mathbf{j}|l,\, 0))|^q 2^{-q(l+1)/2}}{Z(q,\, l+1)} + \frac{|d(-l,\, (\mathbf{j}|l,\, 1))|^q 2^{-q(l+1)/2}}{Z(q,\, l+1)} - \frac{|d(-l,\, (\mathbf{j}|l)|^q 2^{-ql/2}}{Z(q,\, l)}\right]^2, \tag{8.15}$$

showing that $D^2(q,\, l)$ is an observable statistic.

Corollary 8.2. *Suppose that* $2q < q^*$. *Then*

$$\frac{\hat{\tau}_2(q) - \tau(q)}{D(q, l)/\log 2} \Rightarrow N(0, 1) \tag{8.16}$$

as $l \to \infty$.

Proof. Because of (8.11), it suffices to show

$$\frac{D^2(q, l)}{M(2q, \infty)a^l(q)\sigma_3^2(q)/M^2(q, \infty)} \xrightarrow{P} 1,$$

as $l \to \infty$. This is equivalent to showing

$$\frac{[Z^2(q, l)/c^{2l}(q)b^2(q)]D^2(q, l)}{M(2q, l)a^l(q)\sigma_3^2(q)} \xrightarrow{P} 1.$$

After some simple algebra, this ratio is the same as

$$\frac{\sum_{\mathbf{j}|l}\prod_{i=1}^{l}[W^{2q}(\mathbf{j}|i)/c(2q)]}{M(2q, l)\sigma_3^2(q)}\left[\frac{A+B}{bc}\left(\frac{Z(q, l)/c^l b}{Z(q, l+1)/c^{l+1}b}\right) - \frac{C}{b}\right]^2.$$

Since $M(2q, l) \to M(2q, \infty)$, it suffices to show that the numerator converges in probability to $M(2q, \infty)\sigma_3^2(q)$. Due to (7.3), we write the numerator as

$$\sum_{\mathbf{j}|l}\prod_{i=1}^{l}\frac{W^{2q}(\mathbf{j}|i)}{c(2q)}\left(\left[\frac{A+B}{bc} - \frac{C}{b}\right]^2 + o_p(1)2\left(\frac{A+B}{bc} - \frac{C}{b}\right) + o_p(1)^2\left(\frac{A+B}{bc}\right)^2\right)$$

$$= I + II + III.$$

As in Theorem 3.5 of Ossiander and Waymire (2000),

$$I \to M(2q, \infty)\mathrm{E}\left(\frac{A+B}{bc} - \frac{C}{b}\right)^2 = M(2q, \infty)\sigma_3^2(q),$$

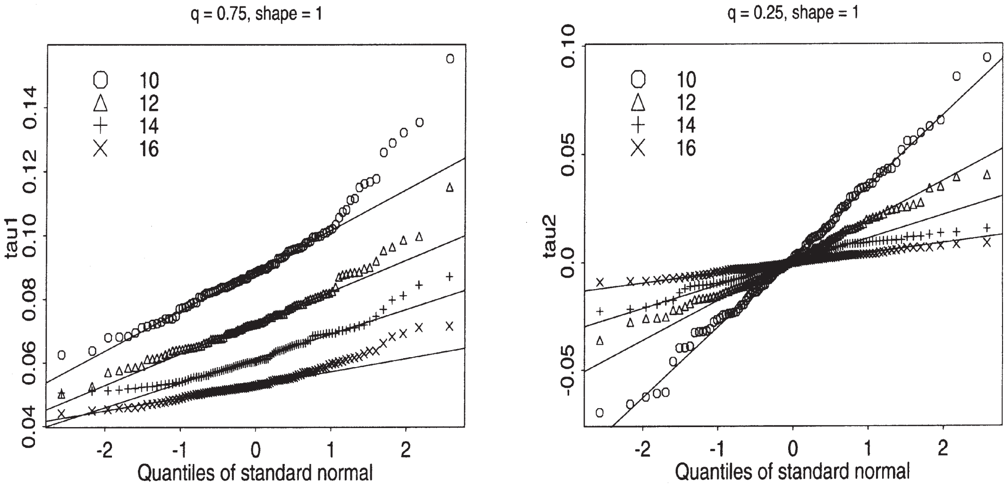
as desired. The terms *II* and *III* can readily be shown to go to 0. □

Figure 1 shows normal quantile–quantile plots of $\hat{\tau}_i(q) - \tau(q)$, $i = 1, 2$, from simulated cascade data with beta-distributed cascade generator with shape parameter 1 (this makes the distribution uniform). In Figure 1(a) $q = 0.75$, and in Figure 1(b) $q = 0.25$. Each plot presents four graphs as the depth l increases to 16. Note the better agreement of $\hat{\tau}_2(q)$ to normality compared with $\hat{\tau}_1(q)$.

