Modeling Infinitely Divisible Scaling: beyond Powerlaws

Rolf Riedi

Dept of Statistics

Cornell, April 2005
Scaling Analysis

An empirical view
Why Cascades

Turbulence: models wanted

• Kolmogorov 1941:
  \[ \langle [v(x+r)-v(x)]^q \rangle \sim r^{q/3} \]

• Kolmogorov 1962:
  \[ \langle [v(x+r)-v(x)]^q \rangle \sim r^{H(q)} \]

• Data is non-Gaussian

• ...presents structure on all scales

Courtesy P. Chainais
Powerlaws?

Real world data
• can deviate from powerlaws: network traffic
• Lukacs:
if the data does not fit to the model then too bad for the data.
Infinitely divisible scaling

- Self-similarity: $\mathbb{E}[|B(t+\delta)-B(t)|^q] \simeq \delta^{qH}$
- Multifractal scaling: $\mathbb{E}[|M(t+\delta)-M(t)|^q] \simeq \delta^{\tau(q)}$
- IDC scaling: $\mathbb{E}[|X(t+\delta)-X(t)|^q] \simeq \exp[n(\delta)\zeta(q)]$

- Multifractal scaling reduces to self-similarity if $\tau$ is linear in $q$. (sometimes called mono-fractal)
- IDC reduces to multifractal scaling if $n(\delta)=-\log(\delta)$
  - thus, for powerlaws
- In general $n(\delta)$ gives the speed of the cascade

- IDS suggested as a framework for statistical analysis in turbulence [Castaing] ...but where are the models?
- Greatest potential for models based on multiplication
Multiplicative models

- In distribution, the increment $X(t+\delta) - X(t)$ looks like a $n(\delta)$-fold product of iid multipliers

$$\mathbb{E}[e^{q \log |X(t+\delta) - X(t)|}] = \mathbb{E}[|X(t + \delta) - X(t)|^q]$$

$$\approx \exp[n(\delta) \zeta(q)]$$

$$= (e^{\zeta(q)})^{n(\delta)}$$

- Since $n(\delta)$ is a function of scale, multipliers are over scale not over time

- Warning: this does not mean that there are actual multipliers such that $X(t+\delta) - X(t) = M_1 \ldots M_{n(\delta)}$ a.s.:
  - For Brownian motion these n factors would have to be constant which is non-sense ($B(t)$ would be a.s. a square root)

$$B(2^{-n}) \overset{d}{=} (1/\sqrt{2})^n B(1)$$
Infinitely Divisible Cascades

Intuition:
A modeling framework for ID-scaling
**Toy: Binomial cascade**

\[ \epsilon := \{\epsilon_k\}_{k} \in \{0, 1\}^{\mathbb{N}} \]

\[ l_n(\epsilon) := \#\{k \leq n : \epsilon_k = 1\} = \sum_{k=1}^{n} \epsilon_k \]

\[ I_n(\epsilon) := \left[ \sum_{k=1}^{n} \epsilon_k/2^k, \sum_{k=1}^{n} \epsilon_k/2^k + 1/2^n \right) \]

\[ \mu_n(I_n(\epsilon)) = p^{n-l_n(\epsilon)}(1-p)^{l_n(\epsilon)} \]

\[ \mu_m(I_n(\epsilon)) = \mu_n(I_n(\epsilon)) \quad m \geq n \]

\[ \rightarrow \mu(I_n(\epsilon)) \]
Binomial as a product of pulses

We may write

\[ \mu(I_n(\epsilon)) = \int_{I_n(\epsilon)} Q_n(t) \, dt \]

where

\[ Q_n(t) = 2^n p^{n-l_n} (1 - p)^{l_n} \]

and \( l_n(\epsilon(t)) \) is constant over each \( I_n \).

Assign multipliers \( 2p \) and \( 2(1 - p) \) to dyadic intervals and compute \( Q_n(t) \) as the product of multipliers contained in the cone

\[ C_r(t) = \{(t_i, r_i) : t - r_i/2 < t_i < t + r_i/2, r_i > r\} \]
From Binomial to IDC

We may write

$$\mu(I_n(\varepsilon)) = \int_{I_n(\varepsilon)} Q_n(t) dt$$

where $Q_n(t)$ is a product of multipliers.

