WAVELET ANALYSIS OF CONSERVATIVE CASCADES

SIDNEY RESNICK, ANNA GILBERT, AND WALTER WILLINGER

Abstract. 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 analyzing multifractal behavior of measured Internet traffic traces, we consider wavelet based statistical techniques for inference about the cascade *generator*, the random mechanism determining the re-distribution of the set's mass at each iteration. We provide two estimators of the structure function, one asymptotically biased and one not, prove consistency and asymptotic normality in a range of values of the argument of the structure function less than a critical value q_*. Simulation experiments illustrate the asymptotic properties of these estimators and provide also interesting conjectures for the uninformative behavior of the estimators beyond the critical value q_*.

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 behavior associated with many naturally occurring phenomena, e.g., fully developed turbulence (see [7, 8, 3, 11] and references therein); spatial rainfall [5]; the movements of stock prices [10]; and Internet traffic dynamics [15, 2].

The *generator* of a cascade determines the re-distribution 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 [9] (also in the turbulence context), conservative cascades have recently been considered in [2] for use in describing the observed multifractal behavior of measured Internet traffic traces. In particular, Feldmann et al. [2] build on empirical evidence that measured Internet traffic is consistent with multifractal behavior by illustrating that "... data networks appear to act as conservative cascades!" They demonstrate that

- multiplicative and measure-preserving structure becomes most apparent when analyzing 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,
- this structure is recovered at the aggregate level (i.e., when considering the superposition process consisting of all IP packets generated by all active TCP connections) and causes aggregate Internet traffic to exhibit multifractal behavior.

The visits of Sidney Resnick 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.

Well short of providing a physical explanation to the all-important networking question of "Why do packets within individual TCP connections conform to a conservative cascade?" the work in [2] is empirical in nature and relies on a number of heuristics for inferring multifractal behavior 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 essence, 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 characterize the singularity behavior of a signal locally in time. Moreover, since the local Hölder exponent at a point in time t_0 describes the local scaling behavior of the signal as we look at smaller and smaller neighborhoods 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 behavior of a process can (under certain assumptions) be fully recovered by studying the singularity behavior in the wavelet domain; i.e., 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 [4]) and hence provides a promising setting for relating the local scaling behavior 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 *Mandelbrot-Kahane-Peyriere (MKP) function* [6]) 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 [2, 4], we have clearly benefited from the recent random cascade work of Ossiander and Waymire [13]. 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 [13] 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 behavior exhibited by random cascades (for related work, see also [16]). While the large-sample properties of the time domain-based estimators considered in [13] 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 is concerned with certain martingales and serves as warm-up for Section 6 where under appropriate conditions we show strong consistency of our two wavelet-based estimators. Asymptotic normality (together with the problem of bias) of the estimators is proven and illustrated with some simulated data in Section 7. We conclude in Section 8 with some observations, open problems and speculations regarding the large sample properties of our estimators when the conditions in Section 6 (or Section 7) are not satisfied. In particular we present heuristic and empirical evidence that the values of the estimators at large values of the argument of the structure function are likely to be uninformative and even misleading, thus providing some practical guidance for properly interpreting the plots associated with the estimation procedure.

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 $(j_1, \dots, j_l) \in \{0,1\}^l$. 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

$$I(j_1, \dots, j_l) = \Big[\sum_{k=1}^{l} \frac{j_k}{2^k}, \sum_{k=1}^{l} \frac{j_k}{2^k} + \frac{1}{2^l}\Big), \quad (j_1, \dots, j_l) \in \{0,1\}^l. \tag{2.1}$$

An infinite path through the tree is denoted by

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

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

$$\boldsymbol{j}|l = (j_1, \dots, j_l).$$

We will sometimes write when convenient

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

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 $1/2$ so that $W \stackrel{d}{=} 1 - W$. The symmetry implies that $E(W) = 1/2$. There is a family of identically distributed random variables

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

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

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

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

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

Note the conservative property entails that

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

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

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

3. Wavelet Coefficients.

We compute the wavelet transform

$$d_{-l,n} = \int_0^1 \psi_{-l,n}(x)\mu_\infty(dx), \quad n = 0, \dots, 2^l - 1;\ l \geq 1, \tag{3.1}$$

using the Haar wavelets

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

We have by examining where the Haar wavelet is constant that

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

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

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

and using the conservative property (2.2), this is

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

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

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

4. Notation Glossary.

Before continuing the analysis, we collect some notation in one place for easy reference. We seek to estimate the distribution of the cascade generator W and this will accomplished if we estimate

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

or equivalently we could estimate the *structure function*

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

We will estimate using estimators constructed from the process

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

and note 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

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

and note the normalization makes

$$E(M(q,l)) = 1.$$

There are further constant functions needed:

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

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

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

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

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

$$\sigma_2^2(q) := \frac{1}{b^2}\mathrm{Var}(|2W-1|^q). \tag{4.10}$$

$$\sigma_3^2(q) := \frac{1}{b^2}\mathrm{Var}\Big(\frac{W_1^q}{c}|2W_2 - 1|^q + \frac{(1-W_1)^q}{c}|2W_3 - 1|^q - |2W_1 - 1|^q\Big), \tag{4.11}$$

where $W_i,\ i = 1,2,3$ are iid with the distribution of the cascade generator.