We conclude this section with a remark on mean square error. Examining Corollary 8.1 and Proposition 8.2 yields that, in the region $2q < q^*$, the conditional mean square error of $\hat{\tau}_1(q)$ is of the form

$$\frac{O_p^{(1)}(a^l(q))}{l^2} + \frac{O_p^{(2)}(1)}{l^2}$$

while that of $\hat{\tau}_2(q)$ is $O_p^{(3)}(a^l(q))$.



(a) (b)

Figure 1. Normal quantile–quantile plots of (a) $\hat{\tau}_1(q) - \tau(q)$ and (b) $\hat{\tau}_2(q) - \tau(q)$

9. Supercritical asymptotics: lack of consistency

A critical issue with both the wavelet-based estimator and the moment-based ones used in Ossiander and Waymire (2000) is that the asymptotic properties of the estimators are only valid in a certain range of q-values. For the wavelet estimators we require $q < q^*$ for consistency, and for the asymptotic normality results we require $2q < q^*$. We now show that the range $q > q^*$ is uninformative for our estimators, and in fact our estimators are misleading when extended to inference for values beyond q^*. A reliable estimate of q^* would be valuable information. In place of such an estimate it is likely that a graphical procedure is possible based on the following.

Let $\hat{\tau}_i^{\vee}(q)$ $(i = 1, 2)$ have the same definition as $\hat{\tau}_i(q)$ except that sum is replaced by max. Thus we can define, by analogy with (4.3),

$$Z^{\vee}(q, l) = \bigvee_{\mathbf{j}|l} \prod_{i=1}^{l} W(\mathbf{j}|i)^q |2W(\mathbf{j}|l, 0) - 1|^q.$$

Note that

$$Z^{\vee}(q, l) = (Z^{\vee}(1, l))^q.$$

For large values of q, namely for $q \geqslant q^*$, $Z(q, l)$ is sufficiently well approximated by its largest summand $Z^{\vee}(q, l)$. Figure 2 presents a density plot of simulated values of $Z(q, l)/Z^{\vee}(q, l)$ as the depth l increases from 10 to 18; note that the densities concentrate most mass around the point 1. The cascade generator is a beta distribution with shape

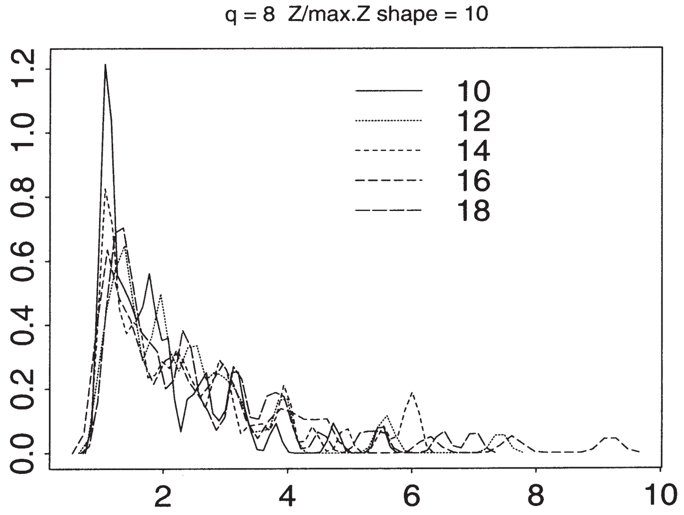


Figure 2. Density plots of $Z(q, l)/Z^{\vee}(q, l)$, $l = 10, 12, \ldots, 18$

parameter 10. Based on the idea of approximating $Z(q, l)$ by $Z^{\vee}(q, l)$, since $\log Z^{\vee}(q, l) = q \log Z^{\vee}(1, l)$ is linear in q, we anticipate that $\hat{\tau}_1(q)$ should also be linear in q, rendering $\hat{\tau}_1(q)$ largely uninformative for inference purposes in the $q \geqslant q^*$ region. A rough estimate of q^* would be provided by the q-value where the plots of $\hat{\tau}_1(q)$ start to look linear.

Computer simulations offer strong support for these remarks. Figure 3 shows overlaid simulated values for $\hat{\tau}_i(q)$, $\hat{\tau}_i^{\vee}(q)$, $i = 1, 2$, for large values of q. In the range of q-values beyond $q^* \approx 3.3$, it is remarkable how linear the plots for $\hat{\tau}_1(q)$ and $\hat{\tau}_1^{\vee}(q)$ look and also how closely $\hat{\tau}_1^{\vee}(q)$ approximates $\hat{\tau}_1(q)$. Note that the values in the plots have been multiplied by -1 to make the plots increasing and that the cascade generator is a beta distribution with shape parameter 1.

