CHAPTER 9

MULTIPLICATIVE MULTIFRACTALS

We have discussed $1/f^\beta$ processes and Levy motions. Now we will study the third type of random fractal model, the random cascade multifractal. This type of multifractal was initially developed to understand the intermittent features of turbulence. Mandelbrot was among the first to introduce this concept. Parisi and Frisch's work has made it widely known. It has been applied to the study of various phenomena, such as rainfall and liquid water distributions inside marine stratocumuli. It has also been applied to the study of finance by Mandelbrot and to network traffic modeling by numerous researchers.

Below, we first present the definition of measure-based multifractals and explain a simple construction procedure for generating a random cascade multifractal. After presenting examples and proving a number of properties of such fractal processes, we explain how the model can be used to understand turbulent motions, to analyze sea clutter and neuronal firing patterns, and to model network traffic.

9.1 DEFINITION

Consider a unit interval. Associate it with a unit mass. Partition the unit interval into a series of small intervals, each of linear length ϵ. Also, partition the unit mass into a series of weights or probabilities $\{w_i\}$ and associate w_i with the ith interval.

Now consider the moments

$$M_q(\epsilon) = \sum_i w_i^q, \tag{9.1}$$

where q is real. We follow the convention that whenever w_j is zero, the term w_j^q is set to zero as well. We also note that a positive q value emphasizes large weights, while a negative q value emphasizes small weights. If we have, for a real function $\tau(q)$ of q,

$$M_q(\epsilon) \sim \epsilon^{\tau(q)}, \quad as \quad \epsilon \to 0 \tag{9.2}$$

for every q, and the weights $\{w_i\}$ are nonuniform, then the weights $w_i(\epsilon)$ are said to form a multifractal measure. Note that the normalization $\sum_i w_i = 1$ implies that $\tau(1) = 0$.

Note that if $\{w_i\}$ are uniform, then $\tau(q)$ is linear in q. When $\{w_i\}$ are weakly nonuniform, visually $\tau(q)$ may still be approximately linear in q. The nonuniformity in $\{w_i\}$ is better characterized by the so-called generalized dimensions D_q, defined as

$$D_q = \frac{\tau(q)}{q-1}. \tag{9.3}$$

D_q is a monotonically decreasing function of q. It exhibits a nontrivial dependence on q when the weights $\{w_i\}$ are nonuniform.

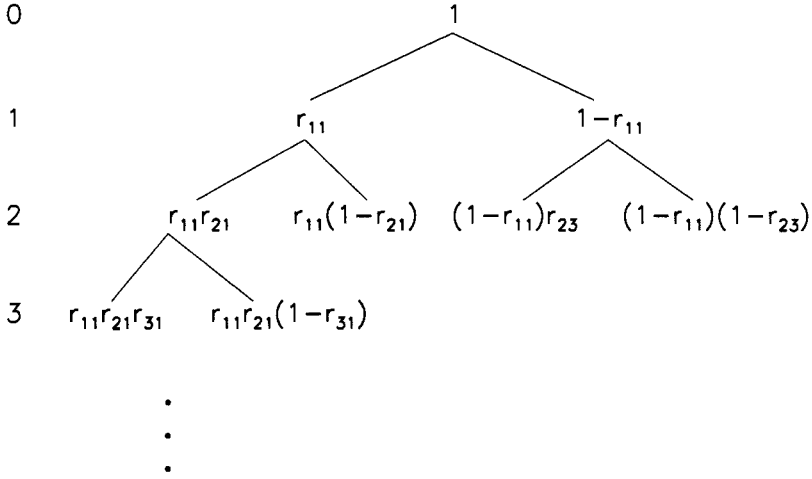## 9.2 CONSTRUCTION OF MULTIPLICATIVE MULTIFRACTALS

To better appreciate the construction rules, we point out that these rules essentially involve dyadic partitions.

Consider a unit interval. Associate it with a unit mass. Divide the unit interval into two, say, left and right segments of equal length. Also, partition the associated mass into two fractions, r and $1 - r$, and assign them to the left and right segments, respectively. The parameter r is in general a random variable, governed by a PDF $P(r), 0 \le r \le 1$. The fraction r is called the multiplier. Each new subinterval and its associated weight are further divided into two parts following the same rule. This procedure is schematically shown in Fig. 9.1, where the multiplier r is written as r_{ij}, with i indicating the stage number and j (assuming only odd numbers, leaving even numbers for $1 - r_{ij}$) indicating the positions of a weight on that stage. Note that the scale (i.e., the interval length) associated with stage i is 2^{-i}. We assume that $P(r)$ is symmetric about $r = 1/2$ and has successive moments $\mu_1, \mu_2, \cdots$. Hence r_{ij} and $1 - r_{ij}$ both have marginal distribution $P(r)$. The weights at the stage N, $\{w_n, n = 1, ..., 2^N\}$, can be expressed as

$$w_n = u_1 u_2 \cdots u_N,$$

where $u_l, l = 1, \cdots, N$, are either r_{ij} or $1 - r_{ij}$. Thus, $\{u_i, i \ge 1\}$ are iid random variables having PDF $P(r)$.

Stage

Figure 9.1.   Schematic illustrating the construction rule of a multiplicative multifractal.

Note that the rules described above generate a conservative process. To generate a nonconservative process, one can simply require that a mass m_i at stage i is split into two parts, $W_1 m_i$ and $W_2 m_i$, where W_i, $i = 1, 2$ are independent, follow the same PDF, and both have mean 1/2. Since the latter does not ensure that $W_1 + W_2 = 1$, the sum of the weights at stage n is not necessarily equal to 1.

In the following, we illustrate this process by selecting a specific pdf $P(r)$.

• **Deterministic binomial multiplicative process**

In this case, the PDF is set to be equal to $P(r) = \delta(r - p)$, where $\delta(x)$ is the Kronecker delta function. Thus, $r = p$ with probability 1, where $0 < p < 1$ is a fixed number. The weights obtained for the first several stages are schematically shown in Fig. 9.2.
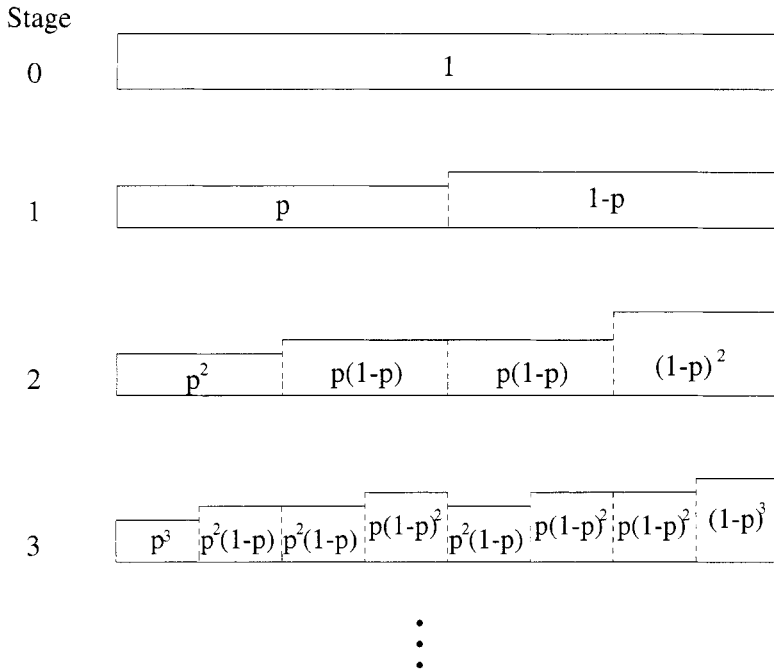
For this process, at stage n, we have

$$M_q(\epsilon) = \sum_{i=0}^{n} C_n^i p^{qi} (1 - p)^{q(n-i)} = [p^q + (1 - p)^q]^n. \qquad (9.4)$$

Since, at stage n, $\epsilon = 2^{-n}$, we obtain

$$\tau(q) = -\ln[p^q + (1 - p)^q]/\ln 2, \qquad (9.5)$$

which is independent of n (or ϵ). Hence, this weight process constitutes a multi-fractal.

It is interesting to note that a similar model has been proposed by meteorologists to describe the stochastic growth process of raindrops. Fig. 9.3 schematically shows that from one time step to another, 90% of raindrops remain their original size, while

Stage

0

1

1

p

$1-p$

2

p^2

$p(1-p)$

$p(1-p)$

$(1-p)^2$

3

p^3 $p^2(1-p)$ $p^2(1-p)$ $p(1-p)^2$ $p^2(1-p)$ $p(1-p)^2$ $p(1-p)^2$ $(1-p)^3$

Figure 9.2.   A schematic showing the weights at the first several stages of the binomial multiplicative process $(r = p$ w.p.1$)$.

10% of them increase their size by a fixed ratio. The physics behind this picture is that a raindrop attracts the water vapor in the surrounding air in an amount proportional to its surface area with probability 0.1 but fails to attract any water vapor with probability 0.9. Although simple, this model is able to account for the observed rapid growth of raindrops.

• **Random binomial multiplicative process**
   To make the weight series random, we modify $P(r)$ to become

$$P(r) = [\delta(r - p) + \delta(r - (1 - p))]/2 \qquad (9.6)$$

so that $P(r = p) = P(r = 1-p) = 1/2$. Hence, $P(r)$ is symmetric about $r = 1/2$. A realization of the weights at stage 12 (with $p = 0.3$) is shown in Fig. 9.4(a). It is quite obvious that the $\tau(q)$ spectrum for this process is identical to that for the deterministic binomial process, since the weight sequence for this process is simply a shuffled version of that for the deterministic case.

