Modeling Telecommunications Traffic

Long-range dependence

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

Admin

Exam, Thursday the 11th, in the morning.

Check your timetables!

Quidquid latine dictum sit, alutum viditur.

Self-similarity

Self-similarity

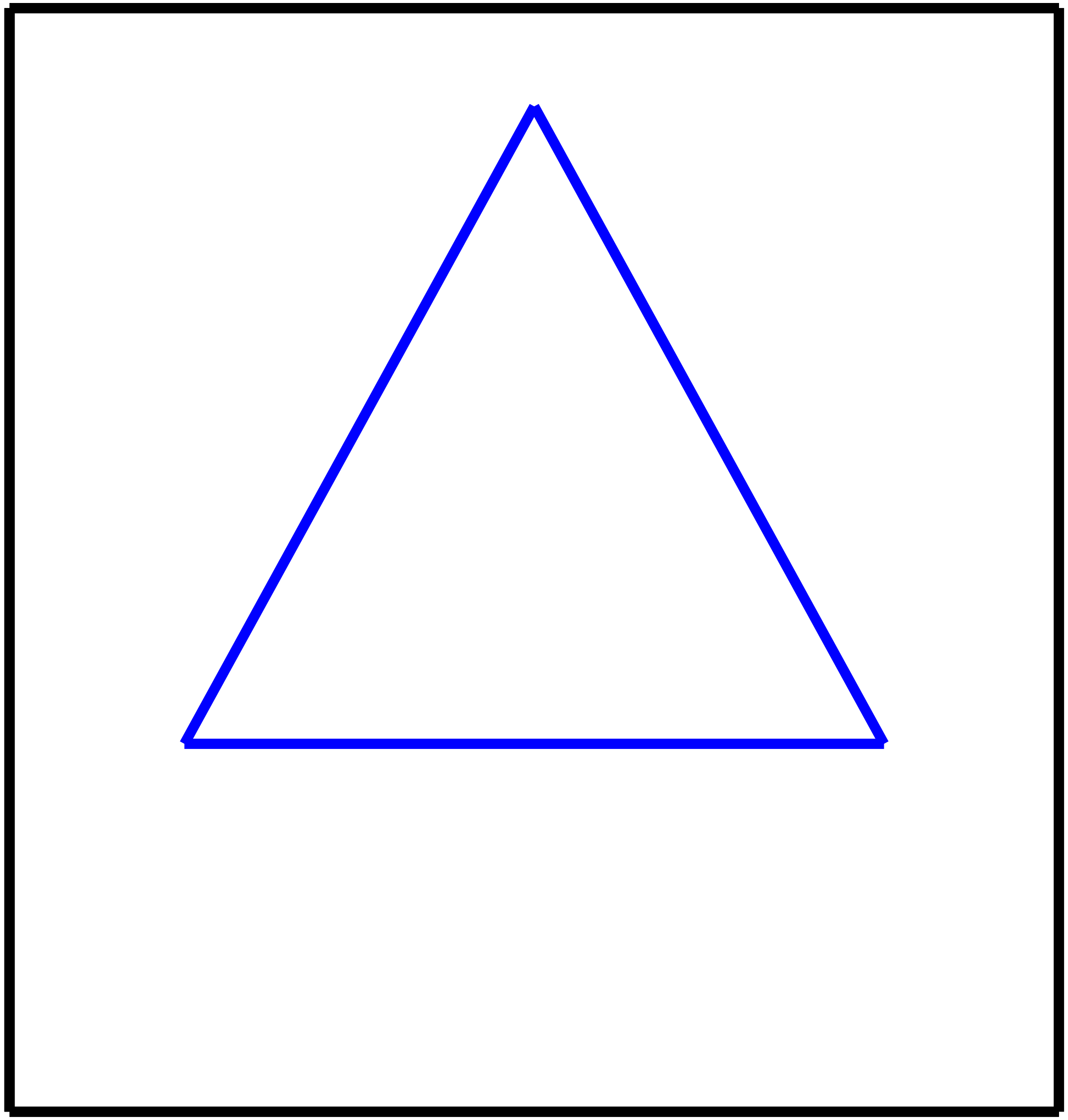
So, Nat'ralists observe, a flea
Hath smaller fleas that on him prey;
And these have smaller still to bite 'em
And so proceed ad infinitum

Jonathon Swift, 1733

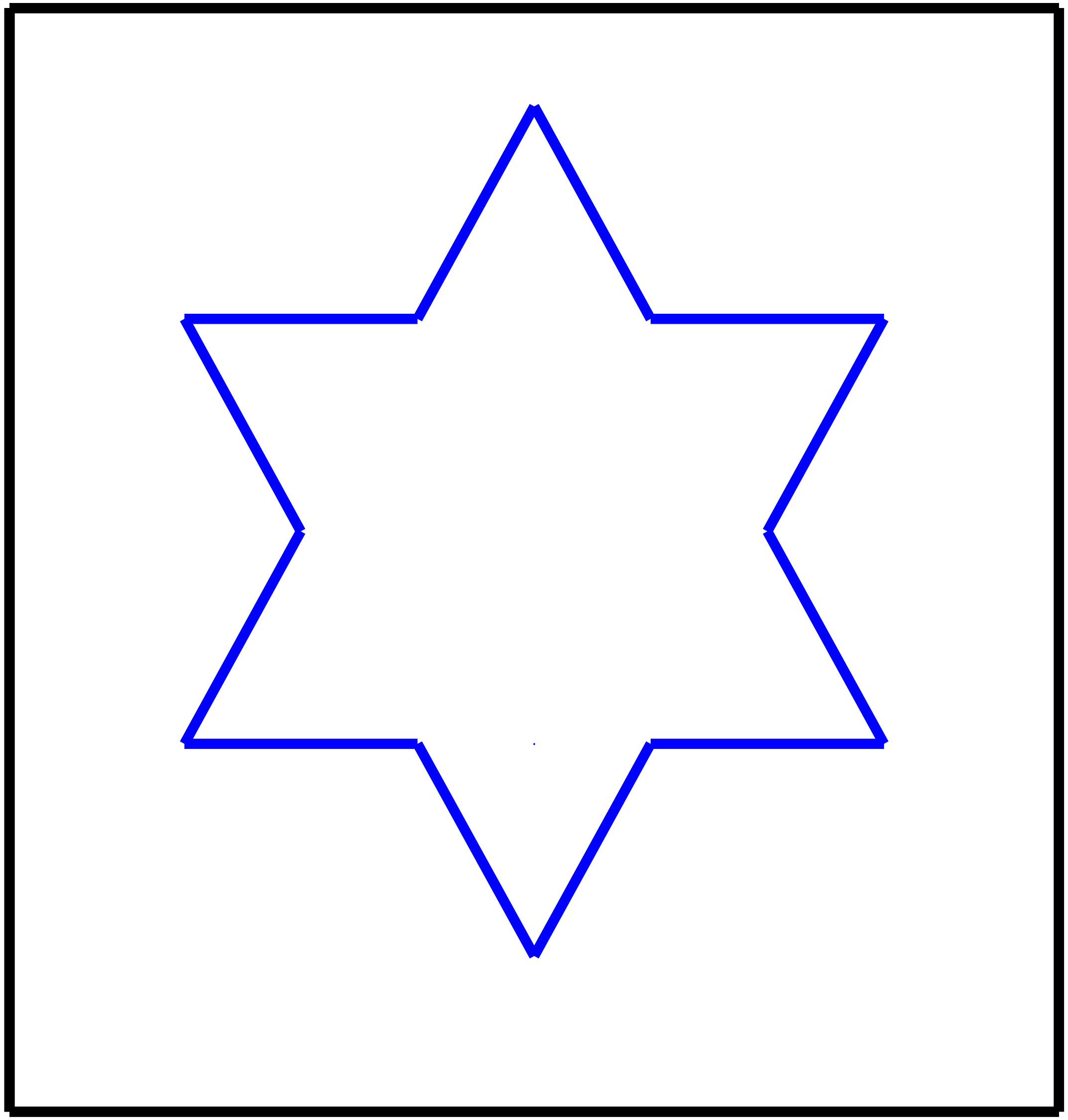
Great fleas have little fleas upon their backs to bite 'em,
And little fleas have lesser fleas, and so ad infinitum.
And the great fleas themselves, in turn, have greater fleas to go on;
While these again have greater still, and greater still, and so on.