We now assume that $q^* < \infty$ and examine this supercritical phenomenon when $q \geqslant q^*$ in more detail. We will prove the asymptotic linearity of the estimator $\hat{\tau}_1(q)$ for $q \geqslant q^*$. In particular, the estimator $\hat{\tau}_1(q)$ is not consistent when $q > q^*$, and neither is the estimator $\hat{\tau}_2(q)$.

We start by introducing new notation. Let

$$U(q, l) = c(q)^l M(q, l) = \sum_{\mathbf{j}|l} \prod_{i=1}^{l} W(\mathbf{j}|i)^q, \qquad q > 0,\ l \geqslant 1, \tag{9.1}$$

$$U^*(l) = \max_{\mathbf{j}|l} \prod_{i=1}^{l} W(\mathbf{j}|i), \qquad l \geqslant 1, \tag{9.2}$$

and define, for $q > 0$,

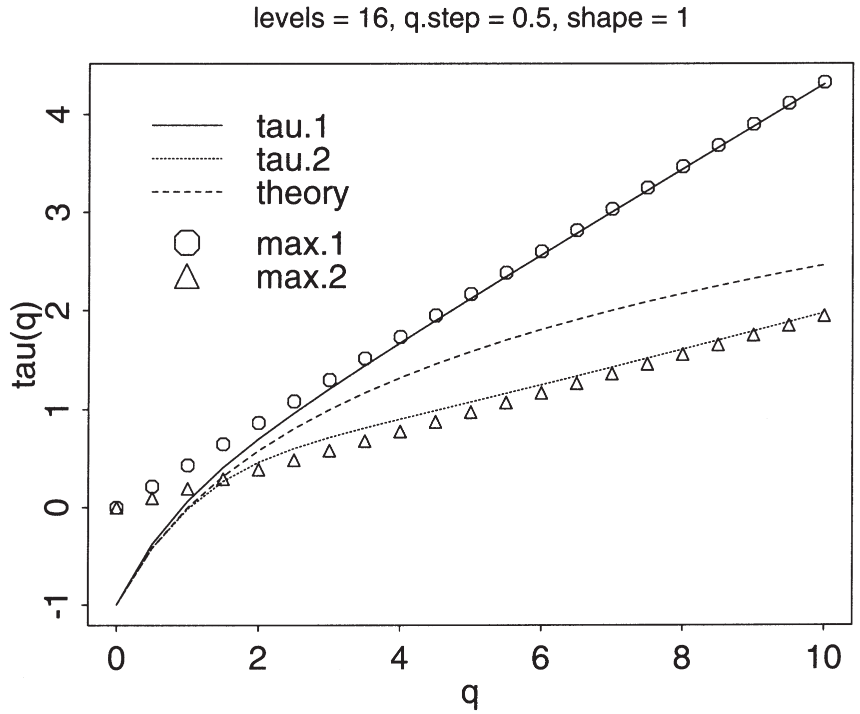


Figure 3. Plots of $\hat{\tau}_i(q)$, $\hat{\tau}_i^{\vee}(q)$ for $q_* \approx 2.4$

$$\overline{m}(q) = \limsup_{l\to\infty} \frac{1}{l} \log_2 U(q, l), \qquad \underline{m}(q) = \liminf_{l\to\infty} \frac{1}{l} \log_2 U(q, l), \tag{9.3}$$

as well as

$$\overline{m}^* = \limsup_{l\to\infty} \frac{1}{l} \log_2 U^*(l), \qquad \underline{m}^* = \liminf_{l\to\infty} \frac{1}{l} \log_2 U^*(l). \tag{9.4}$$

It is immediate that, for all $q > 0$ and $0 \leqslant \theta \leqslant q$,

$$(U^*(l))^q \leqslant U(q, l) \leqslant (U^*(l))^\theta U(q - \theta, l) \leqslant 2^l (U^*(l))^q. \tag{9.5}$$

In particular, for every $q > 0$,

$$\overline{m}(q) - 1 \leqslant q\overline{m}^* \leqslant \overline{m}(q), \qquad \underline{m}(q) - 1 \leqslant q\underline{m}^* \leqslant \overline{m}(q) \tag{9.6}$$

almost surely.