• **Random multiplicative process**
   The function $P(r)$ can be selected to follow any functional form. The following piecewise linear $P(r)$ function is used to generate the weight realization (at stage 12) shown in Fig. 9.4(b):
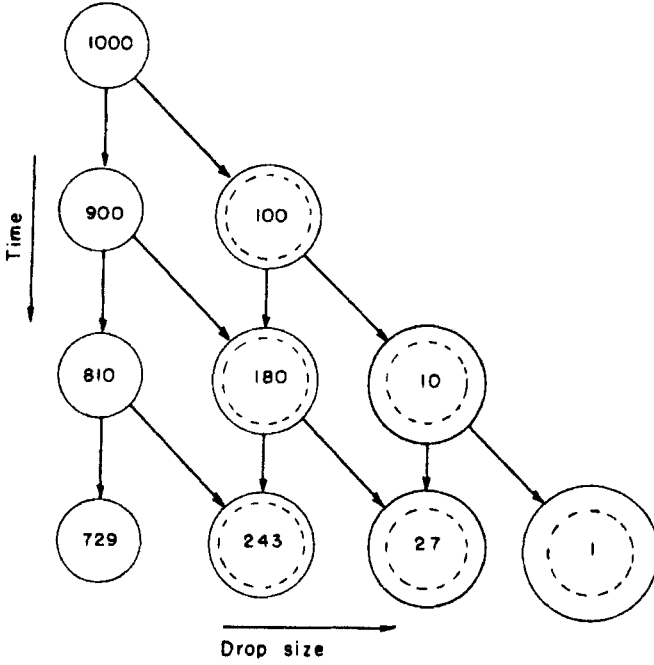
Figure 9.3.   Schematic of raindrop growth. After [47, 224].

$$P(r) = \begin{cases} 2r + 0.5 & \text{if } 0 \le r \le 0.5 \\ -2r + 2.5 & \text{if } 0.5 \le r \le 1. \end{cases} \tag{9.7}$$

9.3   PROPERTIES OF MULTIPLICATIVE MULTIFRACTALS

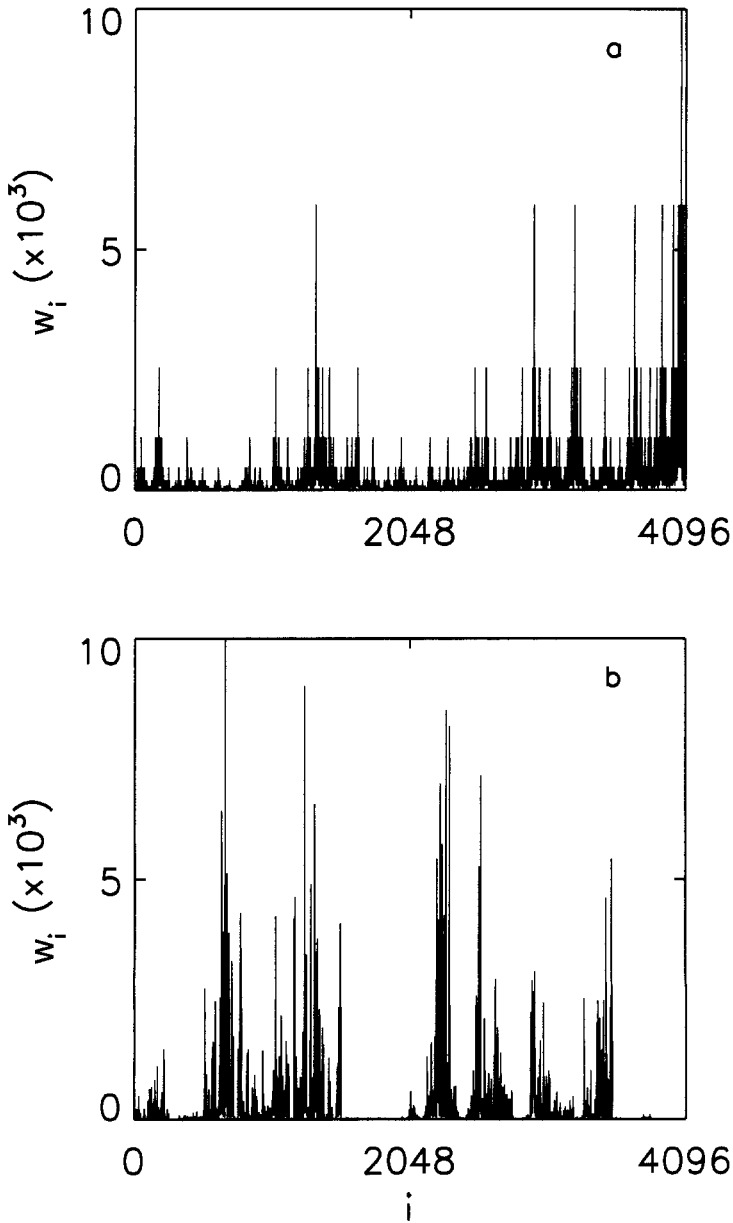For the weights of a conservative cascade model at stage N, we prove the following properties.

(1)
$$M_q(\epsilon) \sim \epsilon^{\tau(q)}, \tag{9.8}$$

with $\epsilon = 2^{-N}$ and
$$\tau(q) = -\ln(2\mu_q)/\ln 2, \tag{9.9}$$

where μ_q is the qth moment of the iid random variables u_i. This follows from the observation that at stage N,

$$M_q(\epsilon) = \sum_{n=1}^{2^N} (w_n)^q \approx 2^N E(w^q) = 2^N E((u_1 u_2 \cdots u_N)^q) = 2^N \mu_q^N.$$

Figure 9.4.   Weight series at stage 12 for (a) the modified binomial multiplicative process ($p = 0.3$) and (b) the random multiplicative process with the multiplier PDF given in Eq. (9.7). The scale of w_i is multiplied by 10^3, as indicated in the parentheses.

This property indicates that a multiplicative process is a multifractal and relates the $\tau(q)$ spectrum to the moments of the multiplier distribution.

Note that for $P(r) = [\delta(r-p) + \delta(r-(1-p))]/2$, we have $\mu_q = [p^q + (1-p)^q]/2$. Hence, $\tau(q) = -\ln[p^q + (1-p)^q]/\ln 2$. This is identical to Eq. (9.5). However, Eq. (9.9) is much more general and, hence, more powerful.

(2)

$$E(w) = E(w_n) = E(u_1 u_2 \cdots u_N) = 2^{-N}, \quad n = 1, \cdots, 2^N. \qquad (9.10)$$

(3)

$$E(w^q) = E((u_1 u_2 \cdots u_N)^q) = \mu_q^N. \qquad (9.11)$$

In particular,

$$E(w^2) = \mu_2^N \qquad (9.12)$$

and

$$Var(w) = Var(w_n) = \mu_2^N - 2^{-2N}, \quad n = 1, \cdots, 2^N. \qquad (9.13)$$

(4) When $N \gg 1$, the weights at stage N have a log-normal distribution. This is deduced directly by taking the logarithm of $w_n = u_1 u_2 \cdots u_N$ and applying the central limit theorem.

Note that log-normality is a salient feature that has been observed extensively in many different types of network traffic processes, including VBR video traffic, the call duration and dwell time in a wireless network, data connection and messages, and page size, page request frequency, and user's think time in WWW traffic. Hence this property can be used to determine if measured traffic trace data are consistent with a multiplicative multifractal process model.

(5)

$$E[(w_n - E(w))(w_{n+m} - E(w))] = (1/2 - \mu_2)\mu_2^{N-1}(4\mu_2)^{-k} - 2^{-2N} \qquad (9.14)$$

for $m = 2^k$, where k is an integer. Hence, the covariance function decays with time lag m in a power-law manner.

To prove assertion (5), we consider two weights, w_{n_1} and w_{n_2}, at stage N. Assume that they share the same ancestor weight x at stage $N - k$, i.e.,

$$w_{n_1} = x \cdot r \cdot \prod_{l=1}^{k-1} r_{1l},$$

$$w_{n_2} = x \cdot (1 - r) \cdot \prod_{l=1}^{k-1} r_{2l},$$

where r and $\{r_{il}, i = 1, 2, l = 1, ..., k-1\}$ are independent random variables with distribution $P(r)$. Then

$$
\begin{aligned}
E[(w_{n_1} - 2^{-N})(w_{n_2} - 2^{-N})] &= E(x^2)E[r(1-r)]E[\prod_{l=1}^{k-1} r_{1l}r_{2l}] - 2^{-2N} \\
&= 2^{-2(k-1)}\mu_2^{N-k}(1/2 - \mu_2) - 2^{-2N}.
\end{aligned}
$$

For $m = 2^k$, all pairs of $\{w_n, w_{n+m}\}$, for $n \geq 1$, share an ancestor at stage $N - k - 1$. Hence,

$$
\begin{aligned}
E[(w_n - E(w))(w_{n+m} - E(w))] &= 2^{-2k}\mu_2^{N-k-1}(1/2 - \mu_2) - 2^{-2N} \\
&= (1/2 - \mu_2)\mu_2^{N-1}(4\mu_2)^{-k} - 2^{-2N}.
\end{aligned}
$$

(6)

$$
Var(W^{(m)}) = \mu_2^N(4\mu_2)^{-k} - 2^{-2N}, \tag{9.15}
$$

where

$$
W^{(m)} = (w_{im-m+1} + \cdots + w_{im})/m, \ m = 2^k, \ k = 1, 2, \cdots, \text{ and } i \geq 1.
$$

This is proved by expressing $W^{(m)} = 2^{-k}x$, where x is a weight at stage $N - k$.