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

Proposition 4.1. (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 $r > 1$.*

Proof. (i) It is convenient to define $W = e^{-Y}$ so that the Laplace transform of Y is

$$\phi(q) := Ee^{-qY} = E(W^q)$$

and

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

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 iff

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

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

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

and this is strictly positive iff

$$\phi(q)\phi''(q) > (\phi'(q))^2, \tag{4.12}$$

that is iff

$$E(e^{-qY})E(Y^2e^{-qY}) > \left(E(Ye^{-q/2Y} \cdot e^{-q/2Y})\right)^2$$

which follows from the Cauchy–Schwartz inequality.

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

$$\frac{d^2}{dr^2}(\log a_r(q)) = \frac{q^2}{\phi^2(rq)}\left[\phi''(rq)\phi(rq) - (\phi'(rq))^2\right]$$

which is positive by (4.12). □

We now define the quantity

$$q_* := \sup\{q > 0 : a(q) < 1\} \tag{4.13}$$

so that for $q < q_*$ we have $a(q) < 1$. Our estimators will behave differently on the regions $q < q_*$ and $q \geq q_*$. Observe that it is always the case that $q_* > 1$ since

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

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

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

and

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

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

$$E(W^q) = \frac{\Gamma(2\alpha)\Gamma(\alpha+q)}{\Gamma(\alpha)\Gamma(2\alpha+q)} \tag{4.14}$$

and q_* satisfies

$$4^{-\alpha}\sqrt{\pi}\Gamma(\alpha+2q_*)\Gamma(2\alpha+q_*)^2 - \Gamma(\alpha+1/2)\Gamma(2\alpha+2q_*)\Gamma(\alpha+q_*)^2 = 0.$$

Example 3. Suppose W has the two point distribution concentrating mass $1/2$ at $\pm p$ for some $0 \leq p < 1/2$. Then

$$E(W^q) = \frac{1}{2}\left(p^q + (1-p)^q\right) \tag{4.15}$$

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 $q_* = \infty$.

Example 4. If W does not have a two point distribution but nevertheless has an atom of size $p_1 < 1/2$ at 1 (and hence by symmetry there is an atom of the same size at 0) we have $q_* < \infty$. So see this note as $q \to \infty$

$$a(q) = \frac{EW^{2q}}{2(E(W^q))^2} \to \frac{P[W=1]}{2P[W=1]^2} = \frac{1}{2p_1} = a(\infty) > 1.$$

It is possible to express $a(\infty)$ in terms of $\tau(\cdot)$ defined in (4.2) and hence arrive at a condition that $a(\infty) \leq 1$ and $q_* = \infty$. Write

$$\log_2 a(q) = \log_2\Big(\frac{c(2q)}{c^2(q)}\Big) = \tau(2q) - 2\tau(q),$$

so that

$$\log_2 a(\infty) = \lim_{q \uparrow \infty}\Big[\tau(2q) - 2\tau(q)\Big].$$

After writing $\tau(\cdot)$ as the integral of its derivative, we find the following result.

Proposition 4.2. *If $\tau'(\cdot)$ is the derivative of $\tau(\cdot)$ we have*

$$\int_0^\infty \Big(\tau'(2s) - \tau'(s)\Big)ds = \frac{\log_2 a(\infty) + 1}{2}, \tag{4.16}$$

and hence

$$a(\infty) < 1 \text{ iff } \int_0^\infty \Big(\tau'(2s) - \tau'(s)\Big)ds < \frac{1}{2}. \tag{4.17}$$

The following result will be needed to establish asymptotic normality of our estimators.

Proposition 4.3. *Given $2q < q_*$ so that*

$$a_2(2q) = a(2q) < 1,$$

we have for any $\delta \in (0,2)$ that

$$a_{1+\delta/2}(2q) = \frac{c((2+\delta)q)}{(c(2q))^{1+\delta/2}} < 1. \tag{4.18}$$

Proof. We are given $\log a_2(2q) < 0$ and since $a_1(2q) = 1$ we have $\log a_1(2q) = 0$. Since $\log a_r(2q)$ is convex in $r > 1$, we conclude $\log a_r(2q) < 0$ or equivalently $a_r(2q) < 1$ for $1 < r < 2$. □

5. The Associated Martingale.

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

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

generated by the weights up to and including depth l.

Proposition 5.1. *For each $q > 0$, the family*

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

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

$$M(q,l) \stackrel{a.s.}{\to} M(q,\infty).$$

Proof. The martingale property is easily established:

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

By the martingale convergence theorem (eg, [14], [12]) a non-negative martingale always converges almost surely. ☐