Note that it follows from Proposition 6.1 that, for $0 < q < q^*$,

$$\overline{m}(q) = \underline{m}(q) = \tau(q). \tag{9.7}$$

Since by the triangle inequality, for all $0 < \rho < 1$ and $q > 0$,

$$(U(q, l))^{\rho} < U(\rho q, l), \tag{9.8}$$

we see that

$$\overline{m}(\rho q) \geqslant \rho \overline{m}(q). \tag{9.9}$$

For a $q \geqslant q^*$ and $0 < \rho < q^*/q$ we hence get

$$\overline{m}(q) \leqslant \frac{1}{\rho} \overline{m}(\rho q) = \frac{1}{\rho} \tau(\rho q),$$

and letting $\rho \uparrow q^*/q$ we conclude that, for every $q \geqslant q^*$,

$$\overline{m}(q) \leqslant q \frac{\tau(q^*)}{q^*}. \tag{9.10}$$

On the other hand, it follows from (9.5) that, for all $q > 0$ and $0 \leqslant \theta \leqslant q$,

$$\underline{m}(q) \leqslant \theta \underline{m}^* + \overline{m}(q - \theta).$$

Using (9.6), we obtain

$$\underline{m}(q_1) \leqslant \theta \frac{\overline{m}(q_2)}{q_2} + \overline{m}(q_1 - \theta)$$

for all $q_1, q_2 > 0$ and $0 \leqslant \theta \leqslant q_1$. In particular, if $0 < q_1 < q^*$, then for every $0 < q_3 < q_1$ we choose $\theta = q_1 - q_3$ and conclude, using (9.7), that

$$\frac{\underline{m}(q_2)}{q_2} \geqslant \frac{\underline{m}(q_1) - \overline{m}(q_3)}{q_1 - q_3} = \frac{\tau(q_1) - \tau(q_3)}{q_1 - q_3}.$$

Therefore, for all $q > 0$,

$$\frac{\underline{m}(q)}{q} \geqslant \sup_{0 < p < q^*} \tau'(p). \tag{9.11}$$

However,

$$\sup_{0 < p < q^*} \tau'(p) = \sup_{0 < p < q^*} \frac{\mathrm{E}(W^p \log_2 W)}{\mathrm{E}(W^p)} \geqslant \frac{\mathrm{E}(W^{q^*} \log_2 W)}{\mathrm{E}(W^{q^*})}$$

$$= \frac{1}{q^*}(1 + \log_2 \mathrm{E}(W^{q^*})) = \frac{\tau(q^*)}{q^*}$$

by the definition of q^*. Substituting into (9.11) immediately gives us

$$\underline{m}(q) \geqslant q \frac{\tau(q^*)}{q^*} \tag{9.12}$$

for all $q > 0$. Comparing (9.12) with (9.10), we see that

$$m(q) =: \lim_{l \to \infty} \frac{1}{l} \log_2 U(q, l) = q \frac{\tau(q^*)}{q^*} \tag{9.13}$$

for any $q \geqslant q^*$. Moreover, using (9.6) with $q \to \infty$ and (9.13), we immediately conclude that

$$m^* := \lim_{l\to\infty} \frac{1}{l} \log_2 U^*(l) = \frac{\tau(q^*)}{q^*}. \tag{9.14}$$

Remark 9.1. For non-conservative cascades, for which the random variables

$$\{W(\mathbf{j}|l), \mathbf{j} \in \{0, 1\}^\infty, l \geqslant 1\}$$

are i.i.d., a statement analogous to (9.14) is equivalent to the so-called *first birth problem*; see, for instance, Kingman (1975). For the particular case of uniformly distributed W in the context of conservative cascades, see also Mahmood (1992).

We are now ready to establish the asymptotic behaviour of the estimator $\hat{\tau}_1(q) = \hat{\tau}_1(q, l)$ in the supercritical case.

Theorem 9.1. *Let $q \geqslant q^*$. Then, as $l \to \infty$,*

$$\hat{\tau}_1(q, l) \to q\frac{\tau(q^*)}{q^*} \quad a.s. \tag{9.15}$$

In particular, the estimator $\hat{\tau}_1(q, l)$ is not a consistent estimator of $\tau(q)$ if $q > q^*$.

Proof. Denote

$$\overline{m}_Z(q) = \limsup_{l\to\infty} \frac{1}{l} \log_2 Z(q, l)$$

and

$$\underline{m}_Z(q) = \liminf_{l\to\infty} \frac{1}{l} \log_2 Z(q, l).$$

Since $Z(q, l) \leqslant U(q, l)$ for all q and l, we immediately conclude by (9.13) that

$$\overline{m}_Z(q) \leqslant m(q) = q\frac{\tau(q^*)}{q^*}. \tag{9.16}$$

For the corresponding lower bound on $\underline{m}_Z(q)$, note that since $P(W \neq \frac{1}{2}) > 0$ and $P(W = 0) < \frac{1}{2}$ (otherwise $q^* = \infty$), there is a $\theta > 0$ such that

$$p_1 := P(|2W - 1| \geqslant \theta) > 0, \qquad p_2 := P(\min(W, 1 - W) \geqslant \theta) > 0.$$

Let $0 < \epsilon < 1$. Note that it follows from (9.14) that for all l large enough,

$$P(U^*(l) \geqslant 2^{(1-\epsilon)l\tau(q^*)/q^*}) \geqslant \frac{1}{2}. \tag{9.17}$$