- Assign multipliers according to
  - a marked point process...
  - ...or a random measure

\[ \prod_{(t_i,r_i) \in C_r(t)} W_i \]

\[ \exp[M(C_r(t))] \]
Infinitely Divisible Cascades

Definitions:
Infinitely divisible measures
Cascades
Infinitely Divisible Measure

- **M** is an infinitely divisible measure
  - Randomly scattered:
    - M(E), M(F) independent for disjoint E and F
  - Measure of set E is inf. div. r.v.

\[
\mathbb{E}[e^{qM(E)}] = \exp \left[ -\rho(q)m(E) \right]
\]

**Ex 1: Poisson count measure:**
M(E) = Poisson variable with mean m(E)

\[
\mathbb{E}[e^{qM(E)}] = \sum_{k \geq 0} e^{-m(E)} \frac{m(E)^k}{k!} e^{qk} = \exp \left[ (e^q - 1)m(E) \right]
\]

\[
\rho(q) = 1 - e^q
\]
Infinitely Divisible Measure

- **M** is an infinitely divisible measure
  - Randomly scattered:
    - **M**(E), **M**(F) independent for disjoint E and F
  - Measure of set E is inf. div. r.v.

\[ \mathbb{E}[e^{qM(E)}] = \exp[-\rho(q)m(E)] \]

**Ex 1:** Poisson count measure:
**M**(E) = Poisson variable with mean m(E)
\[ \rho(q) = 1 - e^q \]

**Ex 2:** Poisson process, marked with W
**M**(E) = Compound Poisson, marks W, mean m(E)
\[ \rho(q) = 1 - \mathbb{E}[W^q] \]
Infinitely Divisible Measure

- **M** is an infinitely divisible measure
  - Randomly scattered:
  - Measure of set \( E \) is inf. div. r.v.

\[
\mathbb{E}[e^{qM(E)}] = \exp[-\rho(q)m(E)]
\]

**Ex 1: Poisson:** mean \( m(E) \)

\[
\rho(q) = 1 - e^q
\]

**Ex 2: Compound Poisson:** mean \( m(E) \), mark \( W \)

\[
\rho(q) = 1 - \mathbb{E}[W^q]
\]

**Ex 3: Gaussian measure:** variance factor \( m(E) \)

\[
(\mu = 0 \text{ w.l.o.g. due to later normalization})
\]

\[
\rho(q) = -q\mu - q^2\sigma^2/2
\]
Infinitely Divisible Cascade

• M is an infinitely divisible measure
  – Randomly scattered
  – Measure of set E is inf. div. r.v.
    \[ \mathbb{E}[e^{qM(E)}] = \exp[-\rho(q)m(E)] \]
  – Assume \( \exp(M(E)) \) has finite mean (\( \rho(1) \) defined)

• Parameters:
  – Infinitely divisible law via \( \rho(q) \)
  – Control measure \( dm(t,r) \)
  – Cone: causal
    \[ C_r(t) = \{(t_i, r_i) : t - r_i < t_i < t, r_i > r\} \]
  – Symmetrical
    \[ C_r(t) = \{(t_i, r_i) : t - r_i/2 < t_i < t + r_i/2, r_i > r\} \]

• IDC:
  \[ \tilde{Q}_r(t) = \exp[M(C_r(t))] \quad Q_r(t) = \frac{\tilde{Q}_r(t)}{\mathbb{E}[\tilde{Q}_r(t)]} \]
Infinitely Divisible Cascades

1\textsuperscript{st} Properties:
Scaling
Scaling of Infinitely Divisible Cascades

- **M** is an infinitely divisible measure
  - Randomly scattered
  - Measure of set E is inf. div. r.v.

\[ \mathbb{E}[e^{qM(E)}] = \exp[-\rho(q)m(E)] \]