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

$$W^q + (1-W)^q = p^q + (1-p)^q. \tag{5.1}$$

Recall (4.15) and then observe for $l > 1$

$$\begin{aligned} M(q,l) &= \sum_{\boldsymbol{j}|l}\prod_{i=1}^{l}\frac{W^q(\boldsymbol{j}|i)}{c^l} = \sum_{\boldsymbol{j}|l-1}\prod_{i=1}^{l-1}\frac{W^q(\boldsymbol{j}|i)}{c^{l-1}}\sum_{j_l}\frac{W^q(\boldsymbol{j}|l)}{c} \\ &= \sum_{\boldsymbol{j}|l-1}\prod_{i=1}^{l-1}\frac{W^q(\boldsymbol{j}|i)}{c^{l-1}}\Big[W^q(\boldsymbol{j}|l-1,0) + W^q(\boldsymbol{j}|l-1,1)\Big]/c \end{aligned}$$

and since $W(\boldsymbol{j}|l-1,1) = 1 - W(\boldsymbol{j}|l-1,0)$ we apply (5.1) to get

$$\begin{aligned} &= \sum_{\boldsymbol{j}|l-1}\prod_{i=1}^{l-1}\frac{W^q(\boldsymbol{j}|i)}{c^{l-1}}\Big[p^q + (1-p)^q\Big]/c \\ &= \sum_{\boldsymbol{j}|l-1}\prod_{i=1}^{l-1}\frac{W^q(\boldsymbol{j}|i)}{c^{l-1}} = M(q,l-1). \end{aligned}$$

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

$$d(q,l) := M(q,l) - M(q,l-1), \quad l \geq 1.$$

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

$$\begin{aligned} d(q,l) &= \sum_{\boldsymbol{j}|l-1} \prod_{i=1}^{l-1} \frac{W^q(\boldsymbol{j}|i)}{c^{l-1}} \Big[\frac{W^q(\boldsymbol{j}|l-1,0) + (1-W(\boldsymbol{j}|l-1,0)^q}{c} - 1 \Big] \\ &= \sum_{\boldsymbol{j}|l-1} \prod_{i=1}^{l-1} \frac{W^q(\boldsymbol{j}|i)}{c^{l-1}} \Big[\xi(\boldsymbol{j}|l) \Big]. \end{aligned} \tag{5.2}$$

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

$$\sigma_1^2(q) = \operatorname{Var}(\xi(\boldsymbol{j}|l) = \frac{1}{c^2} \operatorname{Var}(W^q + (1-W)^q).$$

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

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

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

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

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

$$\sum_{i=1}^{l} E(d^2(q,i)|\mathcal{F}_{i-1}) = \sum_{i=1}^{l} M(2q, i-1) a^{i-1}(q) \sigma_1^2(q). \tag{5.3}$$

Furthermore,

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

and thus

$$\operatorname{Var}(M(q,l)) = \sum_{i=1}^{l} E(d^2(q,i)) = \sigma_1^2(q) \sum_{i=1}^{l} a^{i-1}(q). \tag{5.4}$$

This leads to the following facts.

Proposition 5.2. *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 $E(M(q,\infty) = 1$,*

$$M(q,l) = E(M(q,\infty)|\mathcal{F}_l)$$

and $M(q,l) \to M(q,\infty)$ almost surely and in L_2.

Proof. We have from (5.4) that

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

The rest follows from standard martingale theory (eg, [12, page 68], [14]). □

5.1. **The distribution of $M(q,\infty)$.** The distribution of $M(q,\infty)$ satisfies a simple recursion which can be used to get useful information.

Proposition 5.3. *Suppose $\{M(q,\infty), M_1(q,\infty), M_2(q,\infty)\}$ are iid 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*

$$M(q,\infty) \stackrel{d}{=} W^q \frac{M_1(q,\infty)}{c(q)} + (1-W)^q \frac{M_2(q,\infty)}{c(q)}. \tag{5.5}$$

Proof. We write

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

□

From this recursion (5.5) we conclude the following zero/one law.

Proposition 5.4. *For the conservative cascade*

$$P[M(q,\infty) = 0] = 0 \text{ or } 1 \tag{5.6}$$

for any $q > 0$.

Proof. Define

$$\begin{aligned} p_0 =& P[M(q,\infty)=0] \\ p_W(0) =& P[W=0]=P[W=1]. \end{aligned}$$

Then since $c(q)\neq 0$

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

From this we conclude

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

so that

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

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

Knowing the zero/one law allows one to get a criterion for when the two cases hold.

Proposition 5.5. (i) *If $q<q_*$, then $P[M(q,\infty)=0]=0$.*
(ii) *If $q\geq q_*$, then $P[M(q,\infty)=0]=1$.*

Proof. Square terms in (5.5) to get

$$M^2(q,\infty)=W^{2q}\frac{M_1^2(q,\infty)}{c^2}+(1-W)^{2q}\frac{M_2^2(q,\infty)}{c^2}+2W^q(1-W)^q\frac{M_1(q,\infty)M_2(q,\infty)}{c^2}.$$

Let $m_2=EM^2(q,\infty)$ and let $m_1=EM(q,\infty)$ and taking expectations in the previous equation yields

$$\begin{aligned} m_2 =& EW^{2q}\frac{m_2}{c^2(q)}+EW^{2q}\frac{m_2}{c^2(q)}+2E\big(W^q(1-W)^q\big)\frac{m_1^2}{c^2(q)} \\ =& \frac{c(2q)}{c^2(q)}m_2+2E\big(W^q(1-W)^q\big)\frac{m_1^2}{c^2(q)} \\ =& a(q)m_2+2E\big(W(1-W)\big)^q\frac{m_1^2}{c^2(q)}. \end{aligned} \tag{5.7}$$