For $l \geqslant 1$ let

$$N_l = \mathrm{card}\{\mathbf{j}|l : W(\mathbf{j}|i) \geqslant \theta \text{ for all } i = 1, \ldots, l\}.$$

By definition $N_0 = 1$. Observe that, for all $l \geqslant 0$,

$$N_{l+1} = N_l + M_l,$$

where, given $N_0, N_1, \ldots, N_l$, the distribution of M_l is binomial with parameters N_l and p_2. Therefore, (N_l) is a supercritical branching process with progeny mean $m = 1 + p_2 > 1$ and extinction probability 0. By Theorem I.10.3 in Athreya and Ney (1972, p. 30),

$$\lim_{l\to\infty} \frac{N_l}{(1+p_2)^l} = \hat{N} > 0 \quad \text{a.s.} \tag{9.18}$$

Now let $0 < \delta < 1$. It follows by the definition of (N_l) that, for every $l \geqslant 1$,

$$Z(l, q) \geqslant \theta^{[\delta l]q} \max_{k=1,\ldots,N_{[\delta l]}} U_k^*(l - [\delta l])^q |2W_k^{(l)} - 1|^q, \tag{9.19}$$

where

$$(U_k^*(l - [\delta l]),\ k \geqslant 1) \text{ are i.i.d. with the law of } U^*(l - [\delta l])$$

and

$$(W_k^{(l)},\ k \geqslant 1) \text{ are i.i.d. with the law of } \ W.$$

The two sequences are independent, and also independent of $N_{[\delta l]}$. All the random variables defined above can be assumed to be defined, for all l and k, on the same probability space $(\Omega, \mathcal{F}, P)$.

We introduce several events. Let $d = (1 + p_2)^{1/2} > 1$. Put

$$\Omega_1 = \{N_l \geqslant d^l \text{ for all } l \text{ large enough}\}.$$

It follows from (9.18) that $P(\Omega_1) = 1$. Furthermore, let

$$\Omega_2^{(l)} = \bigcup_{k=1}^{d^{[\delta l]}} \{|2W_k^{(l)} - 1| \geqslant \theta \quad \text{and} \quad U_k^*(l - [\delta l]) \geqslant 2^{(1-\epsilon)(l-[\delta l])\tau(q^*)/q^*}\},$$

$l \geqslant 1$. Note that by (9.17) we have $P(\Omega_2^{(l)}) \geqslant 1 - \mathrm{e}^{-c\delta l}$ for some $c > 0$ and all $l \geqslant 1$, and so, letting

$$\Omega_2 = \liminf_{l\to\infty} \Omega_2^{(l)},$$

we see by the Borel–Cantelli lemma that $P(\Omega_2) = 1$. Therefore, $P(\Omega_1 \cap \Omega_2) = 1$ as well. However, for every $\omega \in \Omega_1 \cap \Omega_2$ we have, by (9.19),

$$Z(l, q) \geqslant \theta^{[\delta l]q} 2^{q(1-\epsilon)(l-[\delta l])\tau(q^*)/q^*} \theta^q$$

for all l large enough, which implies that

$$\underline{m}_Z(q) \geqslant q\delta \log_2 \theta + (1-\epsilon)(1-\delta) q \frac{\tau(q^*)}{q^*} \quad \text{a.s.}$$

Letting $\delta \to 0$ and $\epsilon \to 0$, we conclude that

$$\underline{m}_Z(q) \geqslant q \frac{\tau(q^*)}{q^*}. \tag{9.20}$$

Now (9.15) follows from (9.16) and (9.20).

Finally, it follows from Proposition 5.2(iv) that

$$\tau(q) > q\frac{\tau(q^*)}{q^*}$$

for all $q > q^*$. Hence, the estimator $\hat{\tau}_1(q, l)$ is not a consistent estimator of $\tau(q)$ if $q > q^*$.□

Here is an immediate corollary.

Corollary 9.1. *The estimator* $\hat{\tau}_2(q, l)$ *is not a (strongly) consistent estimator of* $\tau(q)$ *if* $q > q^*$.

Proof. Notice that, for every $l \geqslant 1$,

$$\hat{\tau}_1(q, l) = \frac{1}{l}\sum_{j=0}^{l-1}\hat{\tau}_2(q, j),$$

where $Z(q, 0) = 1$. Therefore if, for some $q > q^*$, $\hat{\tau}_2(q, l) \to \tau(q)$ almost surely as $l \to \infty$, then so does $\hat{\tau}_1(q, l)$, which contradicts Theorem 9.1. □

An estimator related to $\hat{\tau}_2(q, l)$ is

$$\hat{\tau}_3(q, l) = \log_2\left(\frac{U(q, l+1)}{U(q, l)}\right) := \log_2 R(q, l), \qquad l \geqslant 1.$$