- **IDC:**

\[ \tilde{Q}_r(t) = \exp[M(C_r(t))] \quad Q_r(t) = \frac{\tilde{Q}_r(t)}{\mathbb{E}[\tilde{Q}_r(t)]} \]

\[ \frac{\mathbb{E}[e^{qM(E)}]}{\mathbb{E}[M(E)]^q} = \exp\left[-\left(\rho(q) - q\rho(1)\right)m(E)\phi(q)\right] \]

- **Scaling:**

\[ \mathbb{E}[Q_r(t)^q] = \exp[-\phi(q)m(C_r(t))] \]

| Self-similarity: | \[ \mathbb{E}||B(t+\delta) - B(t)||^q \approx \delta^{qH} \] |
|-----------------|---------------------------------------------------|
| Multifractal scaling: | \[ \mathbb{E}||M(t+\delta) - M(t)||^q \approx \delta^{q(q)} \] |
| IDC scaling: | \[ \mathbb{E}||X(t+\delta) - X(t)||^q \approx \exp[n(\delta)\zeta(q)] \] |

\[ m(C_r) \to n(\delta), \text{ speed} \]

\[ \phi(q) \to \zeta(q), \text{ law} \]
Infinitely Divisible Cascade: examples

- **IDC Scaling:** \( \mathbb{E}[Q_r(t)^q] = \exp \left[ - \frac{(\rho(q) - q\rho(1)) m(C_r(t))}{\varphi(q)} \right] \)
  - **Note:** \( \varphi(0) = \varphi(1) = 0 \) and \( \varphi \) is convex

**Ex 1: Poisson:** mean \( m(E) \)
\[
\rho(q) = 1 - e^q \quad \varphi(q) = 1 - e^q + q(e - 1)
\]

**Ex 2: Compound Poisson:** mean \( m(E) \), mark \( W \)
\[
\rho(q) = 1 - \mathbb{E}[W^q] \quad \varphi(q) = 1 - \mathbb{E}[W^q] - q(1 - \mathbb{E}[W]) \quad \varphi(2) = -\mathbb{E}[(W - 1)^2]
\]

**Ex 3: Gaussian measure:** variance factor \( m(E) \)
\[
\rho(q) = -q\mu - q^2\sigma^2/2 \quad \varphi(q) = -(q^2 - q)\sigma^2/2
\]

**Ex 4: Stable, \( \alpha < 1, \beta = -1 \)**
\[
\varphi(q) = \sigma^\alpha(q - q^\alpha) \left( 1 - \tan\left( \frac{\pi\alpha}{2} \right) \right) \quad \text{for } q > 0.
\]
Special case: CPC

- Marked Poisson process \((T_i, R_i, W_i)\)
  - \(m(A) = E[\# \text{points in } A]\)
  - Marks \(W_i\) are i.i.d.

**Cascade Process at scale \(r\):**

\[
Q_r(t) = \gamma_r \cdot \prod_{(t_i, r_i) \in C_r(t)} W_i
\]

\[
\gamma_r = \frac{1}{E} \left[ \prod_{(t_i, r_i) \in C_r(l)} W_i \right]
\]
CPC: dual view

• **Pulse** at \((T_i, R_i, W_i)\)

\[
P_i(t) = \begin{cases} 
W_i & \text{if } T_i \leq l < T_i + R_i \\
1 & \text{else}
\end{cases}
\]

• The pulse multipliers active at time \(t\) are the ones with \((T_i, R_i)\) in the cone

• Cascade Process at scale \(r\):

\[
Q_r(t) = \gamma_r \cdot \prod_{r_i > r} P_i(t)
\]

• Compare with infinite Poisson source model from queuing
Infinitely Divisible Cascades

First look at parameters:

Control measure

- Stationary increments
- Scaling
Stationarity of $Q(t)$

- IDC defined as
  \[ Q_r(t) = \frac{\tilde{Q}_r(t)}{\mathbb{E}[\tilde{Q}_r(t)]} \quad \tilde{Q}_r(t) = \exp[M(C_r(t))] \]