De Morgan: A Budget of Paradoxes, p. 377.

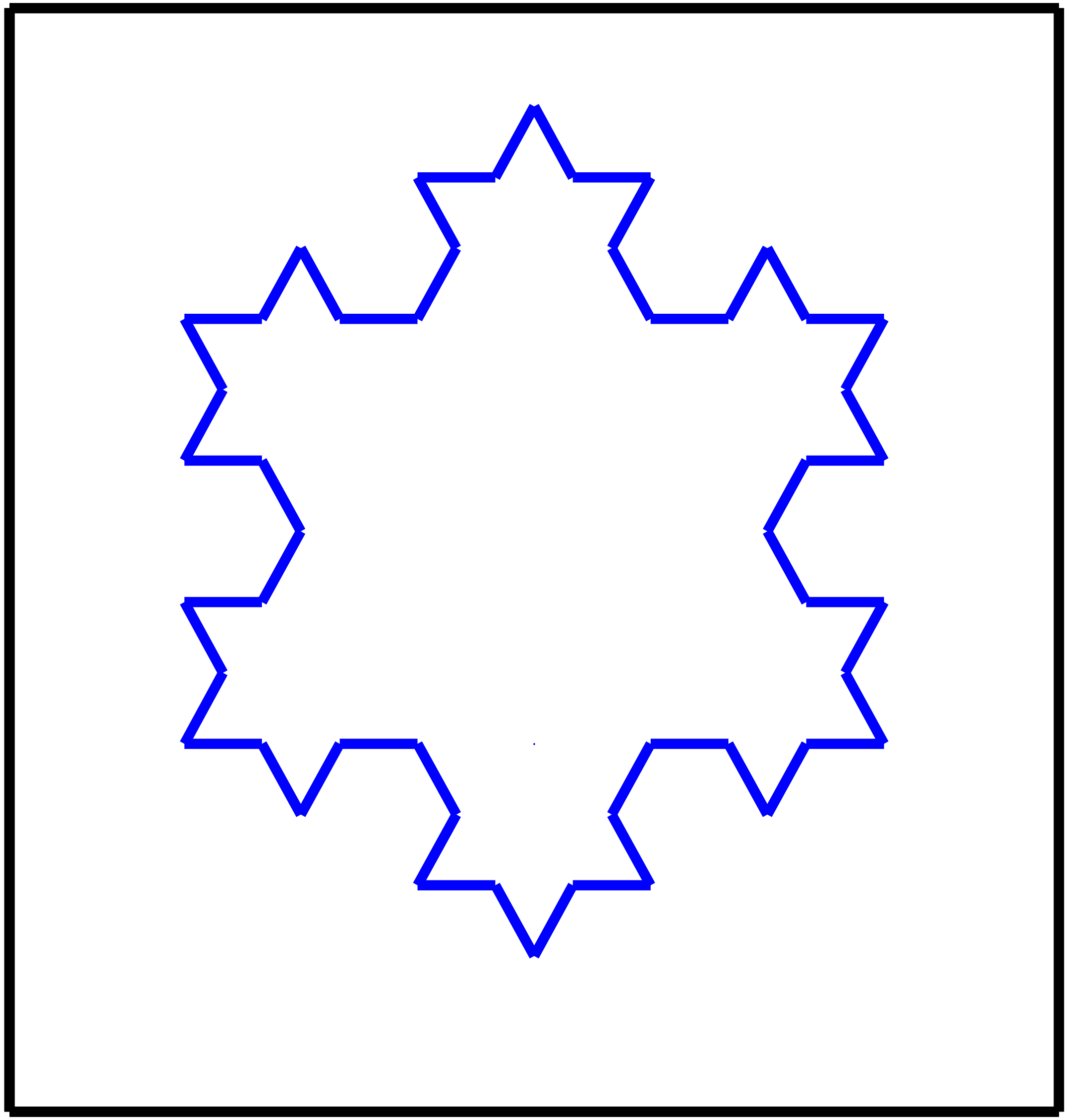
Self-similarity: Koch Snowflake



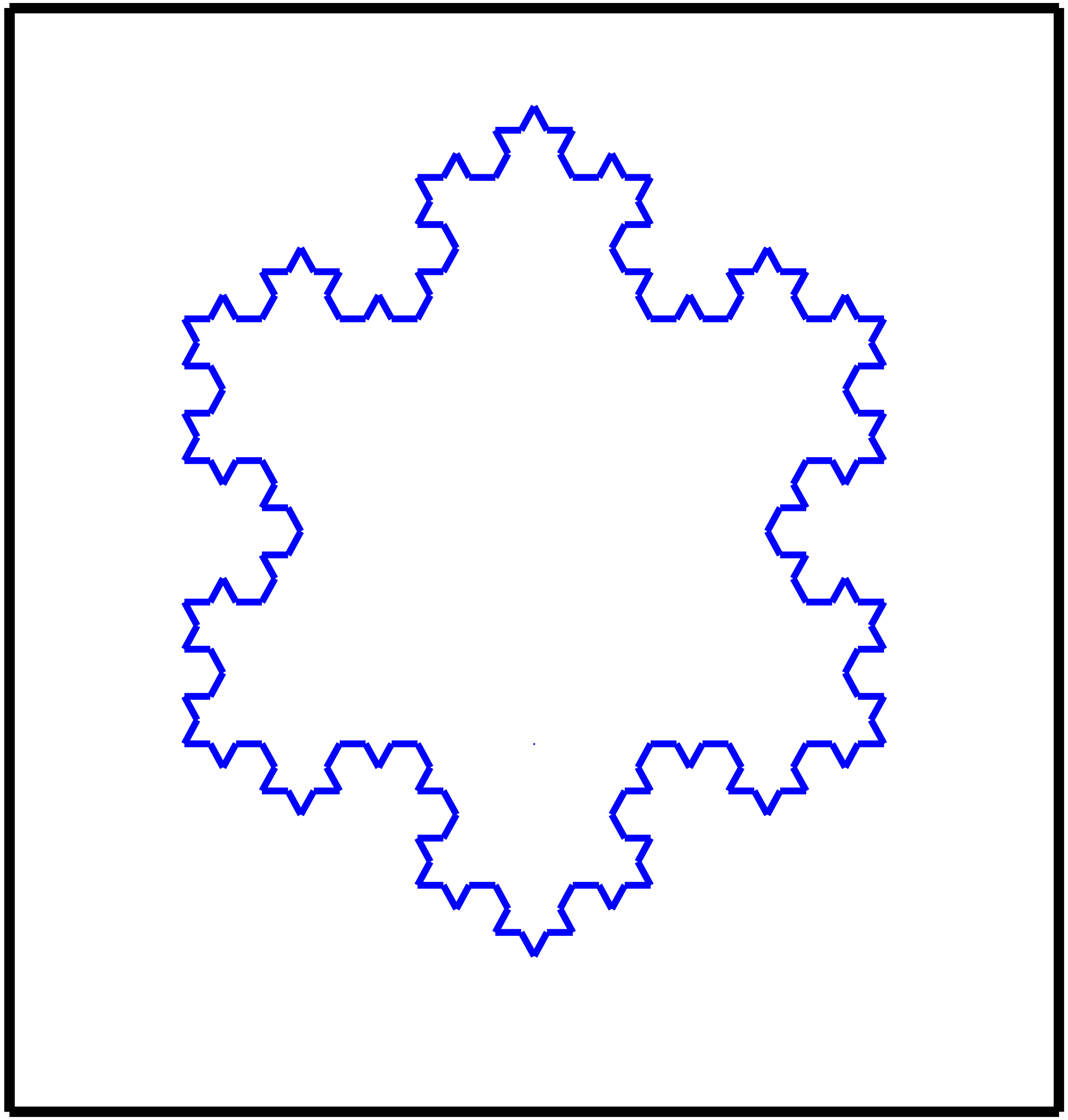
Self-similarity: Koch Snowflake