Since

$$\frac{1}{l}\log_2 U(q, l) = \frac{1}{l}\sum_{j=0}^{l-1}\hat{\tau}_3(q, j), \tag{9.21}$$

where $U(q, 0) = 1$, (9.13) and the same argument as that of Corollary 9.1 show that $\hat{\tau}_3(q, l)$ is not a strongly consistent estimator of $\tau(q)$ if $q > q^*$ (even though it is a strongly consistent estimator of $\tau(q)$ if $q < q^*$). We can say more, however. Note that $0 < R(q, l) < 2$ for all q and l. Furthermore,

$$\mathrm{E}R(q, l) = c(q) = 2^{\tau(q)}, \qquad \text{for all } q \text{ and } l.$$

Therefore if, for some $q > q^*$, $\hat{\tau}_3(q, l)$ converges almost surely to some limit $\tau_3(q)$ as $l \to \infty$, then $2^{\tau_3(q)}$ must have a finite expectation equal to $2^{\tau(q)}$. On the other hand, by (9.13) and (9.21) we must have $\tau_3(q)$ equal to $q\tau(q^*)/q^*$ almost surely. This contradiction shows that $\hat{\tau}_3(q, l)$ cannot converge almost surely as $l \to \infty$ if $q > q^*$.

We conjecture that the same is true for $\hat{\tau}_2(q, l)$, in the sense that it does not converge almost surely as $l \to \infty$ if $q > q^*$. A possibility is that $\hat{\tau}_2(q, l)$ converges in probability and is weakly consistent for $q > q^*$. Whether or not this is true remains an open question.

10. Concluding remarks

While Ossiander and Waymire's estimator for $\tau(q)$ is consistent for random cascades, we also check empirically by simulation that it is an appropriate time-domain method for conservative cascades. By time-domain estimator we mean

$$\hat{\tau}_{\text{time}}(q) = \frac{1}{l}\log_2\left(\sum_{\mathbf{j}|l} |\mu_\infty(I(\mathbf{j}|l))|^q\right),$$

where $\mu_\infty(\cdot)$ is the random measure defined in (2.3). We show in Figure 4 that the time-domain estimator gives equally good results compared with the two wavelet estimators. The cascade generator is a beta distribution with shape parameter 1. The plot is for q-values below $q^* \approx 3.3$. Note that the τ-values are multiplied by -1.

One of the advantages of the wavelet method is its ability to filter deterministic trends because different wavelet families have different vanishing moments; that is, they are orthogonal to low-degree polynomials. The Haar wavelets are 'blind' to additive constants. Figure 5 illustrates the failure of the time-domain method to cope with the presence of an

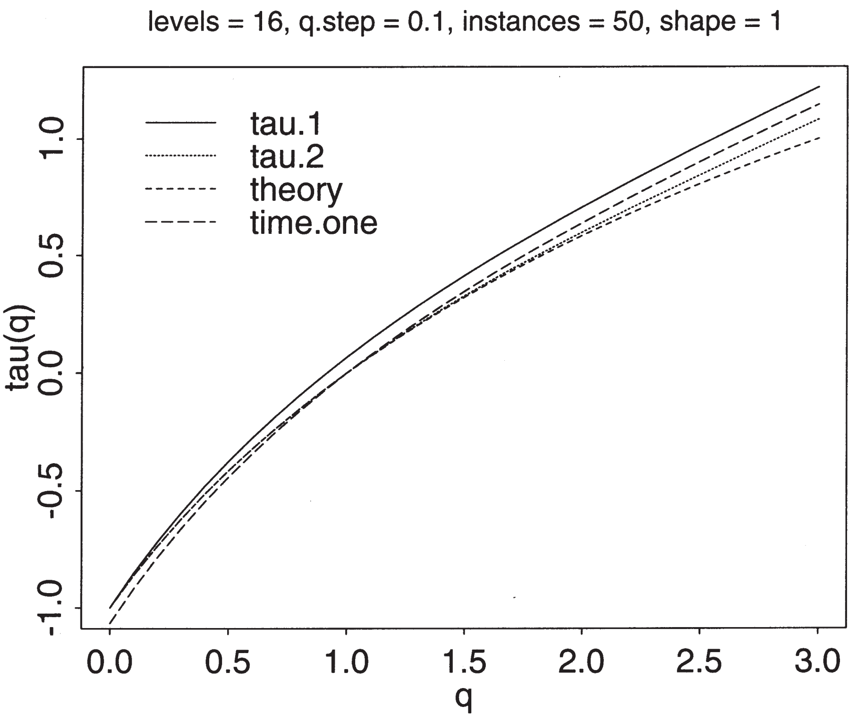


Figure 4. Plots of the two wavelet estimators and the time-domain estimator for $\tau(q)$ with $q < q_* \approx 3.3$