- Recall
  \[ \mathbb{E}[e^{qM(E)}] = \exp[-\rho(q)m(E)] \]

- Stationary if $m$ is time-translation invariant, e.g.:
  \[ dm(t, r) = g(r) dt dr \]
Infinitely divisible nature

Exploit affinity of cone:

\[ C_r^b = \{(t_i, r_i) \in C_r : r < b\} \quad C_r = C_r^b \cup C_b \]

\[
\begin{align*}
Q_r(t) &= \prod_{C_r(t)} W_i \times \prod_{C_b(t)} W_i \times \prod_{C_r^b(t)} 1 \\
&= Q_b(t) \times \prod_{C_r^b} W_i \\
&= Q_b(t) \times Q_r^b
\end{align*}
\]

Note: \( Q_r \) is Martingale

Also:

\[
\mathbb{E} Q_r^b(t)^q = \exp \left[ -\varphi(q) m(C_r^b) \right]
\]
Binomial: perfect rescaling

• We may write

\[ \mu(I_n(\epsilon)) = \int_{I_n(\epsilon)} 2^n p^{n-l_n}(1-p)^{l_n} dt \]

• The number of multipliers in an octave

\[ C^b_r(t) := C_r(t) \setminus C_b(t) \]

depends only on \( b/r \geq 1 \).

• Generalize
  
  - \( Q^b_r(t) = \exp[M(C^b_r(t))] \) depends only on \( b/r \) (in distribution).
  - \( m(C^b_r(t)) \) depends only on \( b/r \).
  - \( m(C^b_r(t)) = c \log(b/r) \) and \( dm(t, r) = c/r^2 dt dr \).
CPC: Invariance and powerlaws

\[ Q_r(t) = Q_b(t) \times Q_r^b(t) \]

\[ \mathbb{E}Q_r^b(t)^q = \exp \left[ -\varphi(q)m(C_r^b) \right] \]

**Invariance:**

- In law:
  - \( Q_r^b \) rescaled version of \( Q_{r/b} \)
  - \( m(C_r^b) = m(C_{r/b}) \)
  - \( m(C_r^b) = c \log(b/r) \)
  - \( dm(t,r) = c/r^2 dt dr \)

- **powerlaws:**
  - \( Q_{b^n}(t) = Q_b(t) \times Q_{b^2}(t) \ldots \times Q_{b^n}(t) \)
  - \( \mathbb{E}Q_r(t)^q = \exp \left[ -\varphi(q)c \log(r) \right] = r^{-c\varphi(q)} \)
Scaling: beyond powerlaws

• Degrees of freedom:
  – Shape of Cone can be compensated by transforming the scale \((t,r) \rightarrow (t,r')\)
  – Corresponds to warped control measure

• Interest in rich behavior at small scale:
  – Control measure “explodes” at \(r=0\)

• Scale-invariance goes along with powerlaws
Credits

- Continuous multiplicative cascades from stochastic equations
  Schmitt [Marsan 2001]

- Multifractal random walk [Bacry Delour Muzy 2001]

- Products of pulses Barral [Mandelbrot 2002]

- Log-infinitely divisible cascades [Bacry Muzy 2002]

- Compound Poisson Cascades, Infinitely divisible cascades
  [Chainais Abry R. 2002; 2003]
Infinitely Divisible Cascades

Advanced Properties:
Convergence and non-degeneracy
Degenerate cascades

- \{Q_r(t)\}_{r>0}: positive, left-continuous martingale
  - Converges almost surely
- Q degenerates
  - As \( r \to 0 \), \( Q_r(t) \to 0 \) for a.a. \( t \), a.s.
  - Reason: LLN for \( \log[Q] \) and \( \mathbb{E}[\log Q] < \log \mathbb{E}[Q] = 0 \)
- Set
  \[ A_r(t) = \int_0^t Q_r(s) \, ds \]
- \( \{A_r(t)\}_{r>0} \) is positive martingale
- Non-degenerate:
  If A converges in \( L_p \) (\( p>1 \)) then
  \[ \mathbb{E}[A(t)] = \lim_r \mathbb{E}[A_r(t)] = \lim_r \int_0^t \mathbb{E}[Q_r(s)] \, ds = t \]
L2 convergence