Equation (9.15) expresses a variance-time relation. For time series with long memory, $Var(W^{(m)}) \sim m^{2H-2}$, where $1/2 < H < 1$ is the Hurst parameter. For multiplicative multifractal processes, when N is large and $\mu_2 > 0$, the term $\mu_2^N(4\mu_2)^{-k}$ dominates. When the term 2^{-2N} in Eq. (9.15) is dropped, the functional variation of $\log Var(W^{(m)})$ vs. $\log m$ is linear. The resulting slope, $-\log(4\mu_2)/\log 2$, provides an estimate of $2H - 2$. A moment's thought will convince us that this slope is an upper bound for $2H - 2$. Hence,

$$
H \leq -\frac{1}{2}\log_2 \mu_2. \tag{9.16}
$$

Since the multiplier distribution $P(r)$ is defined for $0 \leq r \leq 1$ and is symmetric about $1/2$, its mean is $1/2$ and its variance is upper bounded by $1/4$. We thus have

$$
\left(\frac{1}{2}\right)^2 \leq \mu_2 \leq \left(\frac{1}{2}\right)^2 + \frac{1}{4}.
$$

Therefore $1/2 \leq H \leq 1$, with $H = 1$ corresponding to deterministic time series (i.e., $P(r) = \delta(r - 1/2)$). We thus observe that a multiplicative multifractal process also possesses a long-range-dependent property, especially for time scales not too close to the total time span of observation.

Let us see how good the linearity defined by Eq. (9.15) is. For this purpose, we consider three different functional forms for the multiplier function, namely, double exponential with parameter α_e,

$$
P(r) \sim e^{-\alpha_e|r-1/2|} \tag{9.17}
$$
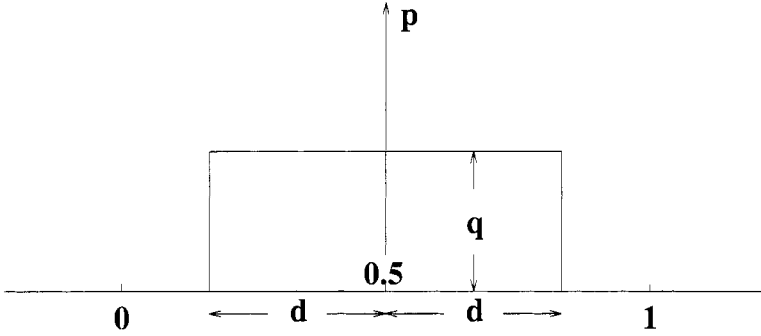
Figure 9.5.   A schematic showing the functional form for the multiplier distribution described by Eq. (9.19).

Gaussian with parameter α_g,

$$P(r) \sim e^{-\alpha_g(r-1/2)^2} \tag{9.18}$$

and a function of the form

$$P(r) = \begin{cases} q + p\delta(r - 1/2) & 1/2 - d \leq r \leq 1/2 + d \\ 0 & \text{otherwise,} \end{cases} \tag{9.19}$$

where $0 \leq d \leq 1/2$. The last function is schematically shown in Fig. 9.5. Note that the three parameters d, p, and q are related by the equation $p + 2qd = 1$. Hence, the function contains two independent parameters. We shall choose p and d as the two basic parameters. Note that one may introduce a parameter equivalent to d for the functions characterized by Eqs. (9.17) and (9.18). Due to exponential decay of Eqs. (9.17) and (9.18), however, such a parameter is not too interesting. Also note that we may rewrite Eqs. (9.17) and (9.18) as $P(r) \sim e^{-\alpha|r-1/2|^\beta}$, with $\beta = 1$ for the double exponential and 2 for the Gaussian. Hence, Eqs. (9.17) and (9.18) really contain two parameters with a prefixed parameter β.

We generate a number of realizations of multiplicative processes with $P(r)$ given by one of the above three forms. Two examples are shown in Fig. 9.6. We then compute the variance-time relation from the generated time series. Some examples of the variance-time curves are shown in Fig. 9.7, with $P(r)$ given by Eq. (9.18) and $N = 18$. We observe that the variance-time curves are indeed approximately linear, with the linearity better defined for larger μ_2.

We can furthermore check how tight the bound determined by Inequality (9.16) is by estimating H from the variance-time curves and comparing it with the right-hand side of Inequality (9.16). Figure 9.8 shows such a comparison, where the dashed curve is generated from Eq. $H = -\frac{1}{2}\log_2 \mu_2$. The points denoted by the diamonds, triangles, and squares are estimated from multiplicative processes with $P(r)$ given by Eqs. (9.17), (9.18), and (9.19), respectively. We observe that the bound given by Inequality (9.16) is very tight, especially for not too small values of μ_2.
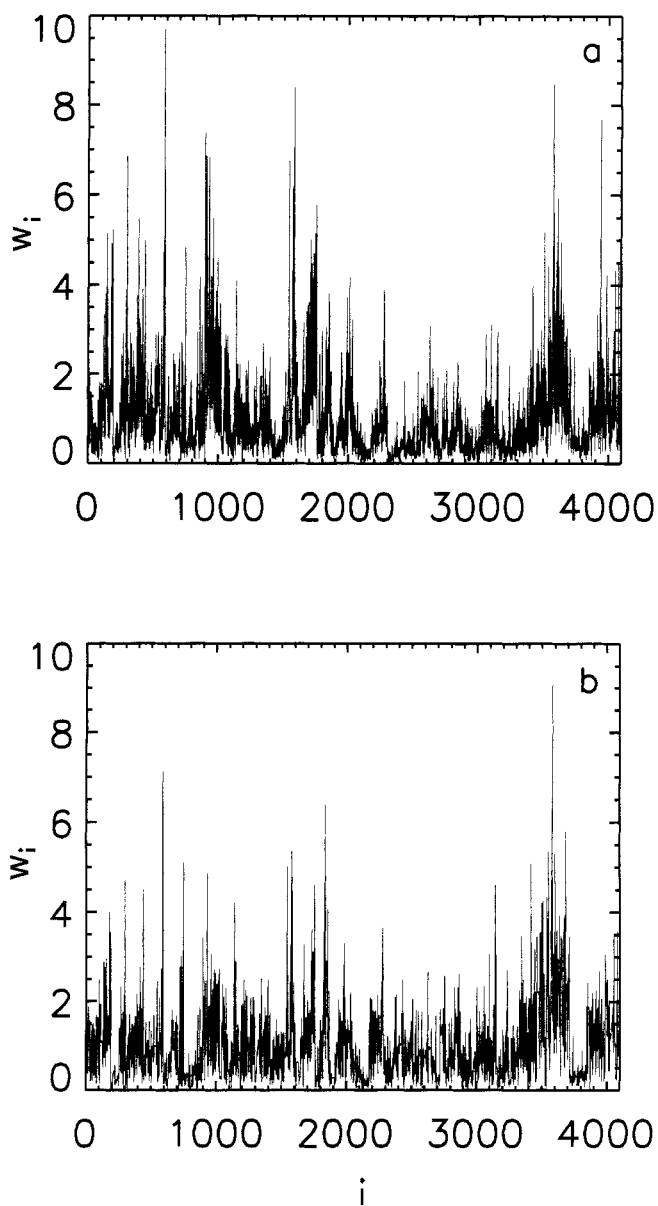
Figure 9.6.   Weight series at stage 12 with the multiplier distributions given by (a) Eq. (9.18) with $\alpha_g = 30$ and (b) Eq. (9.19) with $(p, d) = (0.7, 0.4)$.
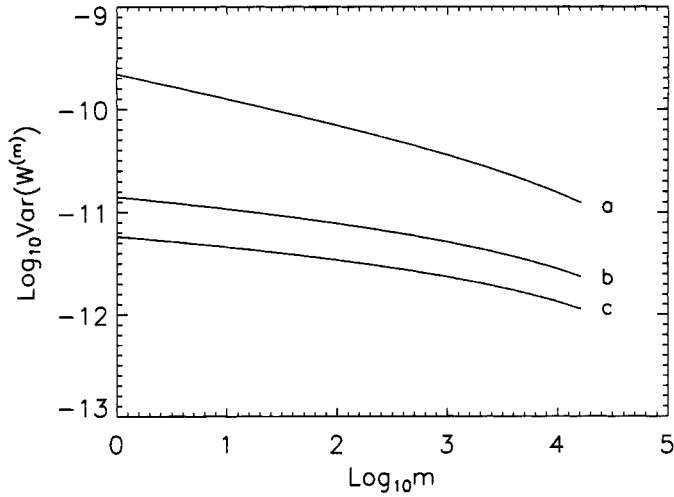
Figure 9.7. $\log Var(W^{(m)})$ vs. $\log m$ curves for (a) $\alpha_g = 10$, (b) $\alpha_g = 50$, and (c) $\alpha_g = 100$.
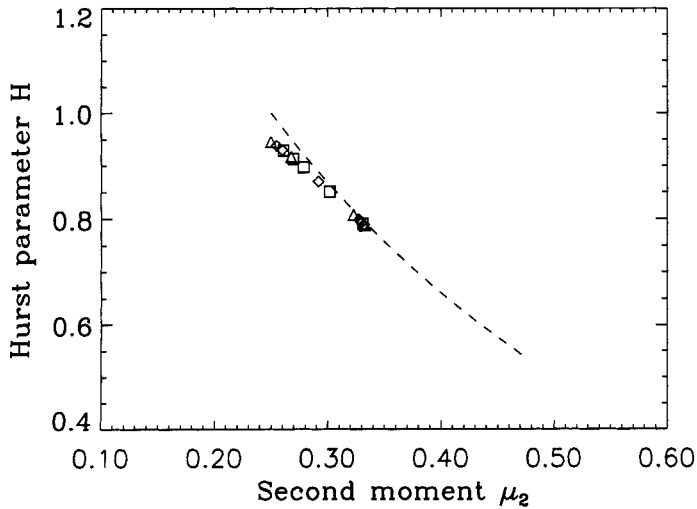
Figure 9.8.   Hurst parameter vs. the second moment μ_2. The dashed line is computed according to the right-hand side of Inequality (9.16), while the points designated by squares, diamonds and triangles are directly estimated from multiplicative processes with their multiplier distributions given by Eqs. (9.17), (9.18), and (9.19), respectively.