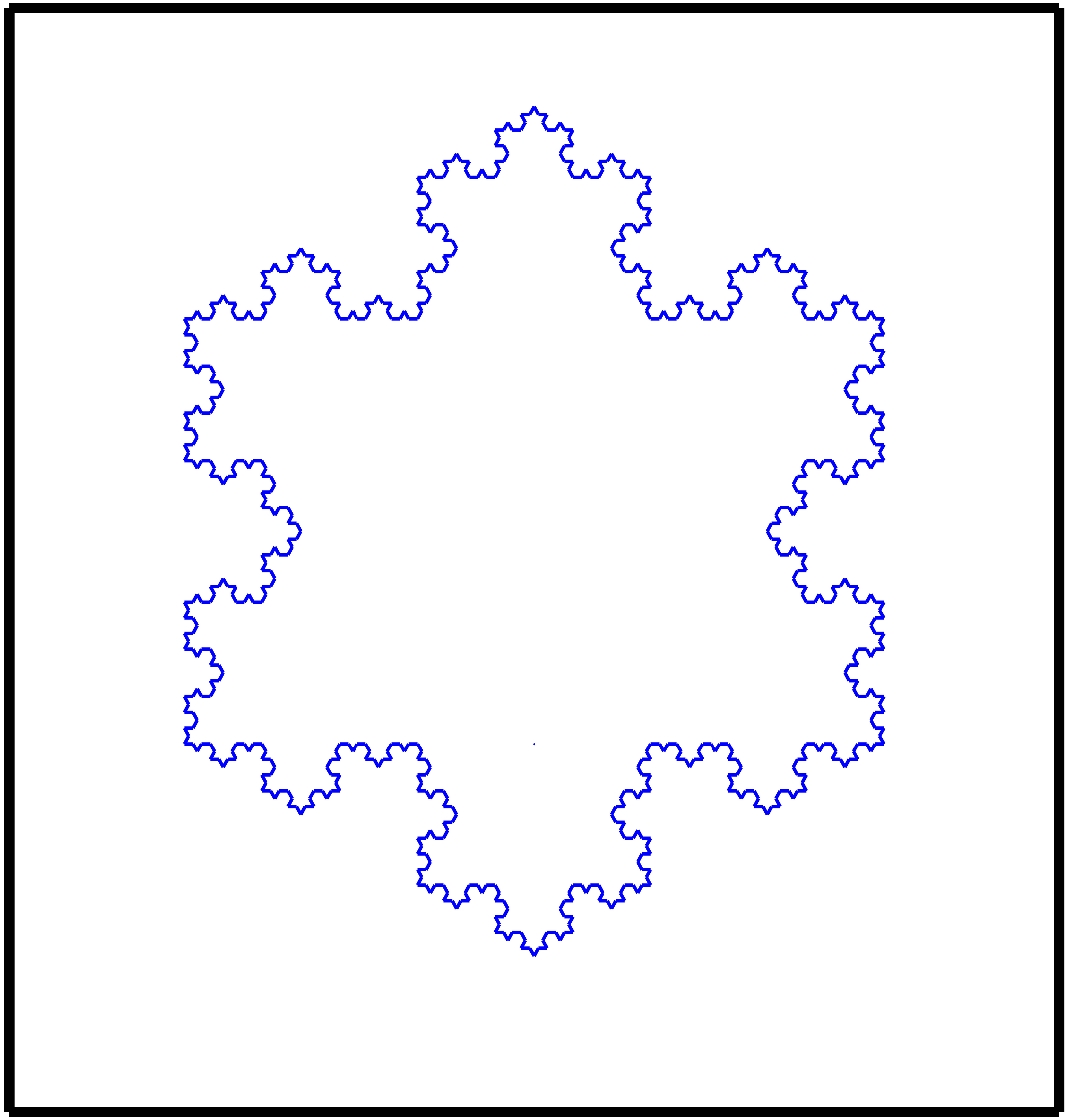
Self-similarity: Koch Snowflake



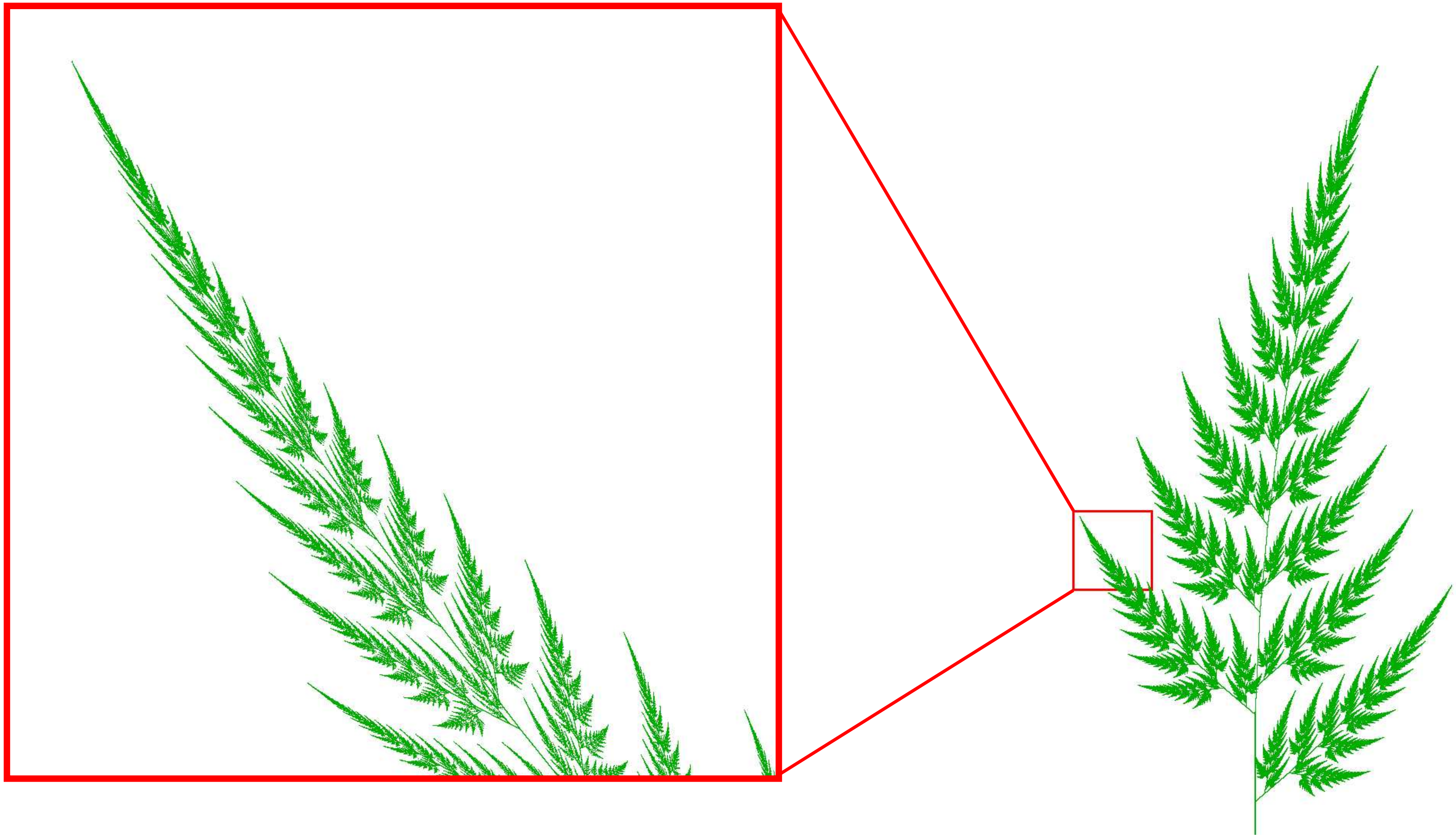
Self-similarity: Koch Snowflake



Self-similarity: Koch Snowflake