• L2 convergence and correlations are governed by control measure of cone intersections, ie, of cones

• A converges in L2 iff

\[ \mathbb{E}[A_r(t)^2] = \int_0^t \int_0^t \mathbb{E}[Q_r(u)Q_r(v)] \, du \, dv < K. \]

\[ \mathbb{E}[Q_r(u)Q_r(v)] = \exp \{ -\varphi(2)m(C_r(u) \cap C_r(v)) \} \]

• Scale invariant m:

\[ \mathbb{E}[Q_r(t)Q_r(s)] = \begin{cases} |t-s|^{\varphi(2)} e^{-\varphi(2)(|t-s|-1)} & \text{for } r \leq |t-s| \leq 1 \\ 1 & \text{for } 1 \leq |t-s|. \end{cases} \]
Convergence in $L^q$

Let $1 < q \leq 2$. A sufficient condition for convergence of $A_r(t)$ in $L^q$ is

$$\limsup_{n \to \infty} \frac{1}{n} m(C_{tk_0^{-n-1}}) < \frac{q - 1}{\varphi(q)} \log(1/k_0) \quad (1)$$

for some integer $k_0 \geq 2$ (recall that $\varphi(q) < 0$).

In the scale invariant case of $m(C_r) = -c \log(r)$ this becomes

$$(q - 1) + c \varphi(q) > 0. \quad (2)$$

[Barral] for CPC, extension to IDC [CAR]
Infinitely Divisible Cascades

Scaling

Parameter estimation
Estimation issues

• In practice we observe only $Q_r$ and not the entire cascade
• To recover hidden scaling information consider

$$A_r(t) = \int_0^t Q_r(s) \, ds$$

• ...and study
  – A.s. limit $A(t) = \lim_{r \to 0} A_r(t)$
  – Scaling of limit

$$\mathbb{E}[A(t)^q] = \mathbb{E}[(A(t + s) - A(s))^q]$$
Auto-correlation

• Recall

\[ \mathbb{E}[Q_r(u)Q_r(v)] = \exp \{ -\varphi(2)m(C_r(u) \cap C_r(v)) \} \]

→ Estimation of \( dm = g(r)dt \, dr \)

• Examples:
  - Poisson \( \varphi(2) = 1 - (1 - e)^2 \)
  - Compound Poisson \( \varphi(2) = -\mathbb{E}[(W - 1)^2] \)
  - Gaussian \( \varphi(2) = -\sigma^2 \)
  - Stable
    \[ \varphi(2) = -\sigma^\alpha(2 - 2^\alpha) \left( 1 - \tan\left( \frac{\pi \alpha}{2} \right) \right) \]
Binomial: scaling of moments

- We may write
  \[ \mu(I_n(\epsilon)) = \int_{I_n(\epsilon)} 2^n p^n(1 - p)^n dt \]

- The density “Q” is constant over dyadic intervals
  - Convergence straightforward
  - Moments scale perfectly...
  - ...at least over perfectly dyadic scales

- Generalize
  \[ A_r(t) = \int_0^t Q_b(s) Q_r^b(s) ds = b \int_0^t Q_b(s) d \left[ A_r/b \left( \frac{s}{b} \right) \right]. \]

  IDC independent of \( Q_b \)
Scaling of $A(t)$

Fix $q > 0$, $b \in (0, 1)$, $\rho(\cdot)$ and $dm = g(r)dt\,dr$.

- Moment condition: $A_t$ converges in $\mathcal{L}^q$.
- Variational condition: “Technical”
  e.g.: CPC and log-normal
- Speed condition: “sub-invariant”,
  e.g., $dm(t, r) = r^\beta dt\,dr$ ($\beta \leq 2$) and log-deviations.