9.4   INTERMITTENCY IN FULLY DEVELOPED TURBULENCE

Understanding turbulent flows is one of the greatest challenges in physics. An intriguing aspect of fully developed turbulence is the possible existence of universal scaling behavior of small-scale fluctuations. In large Reynolds number flows, at

spatial scales l, $\eta \leq l \leq l_0$ (where η is called the molecular dissipation scale, l the inertial-range scale, and l_0 the energy-injection scale), turbulent fluctuations reach a statistically quasi-equilibrium state characterized by a continuous energy flux from large to small scales. The universal scaling behavior refers to the observation that moments of fluctuation at scale l have a power-law dependence on l, and the scaling exponents are universal. Two fluctuating quantities are of special interest: the energy dissipation ϵ_l (per unit mass) averaged over a ball of size l and the longitudinal velocity differences δv_l across a distance l. The scaling behavior of δv_l and ϵ_l is expressed as

$$\langle \delta v_l^q \rangle \sim l^{\zeta_q}, \tag{9.20}$$

$$\langle \epsilon_l^q \rangle \sim l^{\tau_q}. \tag{9.21}$$

Comparing Eq. (9.20) with Eq. (8.13), we see that the velocity is treated as a random walk process and the velocity difference as a noise process. Since ϵ_l has the dimension $L^2 T^{-3}$, where L denotes length and T time, $(l\epsilon_l)^{q/3}$ and δv_l^q have the same dimension. Therefore,

$$\langle \delta v_l^q \rangle \sim \langle \epsilon_l^{q/3} \rangle l^{q/3}. \tag{9.22}$$

One thus has the bridge relation between τ_q and ζ_q:

$$\zeta_q = q/3 + \tau_{q/3}. \tag{9.23}$$

In 1941, Kolmogorov conjectured that

$$\zeta_q = q/3. \tag{9.24}$$

This theory will hereafter be called K41. From this, one sees that $H(q) = \zeta_q/q = 1/3$. Since $H(q)$ does not depend on q, it is a monofractal model for turbulence. The PSD for the velocity difference, called the energy spectrum of turbulence, decays with the wavenumber k as $k^{-(2H+1)} = k^{-5/3}$ (wavenumber is the spatial counterpart of frequency in time). The 5/3 energy spectrum is one of the most famous results in turbulence that have been observed in numerous experimental studies. However, Eq. (9.24) for arbitrary q is not supported by experiments and numerical simulations. See the sixth column of Table 9.1. This is called the intermittency phenomenon of turbulence. Many models have been proposed to explain intermittency. Random cascade multifractal models are among the most successful. The basic idea is

$$\epsilon_{l_2} = W_{l_1 l_2} \epsilon_{l_1}, \tag{9.25}$$

where $W_{l_1 l_2}$ is the cascade multiplicative factor for any arbitrary pair of length scales l_1, l_2, $l_1 > l_2$. Energy conservation requires that

$$\langle W_{l_1 l_2} \rangle = 1. \tag{9.26}$$

Order q	Experiment ζ_q	Logstable ζ_q	P model ζ_q	SL model ζ_q	K41 $\zeta_q = q/3$
1	0.37	0.362	0.361	0.364	0.333
2	0.70	0.694	0.694	0.696	0.667
3	1.00	1.000	1.000	1.000	1.000
4	1.28	1.285	1.282	1.279	1.333
5	1.54	1.551	1.543	1.538	1.667
6	1.78	1.800	1.786	1.778	2.000
7	2.00	2.032	2.014	2.001	2.333
8	2.23	2.249	2.229	2.211	2.667
10	2.60	2.638	2.632	2.593	3.333

Table 9.1   Comparison of the experimentally measured scaling exponents ζ_q and the theoretical predictions

Unlike conservative cascade models discussed earlier, this is a nonconservative model. Using Eq. (9.20), one has $\epsilon_{l_i} \sim l_i^{\tau_q}$, $i = 1, 2$. Therefore,

$$\ln\langle W_{l_1 l_2}^q \rangle = \tau_q(\ln l_2 - \ln l_1). \tag{9.27}$$

Below we discuss the concept of extended self-similarity, which facilitates experimental determination of ζ_q, and describe a few of the better-known models.

9.4.1   Extended self-similarity

To observe fully developed turbulence, the Reynolds number, $R = LV/\nu$, where L, V are the characteristic scale and velocity of the flow and ν is its (kinetic) viscosity, has to be very large. When R is not large, the inertial range is very short, and the scaling behavior described by Eq. (9.21) is either absent or difficult to observe. A few examples are shown at the top of Fig. 9.9. This hinders accurate experimental determination of ζ_q. The concept of extended self-similarity provides a nice solution to this problem. Benzi et al. [42] find that the scaling properties of the velocity increments can be extended up to the dissipation range in the form

$$\langle \delta v_l^q \rangle \sim \langle \delta v_l^3 \rangle^{\zeta_q/\zeta_3}, \quad l \geq 5\eta. \tag{9.28}$$

That is, if one plots $\langle \delta v_l^q \rangle$ against $\langle \delta v_l^3 \rangle$, then the scaling behavior is much better defined, and the scaling exponent is given by ζ_q/ζ_3. For an example of $q = 6$, see the bottom of Fig. 9.9. The fact that $\zeta_3 = 1$ makes this simple procedure even more attractive.
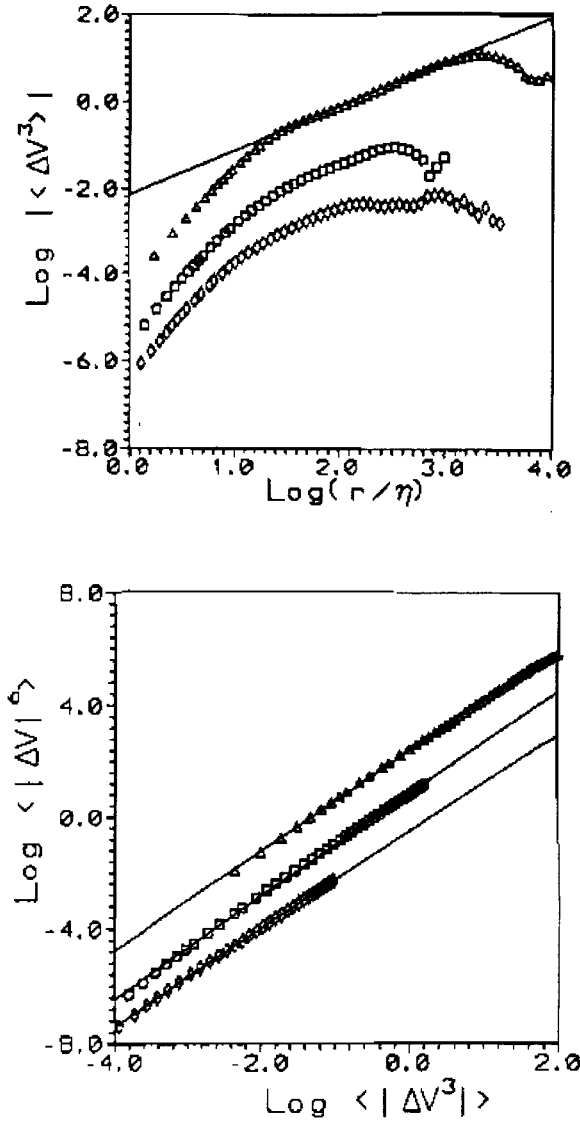
Figure 9.9.   Top: Log-log plot of $\langle \delta v_l^3 \rangle$ against r/η for flows with Reynolds number 300,000 (triangles), 18,000 (squares), and 6000 (diamonds). The solid line correspond to a slope equal to 1. Bottom: Log-log plot of $\langle |\delta v_l^6| \rangle$ against $\langle |\delta v_l^3| \rangle$. Note that when $\log \langle |\delta v_l^3| \rangle$ is plotted against $\log |\langle \delta v_l^3 \rangle|$, one finds an excellent linear scaling relation with a slope very close to 1. From Benzi et al. [42].