If $a(q)<1$ then from Proposition 5.2 we have $m_1=1$ and

$$m_2=\frac{2E(W(1-W))^q}{(1-a(q))c^2(q)}.$$

If m_2 were 0, then $E(W(1-W))^q=0$ would imply $W(1-W)=0$ almost surely and almost surely $W=0$ or 1. Since W has a symmetric distribution we would have $P[W=0]=P[W=1]=1/2$ and thus W would have a two point distribution which requires $M(q,\infty)=1$ almost surely. This contradicts $m_2=0$. Thus $a(q)<1$, implies $m_2=EM^2(q,\infty)>0$. Since $P[M(q,\infty)=0]=0$ or 1, this means $P[M(q,\infty)=0]=0$.

If $a(q)=1$, then from (5.7)

$$2E(W(1-W))^q m_1^2/c^2(q)=0$$

which implies $m_1^2=0$. However, $m_1=EM(q,\infty)=0$ entails $M(q,\infty)=0$ almost surely.

If $a(q) > 1$, then from (5.7)

$$0 = (a(q) - 1)m_2 + 2E(W(1-W))^q \frac{m_1^2}{c(q)}$$

so that $m_2 = 0$ which again entails $M(q, \infty) = 0$ almost surely. □

6. Strong Laws and Strong Consistency

Now with some understanding of the martingale $\{M(q,l), l \geq 1\}$, we show that when $q < q_*$, the scaled summed powers of the wavelet coefficients $\{Z(q,l), l \geq 1\}$ has the same behavior. Recall notational definitions (4.1), (4.2), (4.3), (4.6).

Proposition 6.1. *If $q < q_*$ so that $a(q) < 1$,*

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

almost surely and in L_2. Thus

$$\frac{Z(q,l)}{c^l b} \to M(q,\infty) \tag{6.1}$$

almost surely and in L_2.

Proof. Begin by writing

$$\begin{aligned}\frac{Z(q,l)}{c^l b} - M(q,l) &= \sum_{\boldsymbol{j}|l} \prod_{i=1}^{l} \frac{W^q(\boldsymbol{j}|i)}{c^l} \Big[\frac{|2W(\boldsymbol{j}|l,0) - 1|^q}{b} - 1\Big] \\ &= \sum_{\boldsymbol{j}|l} \prod_{i=1}^{l} \frac{W^q(\boldsymbol{j}|i)}{c^l} \xi(\boldsymbol{j}|l,0)\end{aligned}$$

where $\xi(\boldsymbol{j}|l,0) \perp \xi(\boldsymbol{p}|l,0)$ if $\boldsymbol{j}|l \neq \boldsymbol{p}|l)$. Also $E\xi(\boldsymbol{j}|l,0) = 0$ and recall the notation from (4.10)

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

Similar to the calculations leading to (5.3) and (5.4) we find

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

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

A simple Borel-Cantelli argument now gives the almost sure convergence: We have by Chebychev for any $\delta > 0$

$$\begin{aligned}P[\Big|\frac{Z(q,l)}{c^l b} - M(q,l)\Big| > \delta] \leq& \frac{1}{\delta^2}E\Big|\frac{Z(q,l)}{c^l b} - M(q,l)\Big|^2 \\ =&\frac{1}{\delta^2}\sigma_2^2(q)a^l(q),\end{aligned}$$

and therefore

$$\sum_l P\Big[\Big|\frac{Z(q,l)}{c^l b} - M(q,l)\Big| > \delta\Big] < \infty.$$

This yields the almost sure convergence to 0. □

We can use this strong law to get consistent estimators of the structure function $\tau(q)$.

Proposition 6.2. *Define*

$$\hat{\tau}_1(q) = \frac{\log_2 Z(q,l)}{l} = \frac{\log_2 \sum_{n=0}^{2^l-1} |d_{-l,n}|^q - ql/2}{l} \tag{6.2}$$

$$\hat{\tau}_2(q) = \log_2\left(\frac{Z(q,l+1)}{Z(q,l)}\right) = \log_2\left(\frac{\sum_{n=0}^{2^{l+1}-1} |d_{-(l+1),n}|^q}{2^{q/2}\sum_{n=0}^{2^l-1} |d_{-l,n}|^q}\right). \tag{6.3}$$

Provided $q < q_*$, both estimators are almost surely consistent for $\tau(q)$:

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

as $l \to \infty$.

Proof. In (6.1), take logarithms to the base 2 to get

$$\log_2 Z(q,l) - l\log_2 c(q) - \log_2 b \to \log_2 M(q,\infty), \tag{6.4}$$

almost surely as $l \to \infty$. Divide through by l to get consistency of $\hat{\tau}_1(q)$. To get the consistency of $\hat{\tau}_2(q)$, note from (6.4) 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 $\hat{\tau}_2(q)$. □

7. Central Limit Theorems and Asymptotic Normality of Estimators.

In this section we discuss second order properties of the estimators $\tau_i(q)$, $i = 1, 2$, defined in (6.2) and (6.3). The asymptotic normality for $\tau_1(q)$ requires a bias term which cannot be eliminated. This drawback, is eliminated by using $\tau_2(q)$, whose definition in terms of differencing removes the bias term. For this section it is convenient to write $E^{\mathcal{F}_l}$ and $P^{\mathcal{F}_l}$ for the conditional expectation and conditional probability with respect to the σ-field $\mathcal{F}_l$.