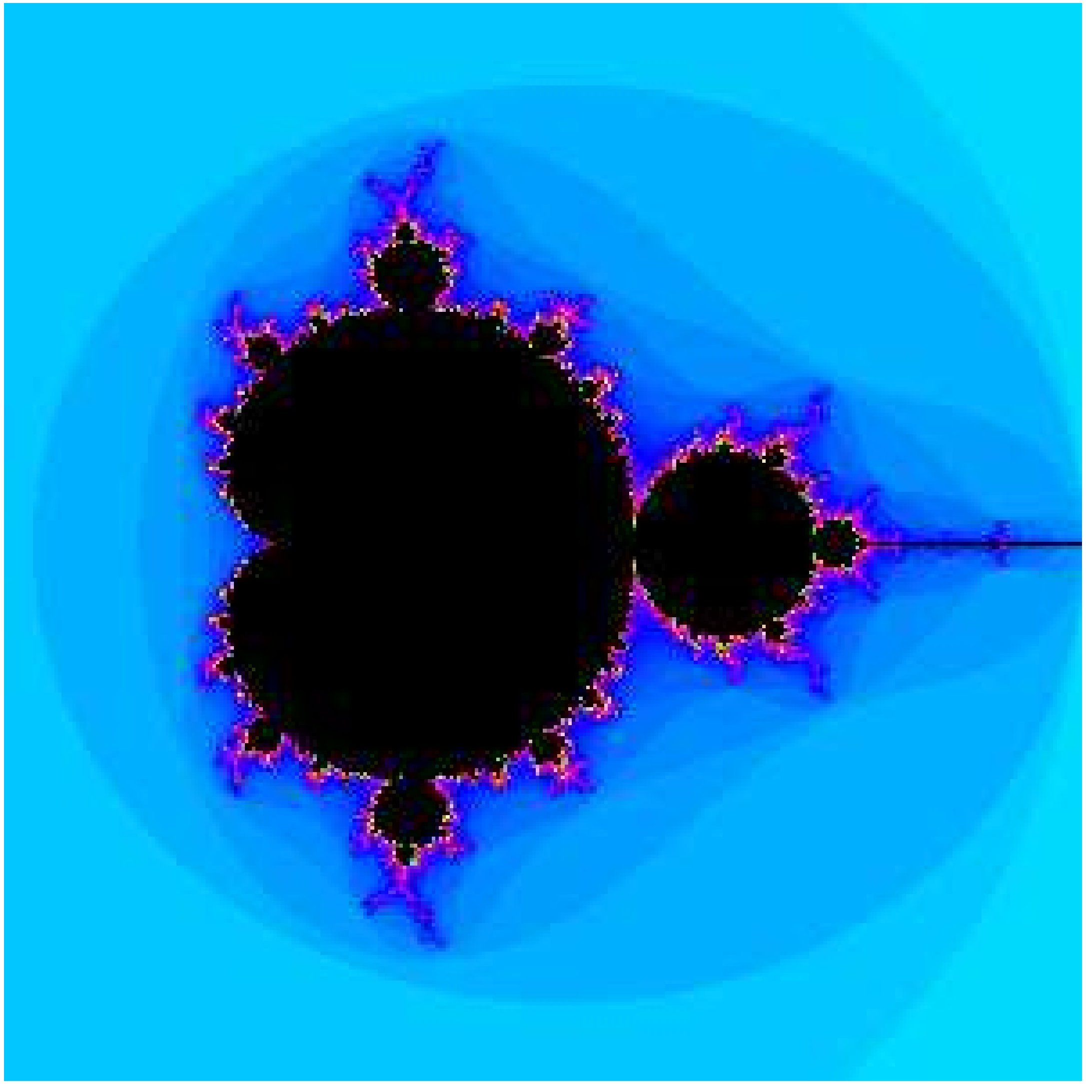


Self-similarity: IFS Fern



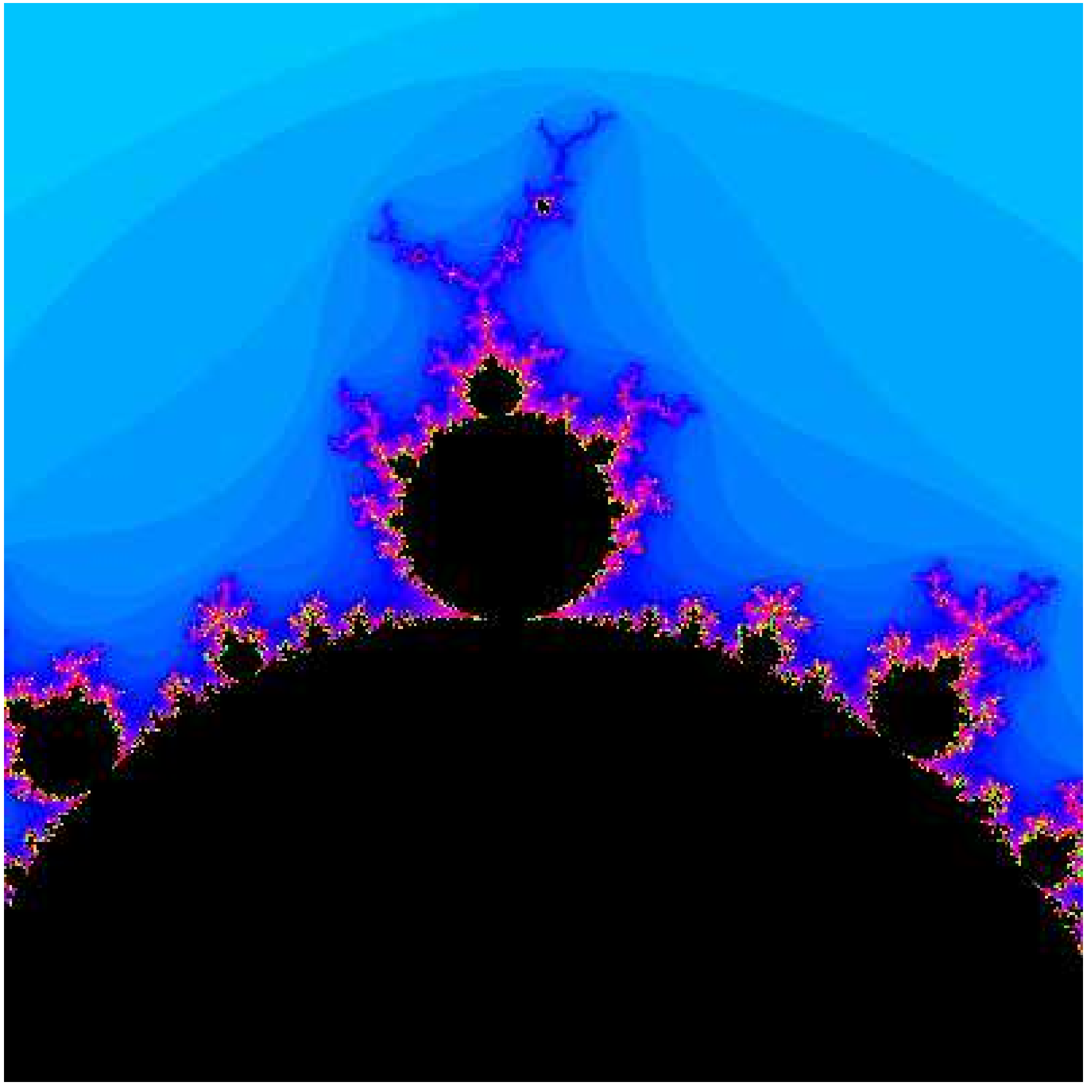
C code from
http://astronomy.swin.edu.au/~pbourke/fractals/