9.4.2 The log-normal model

This is the first and perhaps the most famous (but not necessarily the best) model. The idea is that the scales between l_1 and l_2 can be further partitioned so that $W_{l_1 l_2}$ can be expressed as multiplication of a sequence of iid random variables. Taking the log, then, $\ln W_{l_1 l_2}$ becomes a summation of a sequence of iid random variables. When the central limit theorem is applicable, $W_{l_1 l_2}$ follows the log-normal distribution. More concretely, let $\delta = l_2/l_1$; then

$$W_{l_1 l_2} \stackrel{d}{=} \delta^X, \tag{9.29}$$

where X is $N(\mu_x, \sigma^2)$.

Let us now compute $\langle W^q_{l_1 l_2} \rangle$. It is simply

$$\langle \delta^{qX} \rangle = \langle e^{q \ln \delta X} \rangle.$$

The right-hand side of the equation can be obtained by replacing ju in the characteristic function of $N(\mu_x, \sigma^2)$ (see Eq. (3.18)) by $q \ln \delta$. Therefore,

$$\langle W^q_{l_1 l_2} \rangle = e^{q \ln \delta \mu_x + \sigma^2 q^2 (\ln \delta)^2 / 2} = \delta^{q \mu_x + \sigma^2 q^2 \ln \delta / 2}.$$

When $q = 1$, $\langle W_{l_1 l_2} \rangle = 1$. Therefore, $\mu_x + \sigma^2 \ln \delta / 2 = 0$. Using Eq. (9.27), we then have $\tau_q = \mu_x(q - q^2)$. It is common to write $\mu_x = \mu/2$. Hence,

$$\tau_q = \frac{\mu}{2}(q - q^2), \quad \zeta_q = q/3 + \frac{\mu}{2}[q/3 - (q/3)^2]. \tag{9.30}$$

A good value for μ is 0.3288.

There are a number of problems with the log-normal model. Two are pointed out here: (1) ζ_q is a decreasing function of q when $q > 3/2 + 3/\mu$. Frisch has proven that this causes velocity to diverge (i.e., supersonic flow); and (2) The log-normal model violates Novikov's inequality, which states that, for three dimensions,

$$\tau_q + 3q \geq 0, \text{ for } q \geq 0 \quad \text{and} \quad \tau_q + 3q \leq 0, \text{ for } q \leq 0$$

and for one dimension,

$$\tau_q + q \geq 0, \text{ for } q \geq 0 \quad \text{and} \quad \tau_q + q \leq 0, \text{ for } q \leq 0.$$

Novikov's inequality is a constraint on energy: $l^3 \epsilon_l$ (in three dimensions) or $l \epsilon_l$ (in one dimension) is the total unaveraged energy and, hence, should be a nondecreasing function of l. When either one is raised to the power of q, it should be a nondecreasing function of l when $q \geq 0$ and a nonincreasing function of l when $q \leq 0$.

The log-normal model and many other random cascade models violate Novikov's inequality because of the nonconservative character of the cascade.

9.4.3 The log-stable model

This model applies the generalized central limit theorem to the random variable X in Eq. (9.29). $\langle W_{l_1 l_2}^q \rangle$ can again be simply obtained by replacing ju in the characteristic function of X (see Eq. 7.6) by $q \ln \delta$. Upon simplification, it becomes $\delta^{bq - \beta q^\alpha}$, where β is a complicated coefficient containing the term $|\ln \delta|^\alpha / \ln \delta$. When the condition $\langle W_{l_1 l_2} \rangle = 1$ is used, β is found to be b. Therefore, $\tau_q = b(q - q^\alpha)$. Using Kida's notation, $b = \mu/(2^\alpha - 2)$, we have

$$\tau_q = -\mu \frac{q^\alpha - q}{2^\alpha - 2},$$

$$\zeta_q = q/3 - \mu \frac{(q/3)^\alpha - q/3}{2^\alpha - 2}. \tag{9.31}$$

Kida found that the model gives an excellent fit to experimental data when $\alpha = 1.65$, $\mu = 0.2$. See the third column of Table 9.1. It is clear that the log-normal model is a special case of this model, and, as expected, the log-stable model has the same two problems as the log-normal model discussed in the previous subsection.

9.4.4 The β-model

The β-model is one of the earliest multifractal models proposed to describe intermittency in turbulence. Let a mother eddy at stage i generate N daughter eddies at stage $i + 1$. The size of the mother eddy is $l_i = r^i l_0$, while the size of the daughter eddies is $l_{i+1} = r l_i = r^{i+1} l_0$. It is clear that the fractal dimension of this cascade is

$$D = -\ln N / \ln r. \tag{9.32}$$

Let β denote the space-filling factor, $1 > \beta = Nr^3$. Then

$$3 - D = \ln \beta / \ln r. \tag{9.33}$$

In the β-model, the random cascade factor between scales $i + 1$ and i, denoted by $W = \epsilon_{i+1}/\epsilon_i$, takes only two values, $1/\beta$ with probability β and 0 with probability $1 - \beta$. Using Eq. (9.27), we have

$$\tau_q = \ln(\beta^{1-q}) / \ln r. \tag{9.34}$$

9.4.5 The random β-model

In the random β-model, the space-filling factor β takes two possible values, β_1 and β_2, with probabilities x and $1 - x$, respectively. Therefore, the model is described by the following:

$$W = \begin{cases} 0 & \text{with probability } 1 - \beta_1 x - \beta_2(1 - x), \\ 1/\beta_1 & \text{with probability } \beta_1 x, \\ 1/\beta_2 & \text{with probability } \beta_2(1 - x). \end{cases} \tag{9.35}$$

Using Eq. (9.27), we have

$$\tau_q = \ln \left[x\beta_1^{1-q} + (1 - x)\beta_2^{1-q} \right] \Big/ \ln r. \qquad (9.36)$$

9.4.6 The p model

The p model is a binomial cascade model. Suppose that an eddy of size l breaks down into 2^d eddies of size $l/2$, d being the dimensionality of the space we are analyzing. Furthermore, suppose that the flux of energy to these smaller eddies proceeds as follows: a fraction p is distributed equally among one half of the 2^d new eddies, and a fraction $1 - p$ is distributed similarly among the other half. At this point, the model is the same as the binomial random cascade model discussed earlier. To obtain the relation between $\epsilon_{l/2}$ and ϵ_l, we note that $\epsilon_l \propto E_l/l^d$, $\epsilon_{l/2} \propto E_{l/2}/(l/2)^d$. Therefore, $\epsilon_{l/2} = W\epsilon_l$, where $W = 2p$ and $2(1 - p)$, both with probability $1/2$. Using Eq. (9.27), we have

$$\tau_q = 1 - q - \log_2[p^q + (1 - p)^q]. \qquad (9.37)$$

Using Eq. (9.23), we have

$$\zeta_q = 1 - \log_2 \left[p^{q/3} + (1 - p)^{q/3} \right]. \qquad (9.38)$$

Note that when $p = 1/2$, the model gives the K41 theory (Eq. (9.24)). When $p = 0.7$, Meneveau and Sreenivasan [304] found that the model fit their experimental data very well. In fact, the model with the same p also fits other experimental data amazingly well. For an example, see Table 9.1, fourth column.

9.4.7 The SL model and log-Poisson statistics of turbulence

The SL model, proposed by She and Leveque [399], is one of the most remarkable models for intermittency in turbulence now available. With a few physically very appealing assumptions, it predicts that

$$\tau_q = -2q/3 + 2[1 - (2/3)^q]. \qquad (9.39)$$

Using Eq. (9.23), one finds

$$\zeta_q = q/9 + 2 \left[1 - (2/3)^{q/3} \right]. \qquad (9.40)$$

The prediction fits experimental data amazingly well. See Table 9.1, fifth column. We show that Eq. (9.39) can be derived by assuming log-Poisson statistics for the cascade multiplier factor $W_{l_1 l_2}$.

To see the idea, let the range of the scales between l_1 and l_2 be further divided into n subranges, $l_1, \delta l_1, \cdots, l_2 = \delta^n l_1$, $n \to \infty$. From scale $\delta^i l_1$ to $\delta^{i+1} l_1$, the cascade multiplier factor W_δ takes two values. The first is $\alpha_\delta = \delta^{-\gamma} > 1$,

$\gamma = 0(1)$, characterizing the rate of amplification of the dissipation event ϵ_{l_1}. This event, assumed to be positive, is called a singularity structure event. The most singular structures in turbulence are related to filaments. Filaments exist in localized, rare, nondissipative regions in the flow. See Fig. 9.10. The second value is $\alpha_\delta\beta$ with $\beta < 1$ and characterizes modulation of the singular structure by the factor β. This event is called a modulation-defect event. In the limit, $\delta \to 1$, $\alpha_\delta \to 1$. In order to make $\langle W_\delta \rangle = 1$, the probability of observing the modulation-defect event must go to zero. At this point, the model is similar to the random-β model, taking certain limiting processes. It is clear that over a finite separation of scales l_1 to l_2, the probability of observing $X = m$ modulation defects obeys a Poisson distribution (recall that the binomial distribution, with the probability of success being x, approaches the normal distribution if x stays finite and the number of trials n tends to infinity. However, if $nx \to \lambda$ when $n \to \infty$, then the binomial distribution approaches a Poisson distribution of expectation λ).

Figure 9.10.   Intermittent vortex filaments in a three-dimensional turbulent fluid simulated on a computer (She et al. [398].)