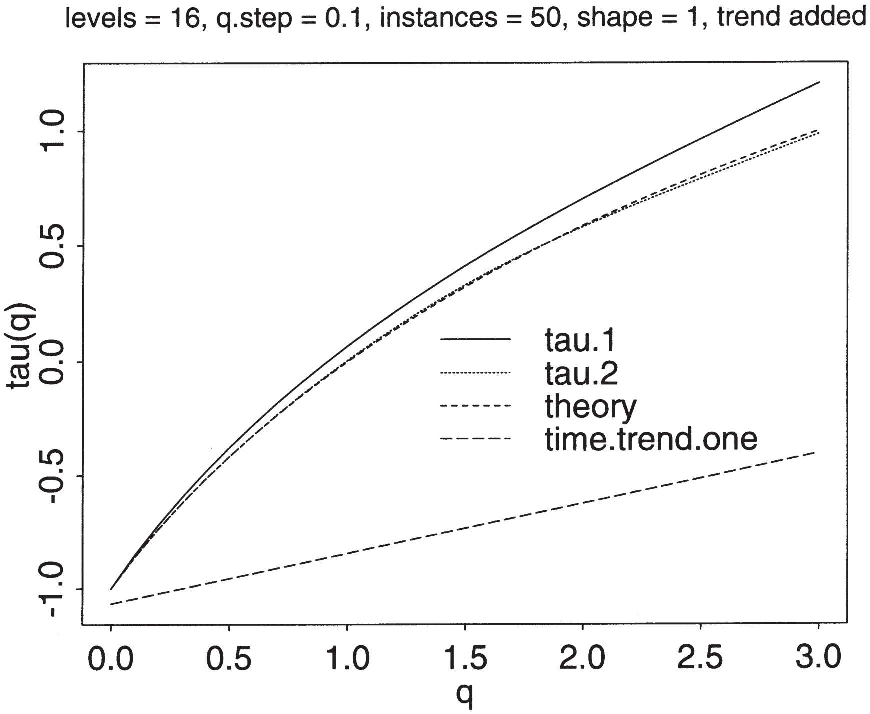


Figure 5. Plots of the wavelet and time-domain estimators for a cascade with an additive constant

additive constant. The cascade is generated with a beta distribution of shape parameter 1 and then a fixed constant 0.1 is added to the cascade. The wavelet estimators give the same values regardless of the presence of the additive constant.

Do our wavelet methods work with other wavelet families? One reason for using other wavelets is that the Haar wavelets have only one vanishing moment and can remove only an additive constant. Other wavelet families with higher vanishing moments can remove higher-degree deterministic trends. Empirical simulated evidence (shown in Figure 6) suggests that other wavelets do indeed work. Figure 6 shows that $\hat{\tau}_1$ works quite well in the case of the D4 wavelet and the presence of an additive linear trend (slope 0.1, intercept 0). Note that the D4 wavelet has four vanishing moments and is hence blind to cubic polynomials). The time-domain method performs poorly, as does $\hat{\tau}_2$. Theoretical investigations are necessary to confirm the validity of the wavelet method for wavelets other than the Haar.

An alternative estimation scheme suggested by an astute referee has coincidentally been implemented in Kulkarni *et al*. (2001) in a study of models of TCP connection traces via products of on–off processes. The idea is to use the specific structure of the conservative cascade to obtain a sample of *W*s by taking ratios at successive levels. Theoretically, this procedure should have good properties from the point of view of asymptotic variance. In