7.1. Asymptotic Normality of $\tau_1(q)$.

Begin by writing

$$\frac{Z(q,l)}{c^l b} - M(q,l) = \sum_{\boldsymbol{j}|l} \prod_{i=1}^{l} \frac{W^q(\boldsymbol{j}|i)}{c^l}\Big[\frac{|2W(\boldsymbol{j}|l,0) - 1|^q}{b} - 1\Big] \tag{7.1}$$

$$=: \sum_{\boldsymbol{j}|l} Z(\boldsymbol{j}|l) \tag{7.2}$$

where

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

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

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

$$\sum_{\boldsymbol{j}|l} E^{\mathcal{F}_l}(Z(\boldsymbol{j}|l))^2 = M(2q,l)a^l(q)\sigma_2^2(q). \tag{7.3}$$

Our strategy for the central limit theorem is to regard $\frac{Z(q,l)}{c^l b} - M(q,l)$ as a sum of random variables which are conditionally independent given $\mathcal{F}_l$ and then apply the Liapunov condition ([14] for asymptotic normality in a triangular array.

Proposition 7.1. *If $2q < q_*$ so that $a(2q) < 1$, then as $l \to \infty$*

$$P^{\mathcal{F}_l}\left[\frac{\frac{Z(q,l)}{c^l b} - M(q,l)}{\sqrt{M(2q,l)a^l(q)\sigma_2^2(q)}} \le x\right] \to P[N(0,1) \le x] \tag{7.4}$$

where $N(0,1)$ is a standard normal random variable with mean 0 and variance 1. Taking expectations in (7.4) yields

$$P\left[\frac{\frac{Z(q,l)}{c^l b} - M(q,l)}{\sqrt{M(2q,l)a^l(q)\sigma_2^2(q)}} \le x\right] \to P[N(0,1) \le x]. \tag{7.5}$$

Proof. Pick $\delta > 0$ so that $2q + \delta < q_*$. From Proposition 4.3 we have

$$a_{1+\delta/2}(2q) < 1. \tag{7.6}$$

Asymptotic normality in (7.4) will be shown if we establish the Liapunov condition

$$\frac{\sum_{\boldsymbol{j}|l} E^{\mathcal{F}_l}|Z(\boldsymbol{j}|l)|^{2+\delta}}{(M(2q,l)a^l(q))^{(2+\delta)/2}} \to 0, \tag{7.7}$$

where the denominator comes from (7.3). The numerator in the left side of (7.7) is bounded above by

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

where

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

So the ratio in (7.7), 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 5.5. Check that

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

So the Liapunov ratio is asymptotic to a finite nonzero random variable times $(a_{1+\delta/2}(2q))^l$ where $a_{1+\delta/2}(2q) < 1$ and the result is proven. ∎

Remark 7.1. In the denominator of (7.5) 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

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

by [1]. The desired result is obtained by multiplying components. □

Remark 7.2. Set

$$N_l := \frac{\frac{Z(q,l)}{c^l b} - M(q,l)}{\sqrt{M(2q,l)a^l(q)\sigma_2^2(q)}}. \tag{7.8}$$

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

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

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

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

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

By the central limit theorem Proposition 7.1,

$$P^{\mathcal{F}_{l+1}}[N_{l+1} \le y] = \Phi(y) + \epsilon_l(y)$$

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

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

from the dominated convergence theorem, and hence we get convergence to

$$\to \Phi(x)\Phi(y).$$

□

We now describe how this central limit behavior transfers to $\tau_1(q)$.

Corollary 7.1. *Under the assumptions in force in Proposition 7.1, we have*

$$\frac{\big(\hat\tau_1(q) - \tau(q)\big) - l^{-1}\log_2 bM(q,l)}{\frac{\sqrt{M(2q,\infty)a^l(q)\sigma_2^2(q)}}{l\log 2\cdot M(q,l)}} \Rightarrow N(0,1). \tag{7.9}$$

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

Proof. For brevity, write

$$d(q) := M(2q,\infty)a^l(q)\sigma_2^2(q) \to 0, \tag{7.10}$$

and using the notation of (7.8) 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

$$l\big(\hat\tau_1(q) - \tau(q)\big) = \log_2 bM(q,l) + \log_2\Big(1 + \frac{N_l\sqrt{d(q)}}{M(q,l)}\Big).$$

Since $N_l\sqrt{d(q)}/M(q,l) \to 0$ the desired result follows by using the relation $\log(1+x) \sim x$ for $x \downarrow 0$. $\square$

The bias term in (7.9) is an unpleasant feature and thus we consider how to remove it by differencing. This leads to consideration of the asymptotic normality of the estimator $\hat\tau_2(q)$.

7.2. **Asymptotic Normality of $\tau_2(q)$.** We now consider the asymptotic normality of $\tau_2(q)$. It is possible to proceed from Proposition 7.1 but it turns out to be simpler to proceed with a direct proof.