Now $W_{l_1 l_2}$ can be written as

$$W_{l_1 l_2} = (\beta \alpha_\delta)^X \alpha_\delta^{n-X} = (l_1/l_2)^\gamma \beta^X, \qquad (9.41)$$

where γ and β are constants and X is a Poisson random variable with mean $\lambda_{l_1 l_2}$. Since $\langle W_{l_1 l_2} \rangle = 1$, $\lambda_{l_1 l_2} = -\gamma(\ln l_1/l_2)/(\beta - 1)$, the qth moment of β^X is given by

$$\langle (\beta^X)^q \rangle = \sum_{m=1}^\infty \beta^{qm} e^{-\lambda_{l_1 l_2}} \frac{\lambda_{l_1 l_2}^m}{m!} = e^{-\lambda_{l_1 l_2}} e^{\lambda_{l_1 l_2} \beta^q}.$$

Therefore,

$$\ln\langle W_{l_1 l_2}^q \rangle = \gamma[q - (\beta^q - 1)/(\beta - 1)] \ln l_1/l_2. \qquad (9.42)$$

Taking $\gamma = \beta = 2/3$, one obtains Eq. (9.39).

9.5   APPLICATIONS

To illustrate the usefulness of cascade models for real-world signal processing applications, in this section we consider three difficult problems: target detection within sea clutter, analysis of neuronal firing patterns, and analysis and modeling of network traffic. For details of these data, we refer to Secs. A.1–A.3 of Appendix A.

For simplicity, we assume the length of the time series to be 2^N. When the original time series is not positive, we can transform it into a positive one by squaring it, or taking the absolute value of it, or adding a sufficiently large positive value to all elements of the time series. Denote the resulting time series by $\{X_i,\ i = 1, 2, \cdots, 2^N\}$. There are two ways to analyze $\{X_i\}$. One method is to compute the moments $M_q(\epsilon)$ at different stages and check whether Eq. (9.2) is valid for certain ϵ ranges. The other method is to compute the multiplier distributions at different stages and check whether they are stage independent. We note that the latter method is typically more useful when constructing a multiplicative process for the time series of interest. In the following, we describe in detail a general procedure for obtaining weight sequences at different stages needed for computing the moments $M_q(\epsilon)$ and the multiplier distributions.

The basic idea in analyzing $\{X_i\}$ is to view $\{X_i, i = 1, \cdots, 2^N\}$ as the weight series of a certain multiplicative process at stage N. Under this scenario, the total weight $\sum_{i=1}^{2^N} X_i$ is set equal to 1 unit, and the scale associated with stage N is $\epsilon = 2^{-N}$. This is the smallest time scale resolvable by the measured data.

Given the weight sequence at stage N, the weights at stage $N - 1$, $\{X_i^{(2^1)}, i = 1, \cdots, 2^{N-1}\}$, are obtained by simply adding the consecutive weights at stage N over nonoverlapping blocks of size 2, i.e., $X_i^{(2^1)} = X_{2i-1} + X_{2i}$, for $i = 1, \cdots, 2^{N-1}$, where the superscript 2^1 for $X_i^{(2^1)}$ is used to indicate that the block size used for the summation at stage $N - 1$ is 2^1. This follows directly from the construction of a multiplicative multifractal process schematized in Fig. 9.1.

Stage

$\vdots$

N - 3 $X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8$ $\cdots$

N - 2 $X_1 + X_2 + X_3 + X_4$ $X_5 + X_6 + X_7 + X_8$ $\cdots$

N - 1 $X_1 + X_2$ $X_3 + X_4$ $X_5 + X_6$ $X_7 + X_8$ $\cdots$

N X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 $\cdots$

Figure 9.11. A schematic showing the weights at the last several stages for the analysis procedure described in the text. It is instructive to compare this figure with Fig. 9.1, where it was shown that a weight at stage i is the sum of the two "daughter" weights at stages $i+1$.

Associated with this stage is the scale $\epsilon = 2^{-(N-1)}$. This procedure is carried out recursively. That is, given the weights at stage $j+1$, $\{X_i^{(2^{N-j-1})}, i = 1, \cdots, 2^{j+1}\}$, we obtain the weights at stage j, $\{X_i^{(2^{N-j})}, i = 1, \cdots, 2^j\}$, by adding consecutive weights at stage $j + 1$ over nonoverlapping blocks of size 2, i.e.,

$$X_i^{(2^{N-j})} = X_{2i-1}^{(2^{N-j-1})} + X_{2i}^{(2^{N-j-1})} \tag{9.43}$$

for $i = 1, \cdots, 2^j$. Here the superscript 2^{N-j} for $X_i^{(2^{N-j})}$ is used to indicate that the weights at stage j can be equivalently obtained by adding consecutive weights at stage N over nonoverlapping blocks of size 2^{N-j}. Associated with stage j is the scale $\epsilon = 2^{-j}$. This procedure stops at stage 0, where we have a single unit weight, $\sum_{i=1}^{2^N} X_i$, and $\epsilon = 2^0$. The latter is the largest time scale associated with the measured data. Figure 9.11 shows this procedure schematically.

After we have obtained all the weights from stages 0 to N, we compute the moments $M_q(\epsilon)$ according to Eq. (9.1) for different values of q. We then plot $\log M_q(\epsilon)$ vs. $\log \epsilon$ for different values of q. If these curves are linear over wide ranges of ϵ, then these weights are consistent with a multifractal measure. Note that, according to Eq. (9.2), the slopes of the linear part of $\log M_q(\epsilon)$ vs. $\log \epsilon$ curves provide an estimate of $\tau(q)$ for different values of q.

Next, we explain how to compute the multiplier distributions at different stages. From stage j to $j + 1$, the multipliers are defined by the following equation, based on Eq. (9.43):

$$r_i^{(j)} = \frac{X_{2i-1}^{(2^{N-j-1})}}{X_i^{(2^{N-j})}} \tag{9.44}$$

for $i = 1, \cdots, 2^j$. We view $\{r_i^{(j)}, i = 1, \cdots, 2^j\}$ as sampling points of the multiplier distribution at stage j, $P_j = \{P_j(r), 0 \le r \le 1\}$. Hence P_j can be determined

from its histogram based on $\{r_i^{(j)}, i = 1, \cdots, 2^j\}$. We then plot $P_j(r)$ vs. r for different stages (j) together. If these curves collapse together so that $P_j \sim P$, then the multiplier distributions are stage independent and the weights form a multifractal measure P.

9.5.1   Target detection within sea clutter

In Sec. 8.9.1, we discussed target detection within sea clutter using structure-function–based multifractal analysis. Here we show that the cascade model provides another simple and effective solution.

Since sea clutter amplitude data are positive, we can simply follow the procedure detailed at the beginning of this section to analyze the data. Figure 9.12 plots $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$ for the primary target bin and a bin without a target for two polarizations, HH and VV. We observe that the scaling between $M_q(\epsilon)$ and ϵ (i.e., the degree of linearity between $\log_2 M_q(\epsilon)$ and $-\log_2 \epsilon$) for all datasets is quite good up to the 15th stage. To further check whether these datasets are truly multifractals, we compute D_q for a certain ϵ range. The results are shown in Fig. 9.13. We observe that in all cases D_q has a nontrivial dependence on q. Therefore, we conclude that these time series are consistent with multifractals. Furthermore, when q is large, D_q can be used to effectively distinguish the data with and without a target.

9.5.2   Modeling and discrimination of human neuronal activity

Understanding neuronal firing patterns is one of the most important problems in theoretical neuroscience. It is also very important for clinical neurosurgery. Events in extracellular neuronal recording generate two types of time series: (1) the time interval between successive firings, called the interspike interval data, and (2) a counting process representing the number of firings in a chosen time period. In the past few decades, a lot of effort has been dedicated to the analysis of these two types of time series, using Fano factor analysis and wavelet analysis to characterize the difference between firing data and Poisson-based models, to quantify long-range-correlations, and to characterize neuronal dynamics.