Then there exist constants $\overline{C}_q$ and $\underline{C}_q$ such that for any $t < 1$

$$C_q t^q \exp \left[ -\varphi(q)m(C_t) \right] \leq \mathbb{E} A(t)^q \leq \overline{C}_q t^q \exp \left[ -\varphi(q)m(C_t) \right].$$
Convergence in $L^p$

- Under assumptions of “scaling” theorem

- Scale-invariant case ($dm(t,r) = c/r^2 \, dt \, dr$)
  - Sufficient condition for convergence in $L^p$
    $$(p - 1) + c\varphi(p) > 0$$
  - Necessary condition for convergence in $L^p$
    $$(p - 1) + c\varphi(p) \geq 0$$

Proof idea

$$\mathbb{E}[A(t)^p] \geq 2\mathbb{E}[A(t/2)^p] \asymp 2(t/2)^p (t/2)^{c\varphi(p)}$$
Summary IDC scaling

- Multifractal formalism holds in self-similar case [Barral-Mandelbrot]
- Infinitely Divisible Scaling

\[ \mathbb{E} Q_r(t)^q = \exp \left[ -\varphi(q) m(C(r, 0)) \right] \]

\[ \mathbb{E} A(t)^q \sim t^q \exp \left[ -\varphi(q) m(C(t, 0)) \right] \]

- powerlaw only if \( m(C(r,0)) = -c \log(r) \)
- Established
  - for IDC in self-similar case [Bacry-Muzy, Barral]
  - for CPC and log-normal IDC in certain non-powerlaw cases [Chainais-R-Abry]
Simulation and Estimation
Simulations of CPC

• Stationary Cascade:

• Non-powerlaw scaling
Empirical analysis of network traffic

**IDC scaling:**

\[ S_q := \mathbb{E}[|X(t + \delta) - X(t)|^q] \]

\[ H(q) := \log S_q \simeq n(\delta)\zeta(q) \]

*Fix \( p \):*

\[ H(q)/H(p) \simeq \zeta(q)/\zeta(p) \]

independent of scale

- Clear indication of transition in speed
- Different mechanisms at work at small and large scales
Experimental results (turbulence)

H(q)

n(a): non-powerlaw

Courtesy P. Chainais
Multifractal Subordination

Processes with multifractal oscillations
Multifractal time warp

$B_H(M(t))$: $B_H$ fBm, $dM$ independent measure

A versatile model

- $M(t)$: Multifractal
  Time change
  Trading time

- $B$: Brownian motion
  Gaussian fluctuations
Hölder regularity

- Levy modulus of continuity:
  - With probability one for all $t$
    \[ |B_H(t + \delta) - B_H(t)| \approx |\delta|^H \]
  - Thus, exponent gets stretched:
    \[ |B_H(M(t + \delta)) - B_H(M(t))| \approx |M(t + \delta) - M(t)|^H \approx |\delta|^{H\alpha(t)} \]
  - and spectrum gets squeezed:
    \[ \dim E_a[B_H(M)] = \dim E_a/H[M] \]
Multifractal Estimation for $B(M(t))$

- Weak Correlations of Wavelet-Coefficients:
  (with P. Goncalves)

- Improved estimator due to weak correlations

- Multifractal Spectrum

\[ M(t+s) - M(t) \sim s^{a(t)} \]
\[ B(t+u) - B(t) \sim u^H \quad (\forall \, t) \]
\[ \Rightarrow \]
\[ B(M(t+s)) - B(M(t)) \sim s^{H*a(t)} \]

Estimation
Take home

- **IDC:**
  - Stationary increments
  - Continuous multiplication
  - Versatile scaling

- **Unexplored**
  - Higher dimensions: anisotropy
  - LRD
  - Pulse shapes
  - Estimation Theory
Reading on this talk

• www.stat.rice.edu/~riedi
• This talk
• Intro for the “untouched mind”
  – Explicit computations on Binomial
• Monograph on “Multifractal processes”
  – Multifractal formalism (proofs, references)
  – Multifractal subordination (warping)
• Papers on “Infinitely Divisible Cascades”
  [with Chainais and Abry]
• links
The end