Mandelbrot set I



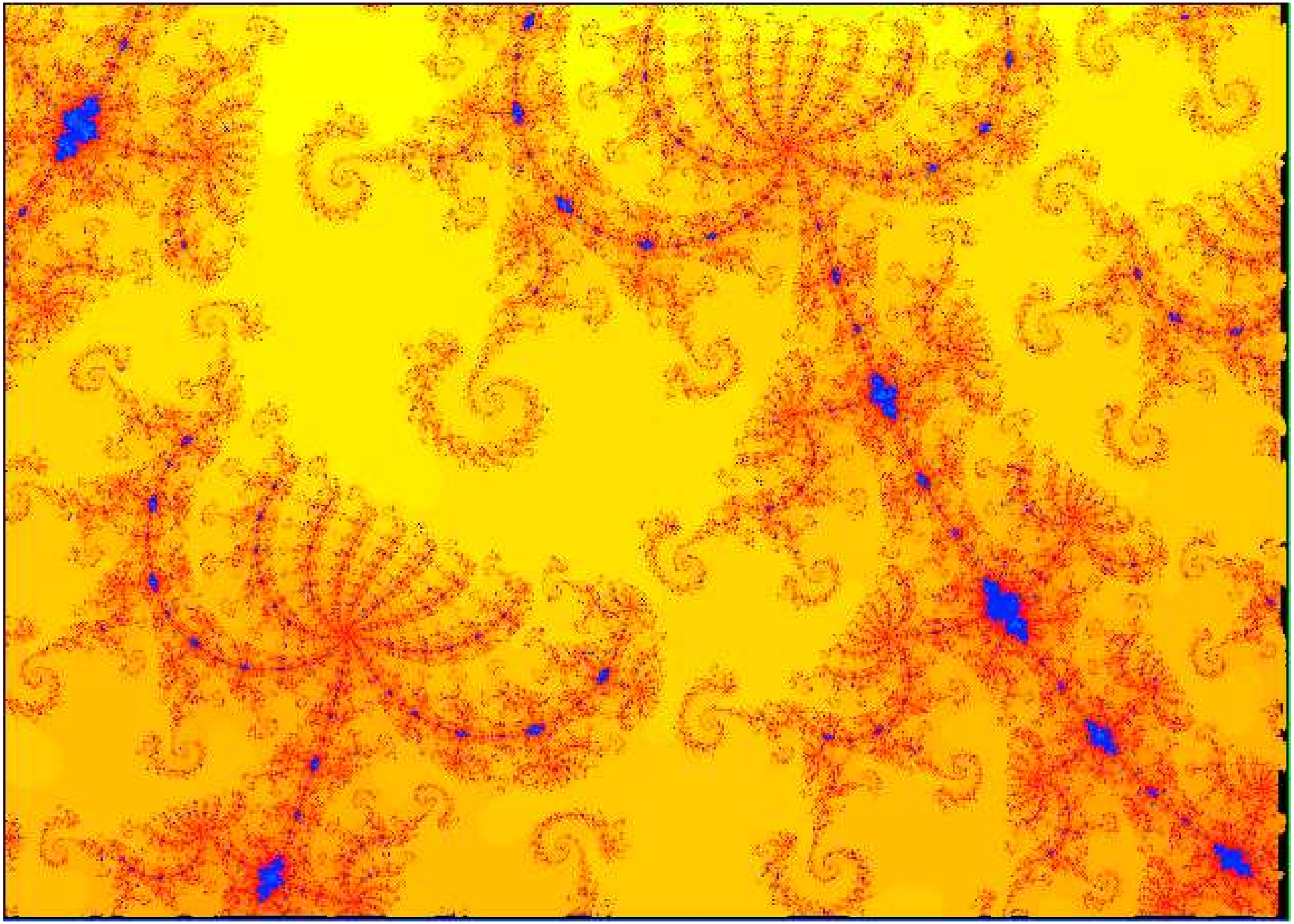
http://aleph0.clarku.edu/~djoyce/julia/julia.html

Mandelbrot set II



http://aleph0.clarku.edu/~djoyce/julia/julia.html

Mandelbrot set III



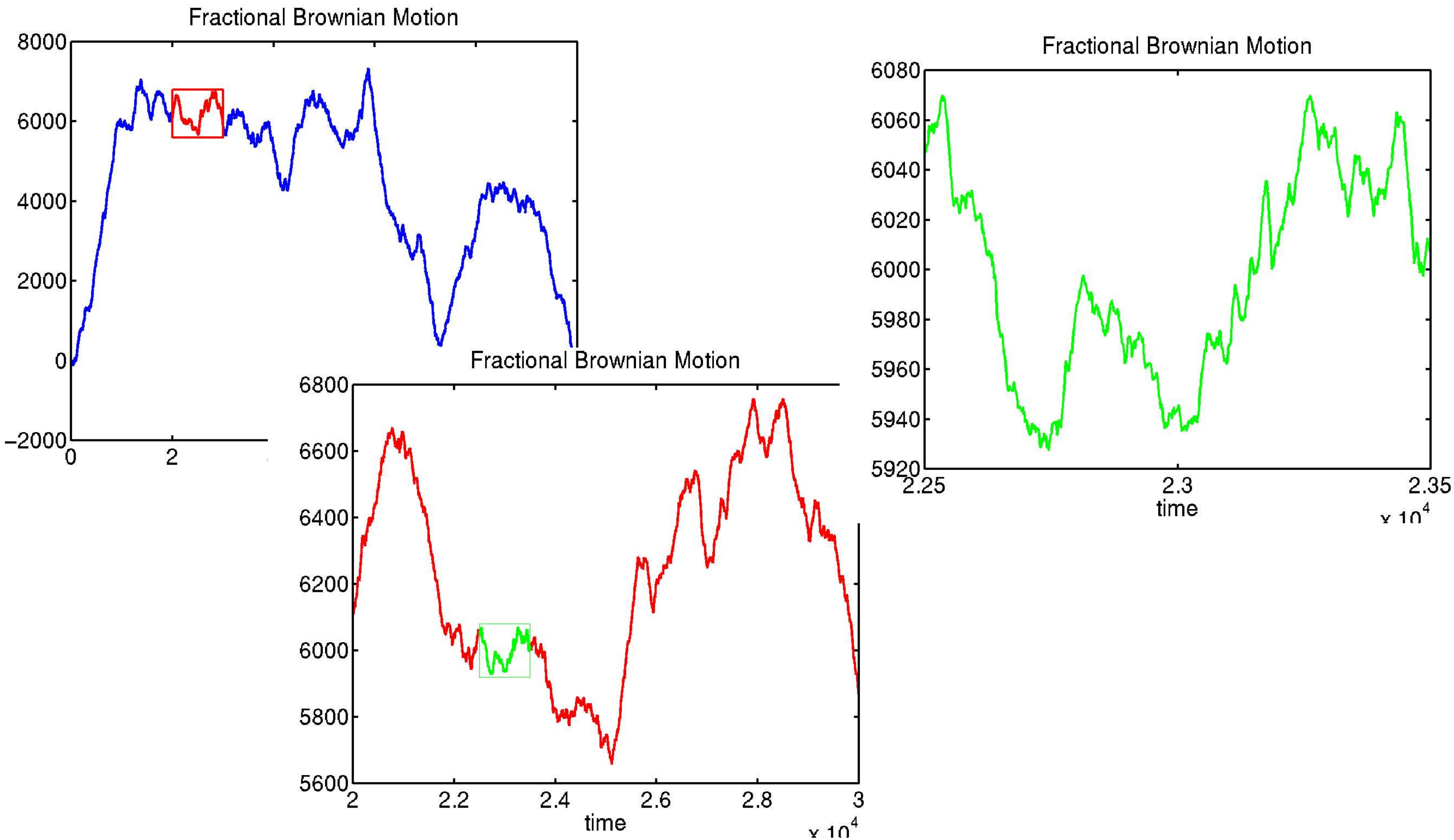
http://www.softsource.com/softsource/fractal.html