Advances in surgical treatment of Parkinson's disease have stimulated much interest in the research of deep brain stimulation (DBS). The globus pallidus (GP) has been targeted in neurosurgical treatment of Parkinson's disease and dystonia. The GP is a complex structure in the basal ganglia and can be divided into two parts: the globus pallidus externa (GPe) and the globus pallidus interna (GPi). Both receive input from the caudate and putamen and communicate with the subthalamic nucleus. The GPi is thought to send the major inhibitory output from the basal ganglia back to the thalamus and also to send a few projections to an area of the midbrain (the PPPA), presumably to assist postural control. Distinguishing GPe from GPi is very important for surgical treatment of Parkinson's disease.
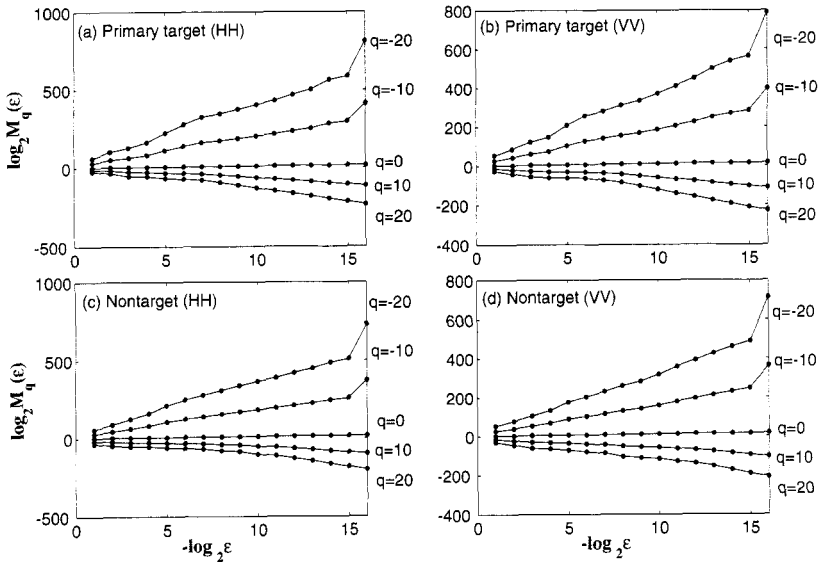
Figure 9.12. $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$ for sea clutter amplitude data for several different q's.
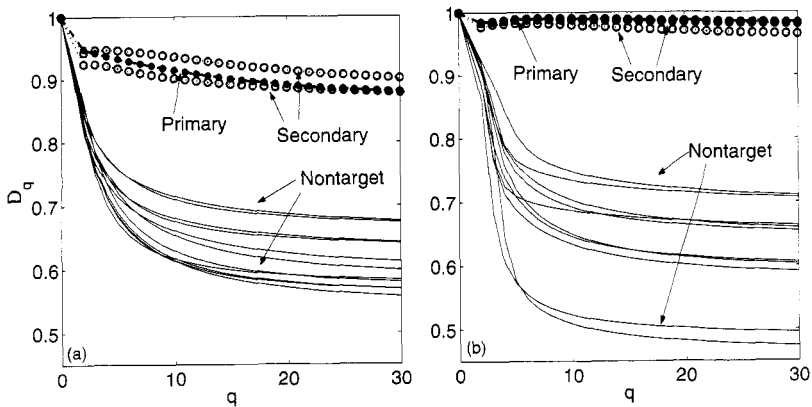
Figure 9.13.   The D_q spectrum for (a) HH and (b) VV datasets of one measurement.

In his pioneering work, DeLong [103] observed that discharge patterns of neurons in the GPi and GPe in an awake monkey (*Macaca mulatta*) at rest appeared to be very different. For GPe neurons, two types of firing patterns were observed: one had recurrent periods of high-frequency discharge with relatively silent periods in between, and the other had a low-frequency discharge with bursts. In contrast, only one firing pattern was found for GPi neurons: a continuous discharge without long periods of silence. DeLong's work has triggered much research on the characterization of different cell types in specific brain regions so that the accuracy of target acquisition of stereotactic electrode placement in the human brain can be automated
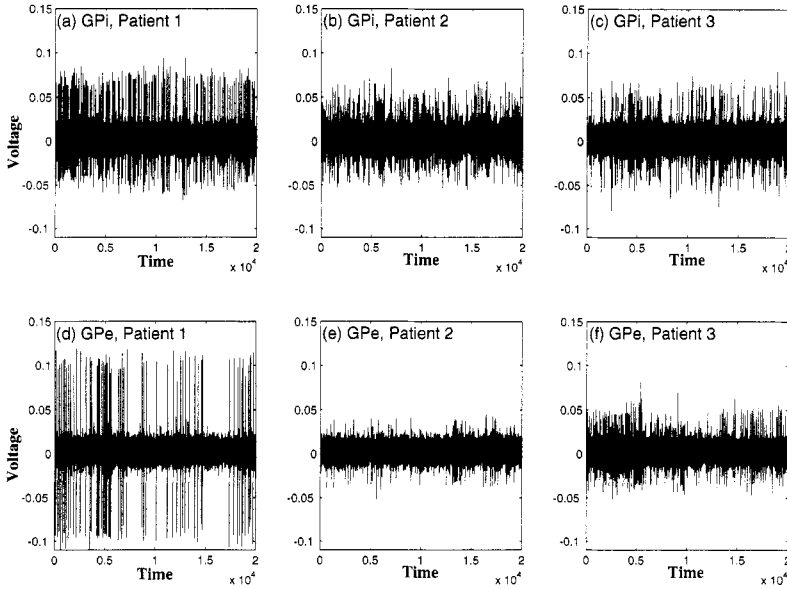
Figure 9.14.   Human deep brain recordings from GPe and GPi.

with high accuracy. It is worth noting that in practice, a well-trained, experienced movement disorders specialist is able to do so by listening to neuronal activity over a loudspeaker through trial and error. We shall show below that the generalized dimension spectrum D_q can effectively differentiate the two brain areas, both within and between patients. Note that the method is effective not only on the spike train data, but also on the raw voltage recording data. The latter is very appealing, since when the raw recording data are very noisy, spike detection/sorting can be difficult or even impossible.

The following discussion is based on the analysis of raw voltage recordings from three patients. See Fig. 9.14. The data were sampled with a frequency of 20 KHz. The mean interspike interval is about 0.01 s, i.e., 200 sample points. Another shorter time scale is the duration of each spike, which is about 0.0008 s, i.e., 16 sample points. It should be emphasized that fractal analysis is meaningful only for time scales greater than the mean interspike interval, since in shorter time scales, only noise and individual spikes can be characterized.

Since the raw data contain both positive and negative values, they are squared first. Then, following the procedures detailed at the beginning of Sec. 9.5, we can obtain the weights at all the stages and compute $M_q(\epsilon)$. Figure 9.15 plots $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$. It is observed that the scaling between $M_q(\epsilon)$ and ϵ (i.e., the degree of linearity between $\log_2 M_q(\epsilon)$ and $-\log_2 \epsilon$) for all datasets is quite good up to the ninth stage, which corresponds to the mean interspike interval. By fitting straight lines to $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$, from stages 9 to 4, which correspond to $t_1 = 0.01$ to $t_2 = 0.40$ s, one can readily obtain τ_q and hence D_q. The latter is shown in
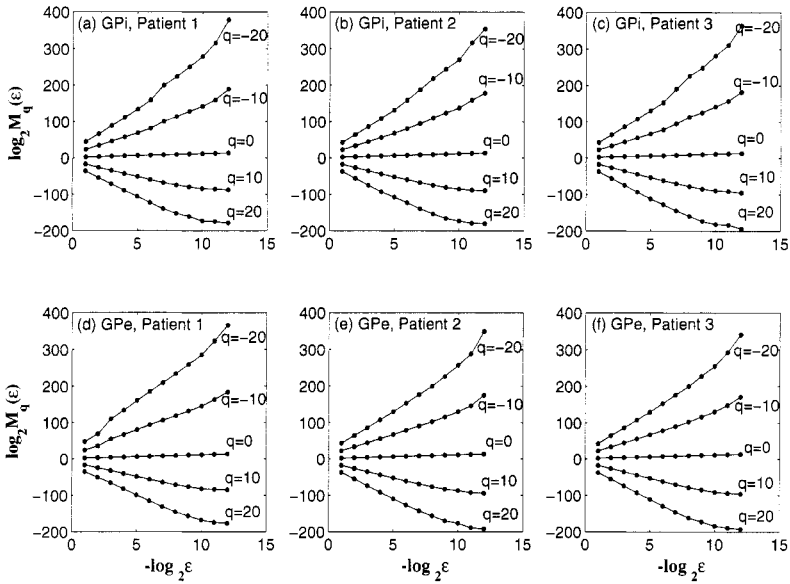
Figure 9.15. $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$ for human brain deep recordings for several different q's.

Fig. 9.16. It is evident that for very negative q, D_q is larger for GPi than for GPe, both within and between patients. This feature is an effective discriminator between GPe and GPi.

It should be noted that the effectiveness of the method in distinguishing between GPe and GPi is due largely to the identification of the two time scales. While we have explained the meaning of t_1, we note that $t_2 = 0.40$ s may be linked to the local and global neuronal interconnectivity of the brain. That is, due to finite propagation of neural signals, time scales longer than about 0.4 s might already correspond to neural interactions beyond GPi and GPe and hence may also be irrelevant to distinguish between the two structures. Indeed, based on these time scales, one may wonder why many traditional signal processing approaches, such as those based on the total spectral density of the signal, fail to distinguish GPe from GPi. The reason is simple: the majority of the energy in the spectral density comes from the noise between spikes and individual spikes. The useful information — the spacing between spikes — contributes little to the total energy.

9.5.3   Analysis and modeling of network traffic

In Sec. 6.7.1, we explained four different representations of network traffic. The purpose of multifractal analysis of network traffic is to check whether the interarrival time series $\{T_i\}$, packet length sequences $\{B_i\}$, and the counting processes $\{\overline{B}_i\}$ and $\{\overline{P}_i\}$ of the traffic data can be viewed as realizations of multiplicative processes.
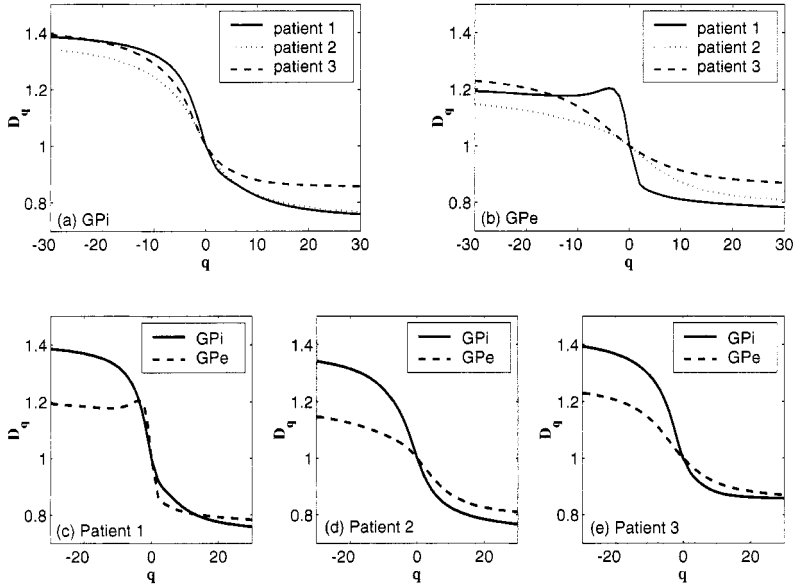
Figure 9.16.   The generalized dimension spectrum for recording data of GPi and GPe area of three patients.