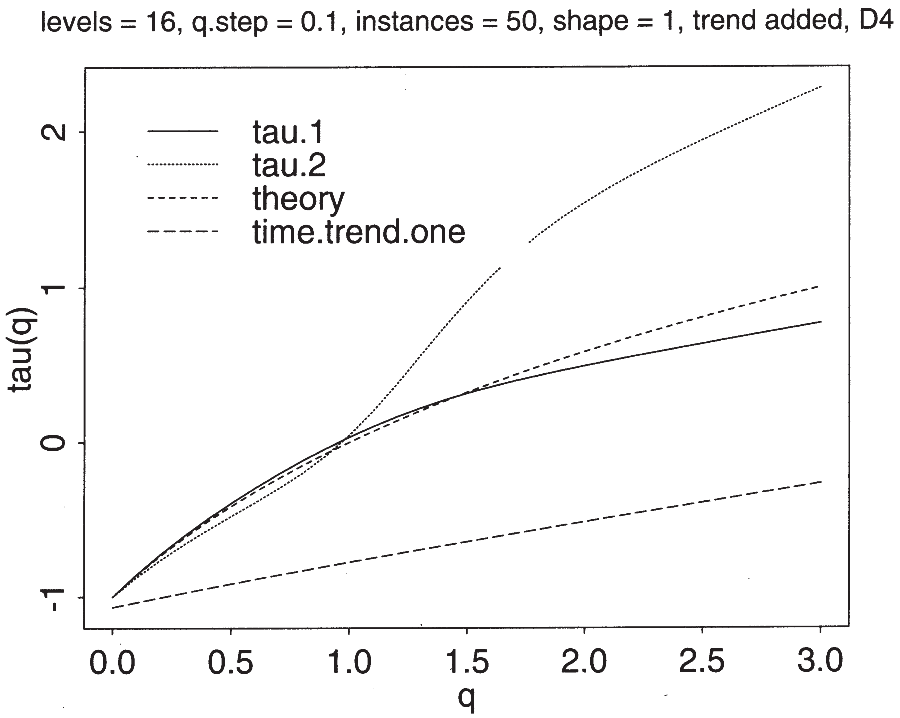


Figure 6. Plots of the wavelet and time-domain estimators using D4 wavelets and in the presence of an additive linear trend

practice the data are culled to ensure independence, resulting in the data set being reduced by half, and this can be a problem. We have preferred to modify a traditional structure function approach by using wavelets.

Acknowledgements

The visits of SR to AT&T Labs – Research were supported by AT&T Labs – Research and a National Science Foundation grant from the Cooperative Research Program in the Mathematical Sciences. SR and GS were also partially supported by NSF grant DMS-97-04982 and NSA grant MDA904-98-1-0041 at Cornell University.

Subsequent to the initial draft of this paper being circulated, one of us (GS) noticed an error which led to revision and a more comprehensive treatment. Independently, Flora Koukiou also noticed the error and provided a proof of (9.13). We thank her for her comments.

References

Abry, P. and Veitch, D. (1998) Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Inform. Theory*, **44**, 2–15.

Athreya, K. and Ney, P. (1972) *Branching Processes*. New York: Springer-Verlag.

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

Feldmann, A., Gilbert, A.C. and Willinger, W. (1998) Data networks as cascades: Investigating the multi-fractal nature of Internet WAN traffic. In *ACM/SIGCOMM '98 Proceedings: Applications, Technologies, Architectures, and Protocols for Computer Communication*, pp. 25–38. New York: Association for Computing Machinery.

Frisch, U. and Parisi, G. (1985) Fully developed turbulence and intermittancy. In M. Ghil (ed.), *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*. Amsterdam: North-Holland.

Gilbert, A.C., Willinger, W. and Feldmann, A. (1999) Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Trans. Inform. Theory*, **45**, 971–991.

Gupta, V.K. and Waymire, E.C. (1993) A statistical analysis of mesoscale rainfall as a random cascade. *J. Appl. Meteorology*, **32**, 251–267.

Holley, R. and Waymire, E.C. (1992) Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Ann. Appl. Probab.*, **2**, 819–845.

Kahane, J.P. and Peyrière, J. (1976) Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.*, **22**, 131–145.

Kingman, J.F.C. (1975) The first birth problem for an age-dependent branching process. *Ann. Probab.*, **3**, 790–801.

Kolmogorov, A.N. (1941) The local structure of turbulence in an incompressible liquid for very large Reynolds numbers. *C.R. (Dokl.) Acad. Sci. URSS (N.S.)*, **30**, 301–305.

Kulkarni, V.G., Marron, J.S. and Smith, F.D. (2001) A cascaded on–off model for tcp connection traces. Technical Report, Department of Statistics, University of North Carolina, Chapel Hill.

Mahmood, H. (1992) *Evolution of Random Search Trees*. New York: Wiley.

Mandelbrot, B.B. (1974) Intermittant turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**, 331–358.

Mandelbrot, B.B. (1990) Limit lognormal multifractal measures. In E. Gotsman, Y. Ne'eman and A. Voronel (eds), *Frontiers of Physics: Proceedings of the Landau Memorial Conference*, pp. 309–340. Oxford: Pergamon.

Mandelbrot, B.B. (1998) *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. New York: Springer-Verlag.

Meneveau, C. and Sreenivasan, K.R. (1987) Simple multifractal cascade model for fully developed turbulence. *Phys. Rev. Lett.*, **59**, 1424–1427.

Neveu, J. (1975) *Discrete-Parameter Martingales*, North-Holland Math. Library 10. Amsterdam: North-Holland.

Ossiander, M. and Waymire, E.C. (2000) Statistical estimation for multiplicative cascades. *Ann. Statist.*, **28**, 1533–1560.

Resnick, S.I. (1998) *A Probability Path*. Boston: Birkhäuser.

Riedi, R.H. and Lévy Véhel, J. (1997) Multifractal properties of TCP traffic: A numerical study. INRIA Technical Report 3129.

Troutman, B.M. and Vecchia, A.V. (1999) Estimation of Rényi exponents in random cascades. *Bernoulli*, **5**, 191–207.

Received August 2000 and revised June 2002