Statistical Self-similarity

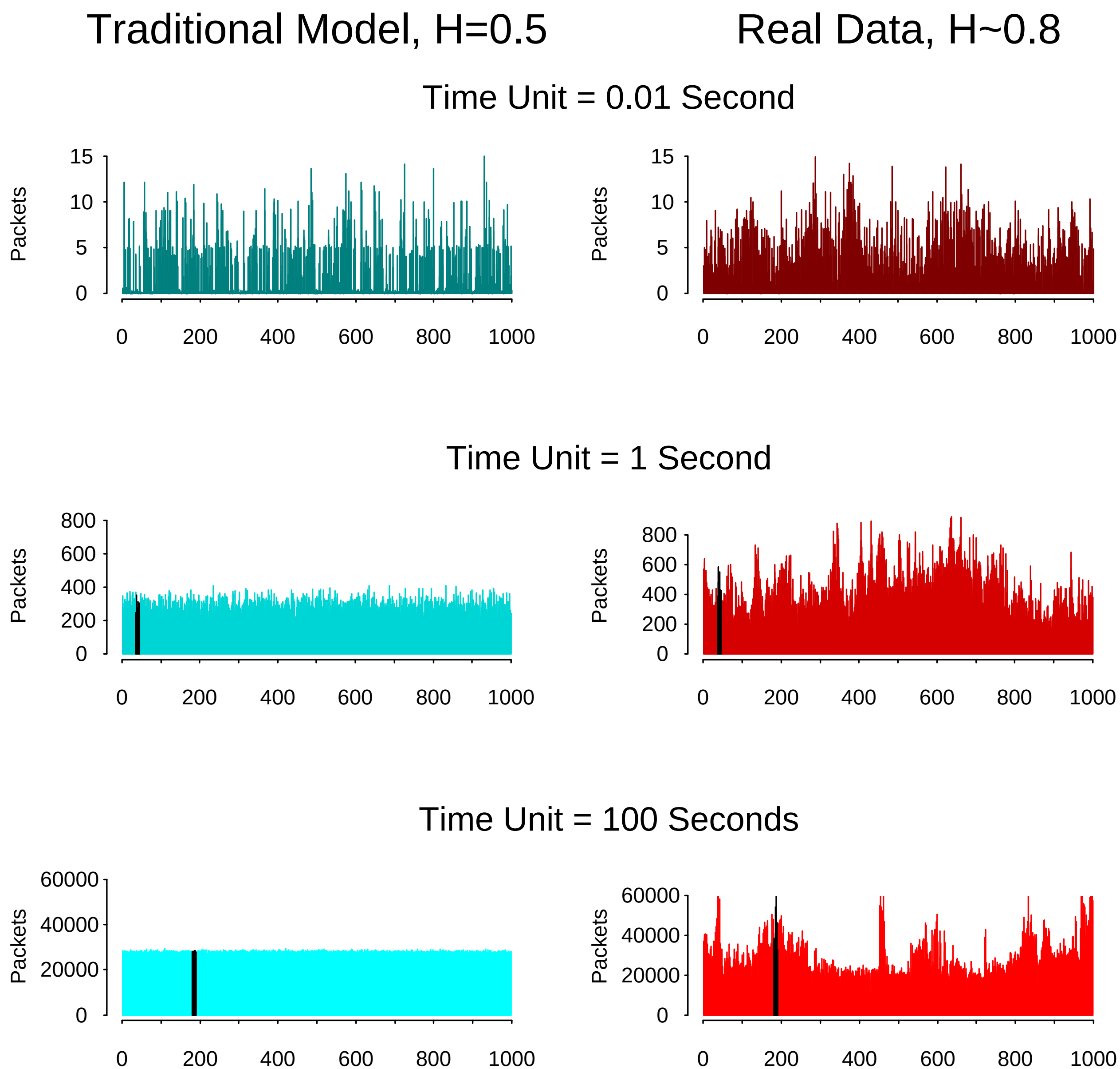
Statistical Self-similarity (SS)

- this is not a course on fractals
- Fractals (such as above) are deterministic
- we are interested in statistical properties of traffic
- look for **statistical** self-similarity

Statistical Self-similarity



Ethernet traffic



Random Processes

Random processes

A discrete time random process is just a random vector $\mathbf{x} = (X_1, X_2, \ldots, X_n)$.

- in general, the x_i may have dependencies, so we need to describe the random sequence, we specify the N-th order distribution functions, for all $N \geq 1$

$$F_{\mathbf{x}}(x_n, x_{n+1}, \ldots, x_{n+N-1}) = P\{X_n \leq x_n, X_{n+1} \leq x_{n+1}, \ldots, X_{n+N-1} \leq x_{n+N-1}\}$$

- typically, we don't need to know all of this, e.g. for **White** noise, the values at each time interval are independent, so we only need to know the first order distribution functions $F_{\mathbf{x}}(x_n)$.

Densities and distributions

Where the distribution function is continuous (e.g. for well-behaved continuous random variables), we can define a density function, e.g.

$$f_X(x) = \frac{dF_X}{dx}$$

with the meaning that

$$f_X(x)\,dx = P\{X \in [x, x+dx)\}$$

We can define $f_X(x)\,dx = dF_X(x)$, where the latter term is more general (applying to badly behaved random variables too). Integrals defined WRT to $dF_X(x)$ are Lebesgue-Stieltjes integrals rather than just Lebesgue integrals.

Moments of the process

Sometimes it is enough to specify the moments of the process, e.g. the mean $\mu_X(n)$ and variance $\sigma_X^2(n)$ at time n.

$$
\begin{aligned}
\mu_X(n) &= E[X(n)] \\
&= \int_{-\infty}^{\infty} x \, dF_X(x_n) \\
&= \int_{-\infty}^{\infty} x f_X(x_n) \, dx, \quad \text{where this is defined} \\
\sigma_X^2(n) &= \mathrm{Var}[X(n)] \\
&= E\left[(X(n) - \mu_X(n))^2\right] \\
&= \int_{-\infty}^{\infty} (x - \mu_X(n))^2 \, dF_X(x_n)
\end{aligned}
$$

Can extend definition to the n-th central moment.

Covariance

The **covariance** of two random variables X and Y if defined by

$$\mathrm{Cov}\{X,Y\} = E\left[(X - E\left[X\right])(Y - E\left[Y\right])\right]$$

It tells us about second-order correlations between X and Y.

The **auto-covariance** of a process is

$$R_{XX}(n;k) = \mathrm{Cov}\{X(n), X(n+k)\}$$

and this tells us about correlations between the process at different times n, and different **lags** k.

Stationarity

- A process is **strictly stationary** if all of its distribution functions are invariant under time shifts, e.g.

$$F_{\mathbf{x}}(x_n, x_{n+1}, \ldots, x_{n+N-1}) = F_{\mathbf{x}}(x_{n+k}, x_{n+k+1}, \ldots, x_{n+k+N-1})$$

- a process is called **wide-sense**, or **weakly**, or **second-order** stationary if its mean, variance and auto-covariance are time-shift invariant, e.g.

$$\begin{aligned} \mu_X(n) &= \mu_X \\ \sigma_X^2(n) &= \sigma_X^2 \\ R_{XX}(n;k) &= R_{XX}(k) = \gamma(k) \end{aligned}$$

for all n.

Auto-correlation

The **auto-correlation** of a process is

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

- not a great term, but standard usage

Marginal distribution

The marginal distribution of a stationary process is defined by the distribution of X_n, e.g.

$$F_{\mathbf{x}}(x_n)$$

which will be identical for all values of n for a stationary process.