Proposition 7.2. *Suppose $2q < q_*$. Then*

$$\frac{\hat\tau_2(q) - \tau(q)}{\frac{\sqrt{M(2q,\infty)a^l(q)\sigma_3^2(q)}}{\log 2\cdot M(q,\infty)}} \Rightarrow N(0,1), \tag{7.11}$$

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

Proof. Begin by observing that

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

where we have set

$$H(\boldsymbol{j}|l) = \frac{W^q(\boldsymbol{j}|l,0)}{c}\frac{|2W(\boldsymbol{j}|l,0,0)-1|^q}{b} + \frac{W^q(\boldsymbol{j}|l,1}{c}\frac{|2W(\boldsymbol{j}|l,1,0)-1|^q}{b} - \frac{|2W(\boldsymbol{j}|l,0)-1|^q}{b}. \tag{7.12}$$

Check that

$$E^{\mathcal{F}_l}H(\boldsymbol{j}|l) = \frac12 + \frac12 - 1 = 0$$

and

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

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 iid random variables with (conditional) variance

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

As in the proof of Proposition 7.1 we may check the Liapunov condition and conclude

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

or equivalently

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

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

$$
\frac{\Big(c^{-1}\frac{Z(q,l+1)}{Z(q,l)}-1\Big)M(q,\infty)}{\sqrt{M(2q,\infty)a^l(q)\sigma_3^2(q)}}\Rightarrow N(0,1). \tag{7.13}
$$

Since

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

it follows that

$$
\begin{aligned}
\hat{\tau}_2(q)-\tau(q)=&\log_2\Big(c^{-1}\frac{Z(q,l+1)}{Z(q,l)}\Big)\\
=&\frac{\log\Big(1+\big(c^{-1}\frac{Z(q,l+1)}{Z(q,l)}-1\big)\Big)}{\log 2}\\
\sim&\frac{c^{-1}\frac{Z(q,l+1)}{Z(q,l)}-1}{\log 2}
\end{aligned}
$$

in probability. Combine this with (7.13) to complete the proof. □

For statistical purposes, the result (7.11) contains unobservables so as in [16, 13], 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 $\{Z(q,l)\}$.

We define the following useful observable quantity

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

$$\left. - \frac{|2W(\boldsymbol{j}|l,0)-1|^q}{Z(q,l)} \right]^2 \tag{7.14}$$

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

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

Note that in this notation,

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

where $H(\boldsymbol{j}|l)$ is defined in (7.12). Recall also that $EH^2(\boldsymbol{j}|l) = \sigma_3^2(q)$. In terms of the wavelet coefficients we have

$$D^2(q,l) = \sum_{\boldsymbol{j}|l} \left[\frac{|d(-l,(\boldsymbol{j}|l,0))|^q 2^{-q(l+1)/2}}{Z(q,l+1)} + \frac{|d(-l,(\boldsymbol{j}|l,1))|^q 2^{-q(l+1)/2}}{Z(q,l+1)} - \frac{|d(-l,(\boldsymbol{j}|l)|^q 2^{-ql/2}}{Z(q,l)} \right]^2, \tag{7.15}$$

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

Corollary 7.2. *Suppose $2q < q_*$. Then*

$$\frac{\hat{\tau}_2(q) - \tau(q)}{D(q,l)/\log 2} \Rightarrow N(0,1) \tag{7.16}$$

as $l \to \infty$.

Proof. Because of (7.11), 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

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

After some simple algebra, this ratio is the same as

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

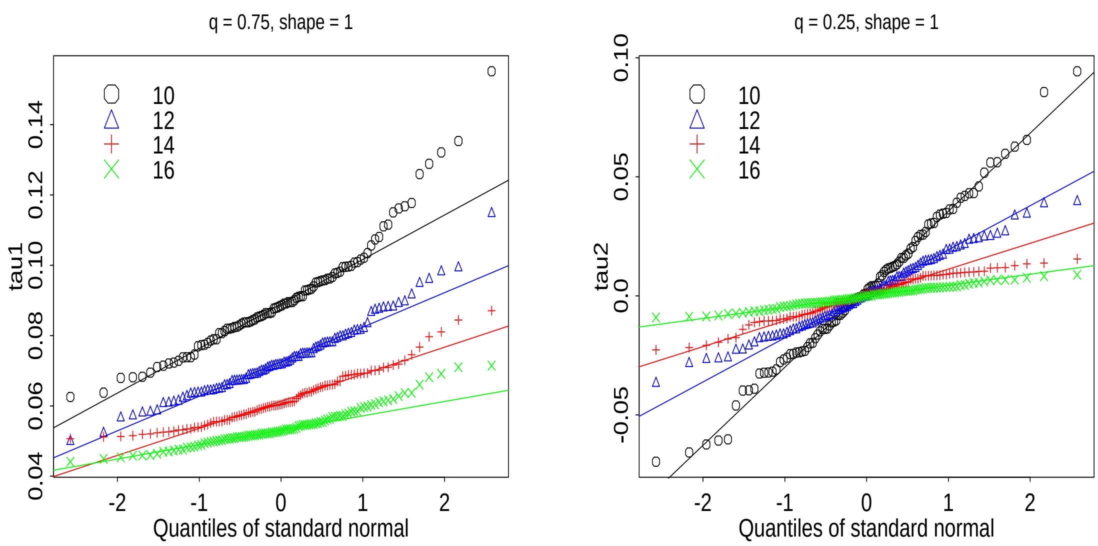

FIGURE 1. Normal QQ-plots of $\hat{\tau}_1(q) - \tau(q)$ (left) and $\hat{\tau}_2(q) - \tau(q)$ (right).