If they are only approximate multifractals, an equivalent multifractal model may still be constructed. To illustrate the idea, we consider video traffic modeling here.

Owing to the increasing demand on video services, variable bit rate (VBR) video traffic modeling has become a critical issue. Simple and accurate video traffic models are needed to solve problems arising from multimedia traffic management, such as traffic policing, shaping, and call admission control. Many services (such as video on demand and broadcast TV) impose on video traffic special variability through the evolution of scenes of different complexity and degrees of motion. We show that a video traffic trace of "star wars" can be very effectively modeled by a cascade process. The trace consists of 174,136 integers, where each integer represents the number of bits per video frame (24 frames/second for approximately 2 hrs). Our analysis uses the first 2^{17} points.

Now that we have explained the usefulness of the $\log_2 M_q(\epsilon)$ vs. $-\log_2 \epsilon$ curves and the D_q spectrum, we work with the multiplier distributions. Following the procedures detailed at the beginning of the section, the multiplier distributions P_j for the video traffic data are shown in Fig. 9.17, where the asterisk curve is generated from $P(r) \sim e^{-\alpha_g |r-1/2|^2}$, with $\alpha_g = 200$. Collapsed on it are P_j curves with $j = 8, \cdots, 11$. We thus see that the video traffic forms a multifractal process over certain time scales.

After the multiplier distributions are found, we can readily generate the multi-fractal video traffic and compare it with the actual trace by feeding both modeled and real traffic to a queuing system. Figure 9.18 shows the comparison of the
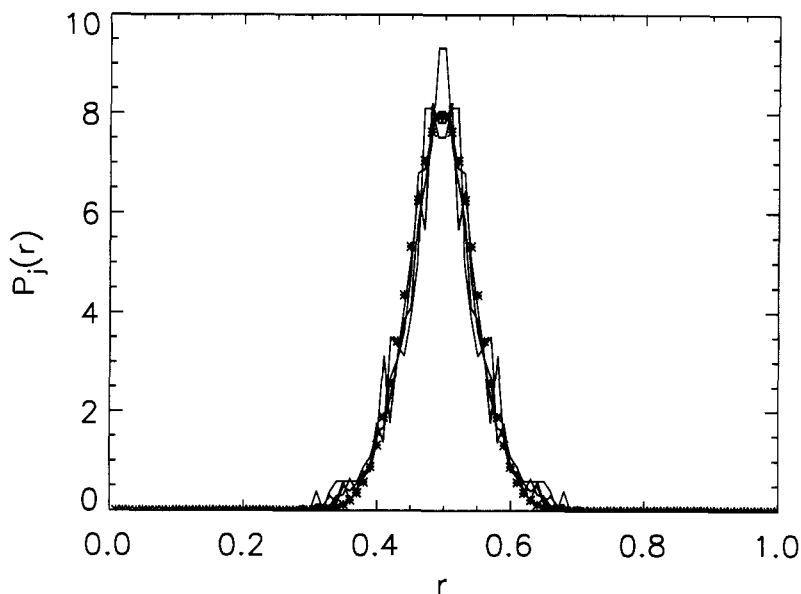
Figure 9.17.   Multiplier distributions P_j for the frame size sequence data MPEG.data. See the text for more detail.

system size tail distributions for a single-server FIFO queuing system loaded on the one hand by the measured traffic (dashed lines) and on the other hand by the multifractal traffic (solid lines). The system size is represented by the queue length, which measures the total number of queued frames normalized by the average frame size. The three curves, from top to bottom, correspond to three different utilization (i.e., normalized loading) levels, $\rho = 0.7$, 0.5, and 0.3. Clearly, the model yields an excellent fit of the system size tail distribution for a queuing system loaded by the measured video traffic.

9.6   BIBLIOGRAPHIC NOTES

The multiplicative cascade multifractal was initially developed for understanding the intermittent features of turbulence. The modern theory of turbulence started with Kolmogorov's theory of 1941 [265]. Since then, various models have been proposed. Early works include [293, 336]. For the various models discussed in Sec. 9.4, in order, we refer readers to [113, 258, 266, 304, 323, 399, 400]. See also [55] for a comparison of intermittency models of turbulence, [141] on experimental data analysis, [42, 43] on extended self-similarity, [398] on filaments in turbulence, and [47, 224] on raindrop formation. Since the discussions on turbulence here are brief, we have to omit many relevant references. Interested readers should read the exquisite books by Frisch [143] and Chorin [75] (in fact, we believe many readers
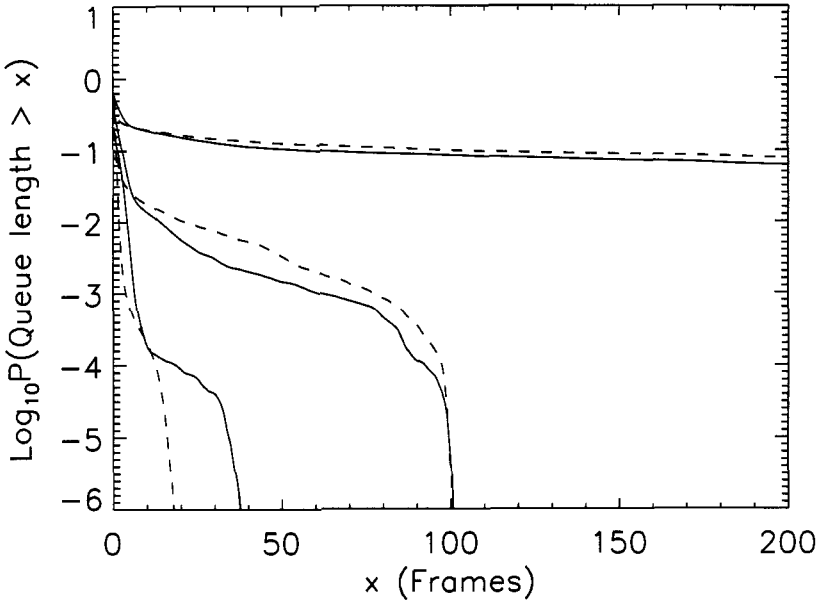
Figure 9.18. System size tail probabilities obtained when VBR video traffic MPEG.data (solid curves) and its corresponding multifractal traffic process (dashed curves) are used to drive the queuing system. Three curves, from top to bottom, correspond to $\rho = 0.7, 0.5$, and 0.3, respectively.

have done so.) Also, it should be emphasized that besides the cascade models discussed here, there exist other ways to tackle the problem of intermittency in turbulence. See, for example, [31,32].

It is interesting to note that the cascade model has been applied to the study of various phenomena such as rainfall [333], tropical convection [454], liquid water distributions inside marine stratocumulus [100, 101], and finance [296]. The generalized dimensions spectrum D_q was defined by Hentschel and Procaccia [219]. Systematic studies of the properties of the multiplicative processes can be found in [168]. Interesting sources on log-normality in network traffic are [30,54,63,311].

Readers interested in applications of cascade models on sea clutter are referred to the references noted in the end of Chapter 8, to [103,390,495] and many references cited therein on analysis of neuronal firing patterns, and to [44, 89, 134, 148, 165–173, 281, 368, 425] on traffic modeling.

9.7 EXERCISES

1. In Sec. 9.3, the properties of conservative cascade models are derived. Show that for a nonconservative model, all the properties except Eq. (9.14) still hold

and that Eq. (9.14) should read as follows:

$$E[(w_n - E(w))(w_{n+m} - E(w))] = \frac{1}{4}\mu_2^{N-1}(4\mu_2)^{-k} - 2^{-2N}.$$

2. Write a simple code to reproduce Figs. 9.4 and 9.6.

3. Analyze the time series generated from a random cascade multifractal and verify Inequality (9.16) (i.e., Fig. 9.8).

4. Compare the various intermittency models of turbulence discussed in Sec. 9.4 by plotting ζ_q vs. q for $-30 \le q \le 30$.

5. Prove that if the multiplier factor is stage-dependent, then the model is not a multifractal.

6. Consider the schematic for raindrop growth shown in Fig. 9.3: from one time step to another, $p = 90\%$ of raindrops remain their original size, while $1 - p = 10\%$ of them increase their size by a fixed ratio r. Now consider an arbitrary raindrop of unit size at time step 0. Answer the following:

    (a) At time step 1, the size of the raindrop can be either 1 or $1 + r$. What are their corresponding probabilities?

    (b) At time step 2, the size of the raindrop can be 1, $1 + r$, or $(1 + r)^2$. What are their corresponding probabilities?

    (c) Determine the sample space of the raindrop size at time step n, and determine the probabilities.

    (d) Prove that at step n, the k-th moment for drop size is

    $$[p + (1 + r)^k(1 - p)]^n$$