Examples:

- Bernoulli process: the marginal distribution takes values 0 or 1 with probabilities p and $1-p$, respectively.
- random dice rolls: the marginal distribution is uniform on $\{1,2,3,4,5,6\}$.

Gaussian processes

- are processes with a Gaussian marginal distribution

$$f_{\mathbf{x}}(x_n) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right)$$

- completely characterized by mean, variance and auto-covariance
- hence the value of second-order stationarity!
- Gaussian processes are the "linear-time invariant" processes of the noise world
  - simple, tractable, sometimes reasonable
  - Central Limit Theorem: sums of well behaved random variables tend towards Gaussian distributions

Useful fact

If X and Y are independent Gaussian random variables, then their sum $X+Y$ (and difference $X-Y$) will also be Gaussian, with mean and variances

$$
\begin{aligned}
\mu_{X+Y} &= \mu_X + \mu_Y \\
\mu_{X-Y} &= \mu_X - \mu_Y \\
\sigma^2_{X+Y} &= \sigma^2_X + \sigma^2_Y \\
\sigma^2_{X-Y} &= \sigma^2_X + \sigma^2_Y
\end{aligned}
$$

We could easily generalize this to include correlations, with the only effect being a modification to σ^2_{X+Y} and σ^2_{X-Y}.

Also note that $\mu_{\alpha X} = \alpha\mu_X$, and $\sigma^2_{\alpha X} = \alpha^2\sigma^2_X$.

Spectral density

The **spectral density** of a random process is given by the Fourier Transform of the Autocovariance function.

$$f(\lambda) = \sum_{h=-\infty}^{\infty} e^{-i2\pi h\lambda} \gamma(h)$$

$$\gamma(k) = \int_{-1}^{1} e^{i2\pi k\lambda} f(\lambda)\, d\lambda$$

- $f(\lambda)$ is non-negative
- $f(\lambda)$ is even
- $f(\lambda)$ is the Fourier transform of the autocovariance.
- the autocovariance is the inverse FT of $f(\lambda)$

White noise

Typically, everyone assumes noise is “white”, or “uncoloured”.

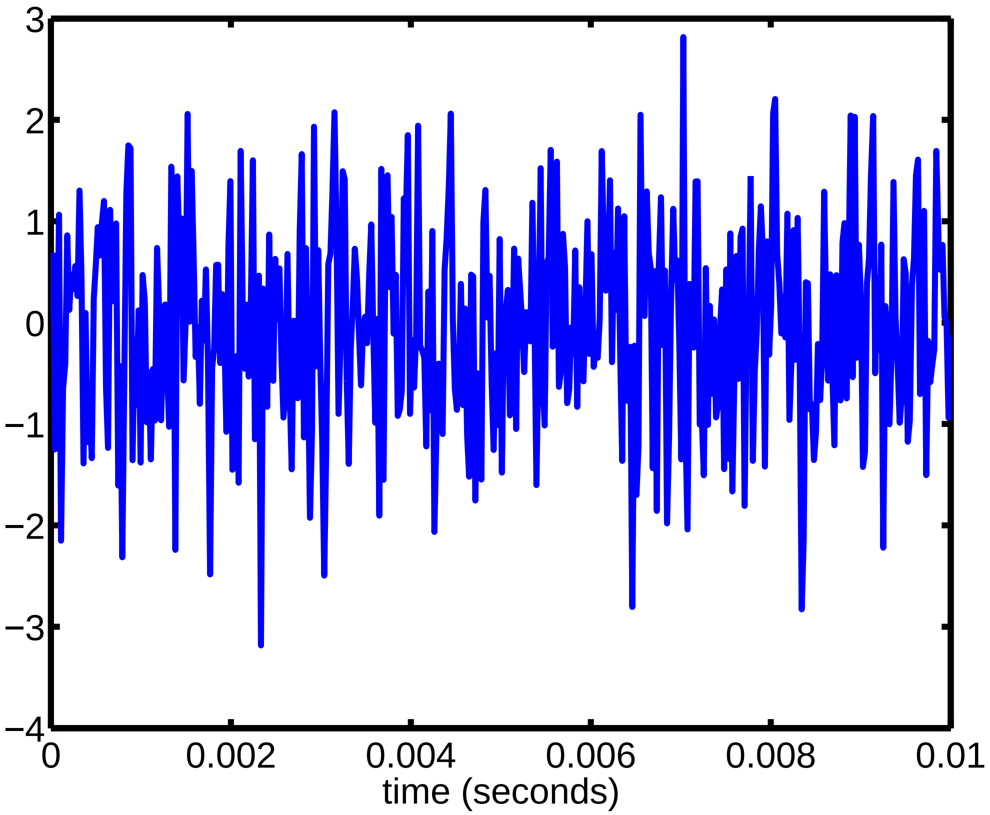
- Gaussian (typically implied, though not necessary)
- its spectral density is flat
  i.e., the noise includes all frequencies (up to f_s)

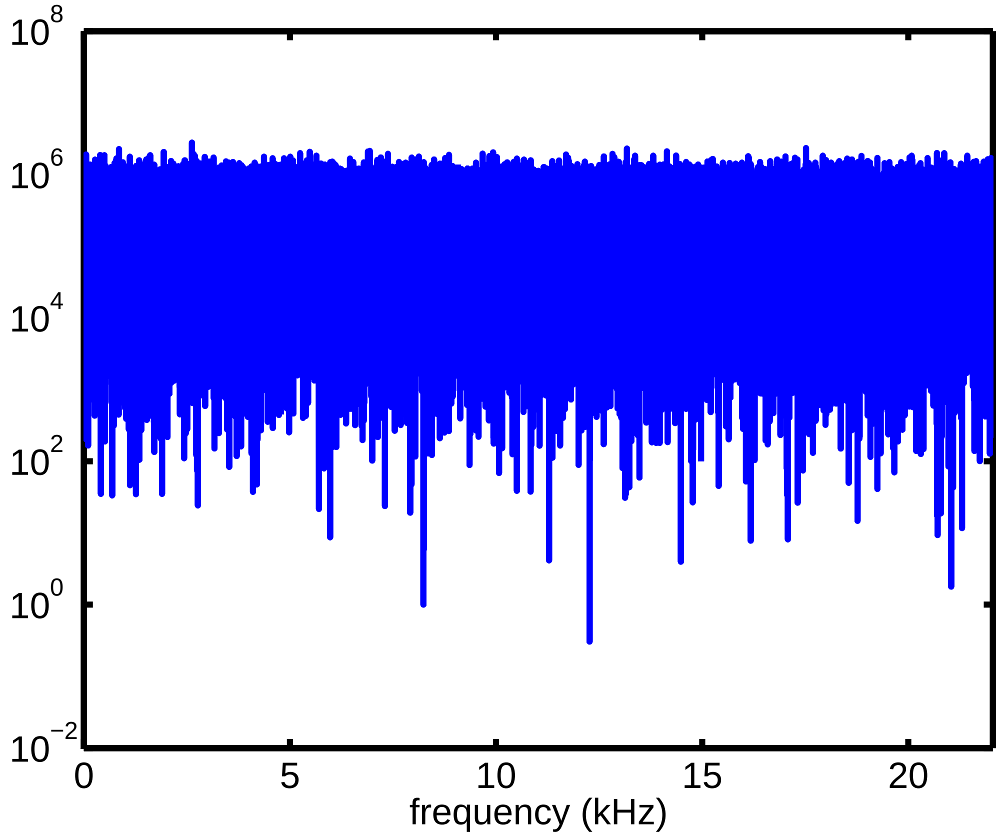
$$f(\lambda) = \sigma^2$$

- Uncorrelated (same as independent for Gaussian)
  - follows from duality of the spectrum and auto-covariance, i.e. flat spectrum implies delta function (at zero) for auto-covariance, and so the autocovariance is zero at non-zero lags.