Since $M(2q, l) \to M(2q, \infty)$, it suffices to show the numerator converges in probability to $M(2q, \infty)\sigma_3^2(q)$. Due to (6.1), we write the numerator as

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

As in Theorem 3.5 of [13],

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

as desired. The terms I and II can readily be shown to go to 0. $\square$

Figure 1 shows normal QQ plos of $\tau_i(q) - \tau(q)$, $i = 1, 2$, from simulatd cascade data with cascade generator with shape parameter 1. The left plot is for $q = 0.75$ and the right is for $q = 0.25$. Each plot presents 4 graphs as the depth l increases to 16. Note the better agreement of $\tau_2(q)$ to normality compared with $\tau_1(q)$.

8. CONCLUDING REMARKS AND SPECULATION

A critical issue with both the wavelet based estimator and the moment based ones used in [13], is that the asymptotic properties of the estimators are only valid in a certain range of q-values. For the wavelet estimators, we require $q < q_*$ and for the asymptotic normality results we require $2q < q_*$. We now suggest that the range $q > q_*$ is uninformative for our estimators and in fact, it may be rather misleading to extend the inference to excessively large values of 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 heuristics.

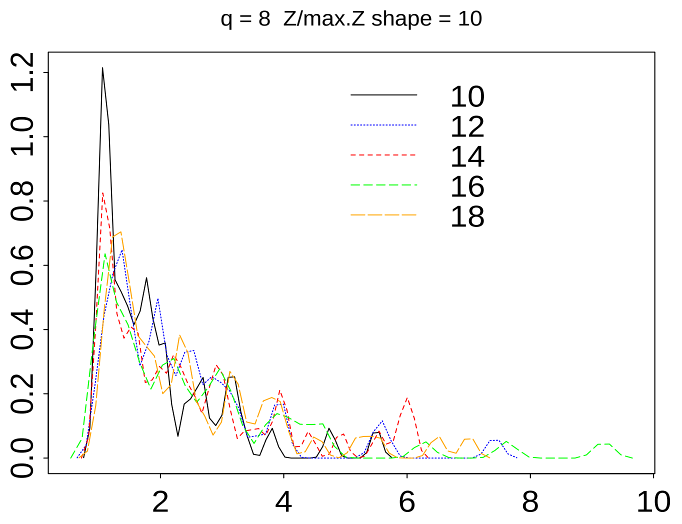

FIGURE 2. Density plots of $Z(q,l)/Z^{\vee}(q,l)$.

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)

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

Note that

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

For large values of q, presumably for $q>q_*$, it is likely that $Z(q,l)$ is determined by its largest summand and thus we expect that as $l\to\infty$

$$Z(q,l)/Z^{\vee}(q,l)\xrightarrow{P}1.$$

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 the densities concentrate most mass around the point 1. If this speculation is true, then $\log Z^{\vee}(q,l)=q\log Z^{\vee}(1,l)$ is linear in q and thus $\hat{\tau}_i(q)$ should also be linear in q $(i=1,2)$ rendering these estimators largely uninformative for inference purposes in the q–region where all these speculative comments are valid. A rough estimate of q_* would be provided by the q-value where the plots of $\hat{\tau}_i(q)$ start to look linear.

We currently have no proofs for these speculative comments but computer simulations offer strong support. 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 $q_*\approx 2.4$, it is remarkable how linear the plots look and also how closely $\hat{\tau}_i^{\vee}(q)$ approximates $\hat{\tau}_i(q)$.

We intend to examine this phenomenon much more closely.

Finally, to emphasize that there are differences between the random cascade models considered in [13] and the conservative cascade processes studied in this paper that affect the inference problem, we consider more closely the critical range of moments of order q for which the large sample

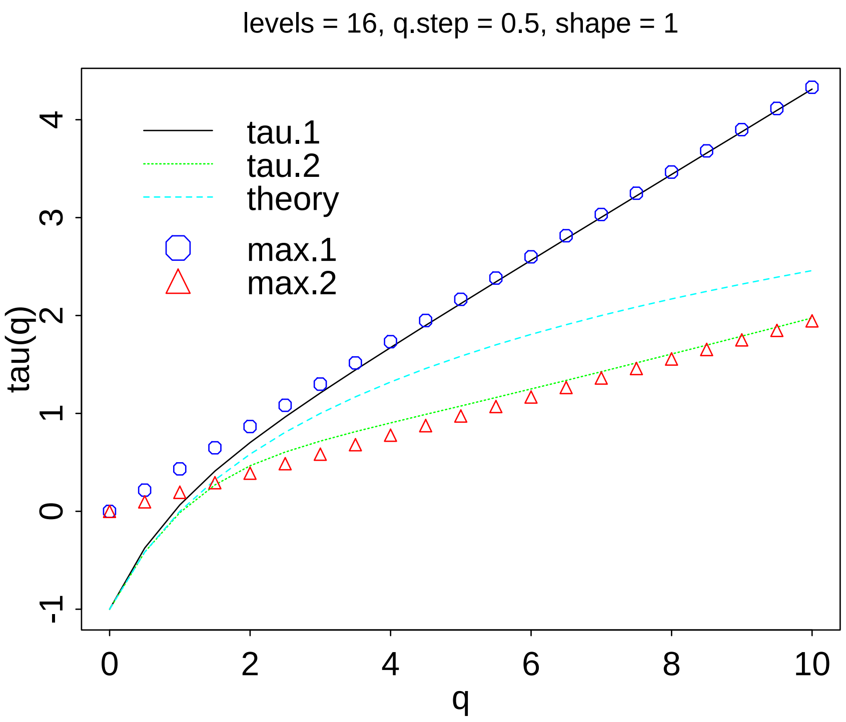

FIGURE 3. Plots of $\hat{\tau}_i(q), \hat{\tau}_i^{\vee}(q)$ for $q_* \approx 2.4$.

properties of Ossiander and Waymire's time domain-based cascade estimators and our wavelet-based estimators hold. In the random cascade case, this critical range is shown in [13] to be an interval of the form $[Q_c^-, Q_c^+]$ with

$$Q_c^+ = \sup\{q \geq 1 \,|\, q\chi_b'(q) - \chi_b(q) < 0\} \quad \text{and}$$
$$Q_c^- = \sup\{q \leq 0 \,|\, q\chi_b'(q) - \chi_b(q) < 0\}.$$

On the other hand, in the case of the wavelet-based estimators for the conservative cascade processes, we have shown that his critical range is of the form $[0, q_*]$ where q_* satisfies

$$q_* = \sup_{q>0}\{\chi_2(2q) - 2\chi_2(q) < 0\};$$

equivalently,

$$q_* = \sup_{q>0}\{q\tilde{\chi}_2(2q) - \chi_2(q) < 0\}$$

where $\tilde{\chi}_2(2q) = (\chi_2(2q) - \chi_2(q))/q$ is the secant of χ_2 at $2q$. Note that with the stipulation on the generator W that $\chi'(1-) < 0$, χ_2 is a convex function and hence, the critical range for the wavelet-based estimator (for conservative cascades) is in general smaller than the positive half of the critical range for the cascade estimator (for random cascades). For example, if the generator W for the random cascade is uniform on $[0, 1]$, we get $Q_c^+ \approx 3.3$; to compare, taking the same W as generator for a conservative cascade, we have from Example 1 in Section 4 that $q_* \approx 2.4$. To enable a direct comparison of the estimators considered in this paper and in [13], we plan to investigate wavelet-based estimation for random cascades. At the same time, it would be of interest to see whether or not the techniques used in [13] yield comparable results to the ones presented in this paper when they are applied to the conservative cascade setting with its subtle dependence structure (see the comment in [13, page 3].

References

[1] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[2] A. Feldmann, A. C. Gilbert, and W. Willinger. Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. In *Proc. of the ACM/SIGCOMM'98*, pages 25–38, Vancouver, B.C., 1998.

[3] U. Frisch and G. Parisi. Fully developed turbulence and intermittancy. In M. Ghil, editor, *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics.* North-Holland, Amsterdam, 1985.

[4] A.C. Gilbert, W. Willinger, and A. Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):971–991, 1999.

[5] V. K. Gupta and E. C. Waymire. A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32:251–267, 1993.

[6] Richard Holley and Edward C. Waymire. Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Ann. Appl. Probab.*, 2(4):819–845, 1992.

[7] A.N. Kolmogorov. Local structure of turbulence in an incompressible liquid for very large Reynolds numbers. *C.R. Doklady Acad. Sci. URSS (N.S.)*, 30:299–303, 1941.

[8] B. B. Mandelbrot. Intermittant turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics*, 62:331–358, 1974.

[9] B. B. Mandelbrot. Limit Lognormal Multifractal Measures. In Gotsman, Ne'eman, and Voronel, editors, *The Landau Memorial Conference*, pages 309–340, Tel Aviv, 1990.

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

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

[12] J. Neveu. *Discrete-Parameter Martingales*, volume 10 of *North-Holland Mathematical Library.* North Holland, Amsterdam, 1975. Translated from the French original by T.P. Speed.

[13] M. Ossiander and E. Waymire. Statistical estimation for multiplicative cascades. To appear; available from {ossiand,waymire}@math.orst.edu, 1999.

[14] S.I. Resnick. *A Probability Path.* Birkhäuser, Boston, 1998.

[15] R. H. Riedi and J. Levy-Vehel. Tcp traffic is multifractal: A numerical study. Preprint, 1997.

[16] B. Troutman and A.V. Vecchia. Estimation of Renyi exponents in random cascades. *Bernoulli*, 1999. To appear.

School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853
E-mail address: sid@orie.cornell.edu

AT&T Labs–Research, 180 Park Avenue, Florham Park, NJ 07932
E-mail address: agilbert@research.att.com, walter@research.att.com