Example: white noise

White Gaussian noise





Statistical Self-similarity

Statistical Self-similarity

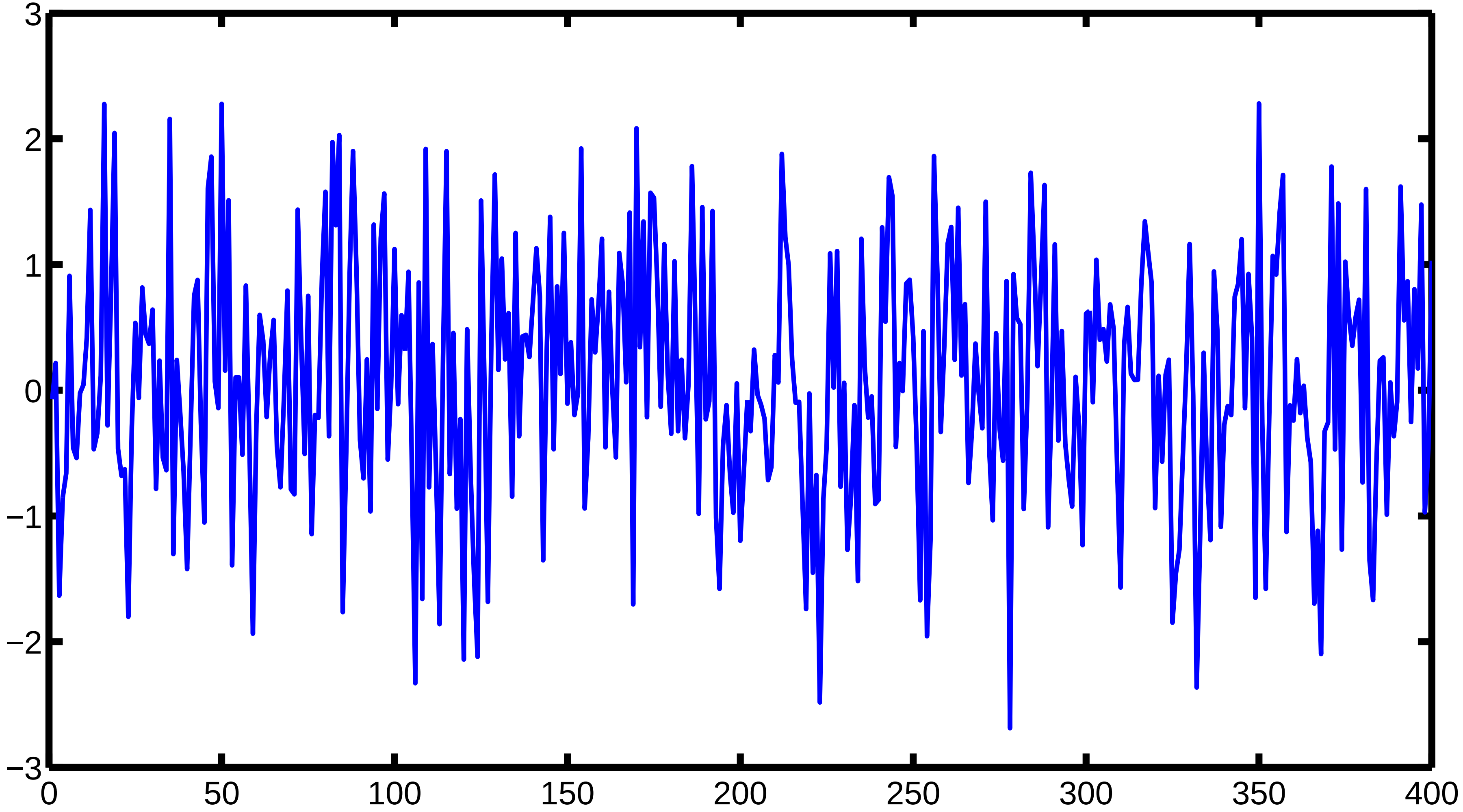
SS block aggregation definition
(another definition exists)

We define he aggregated time series $\{X_k^{(m)}\}$ at level m by

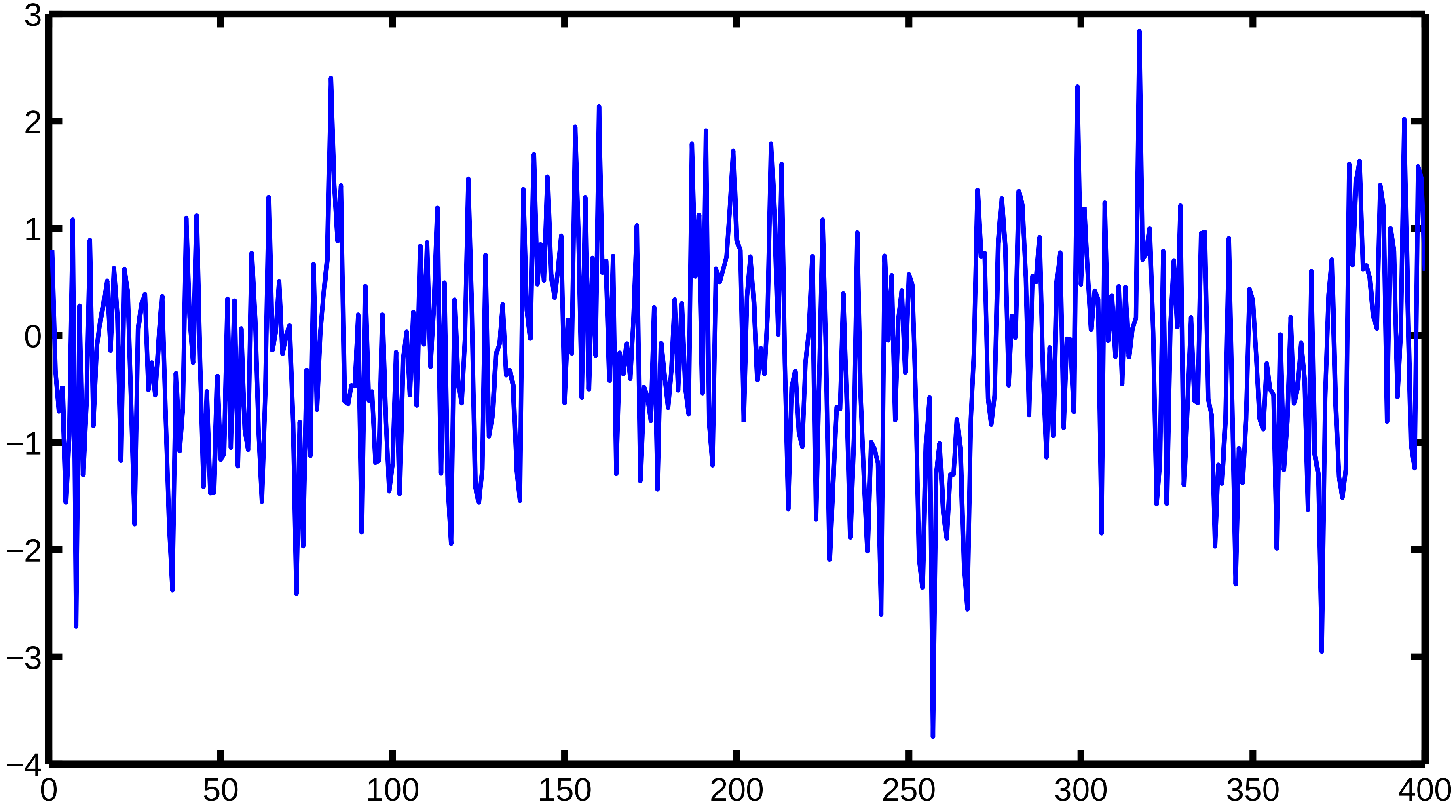
$$X_k^{(m)} := \frac{X_{(k-1)m+1} + \cdots + X_{km}}{m}.$$

A stationary time series $X = \{X_1, X_2, \ldots\}$ is called self-similar with Hurst parameter H if, for all m, the aggregated process $m^{1-H}X^{(m)}$ has the same distributions as X.

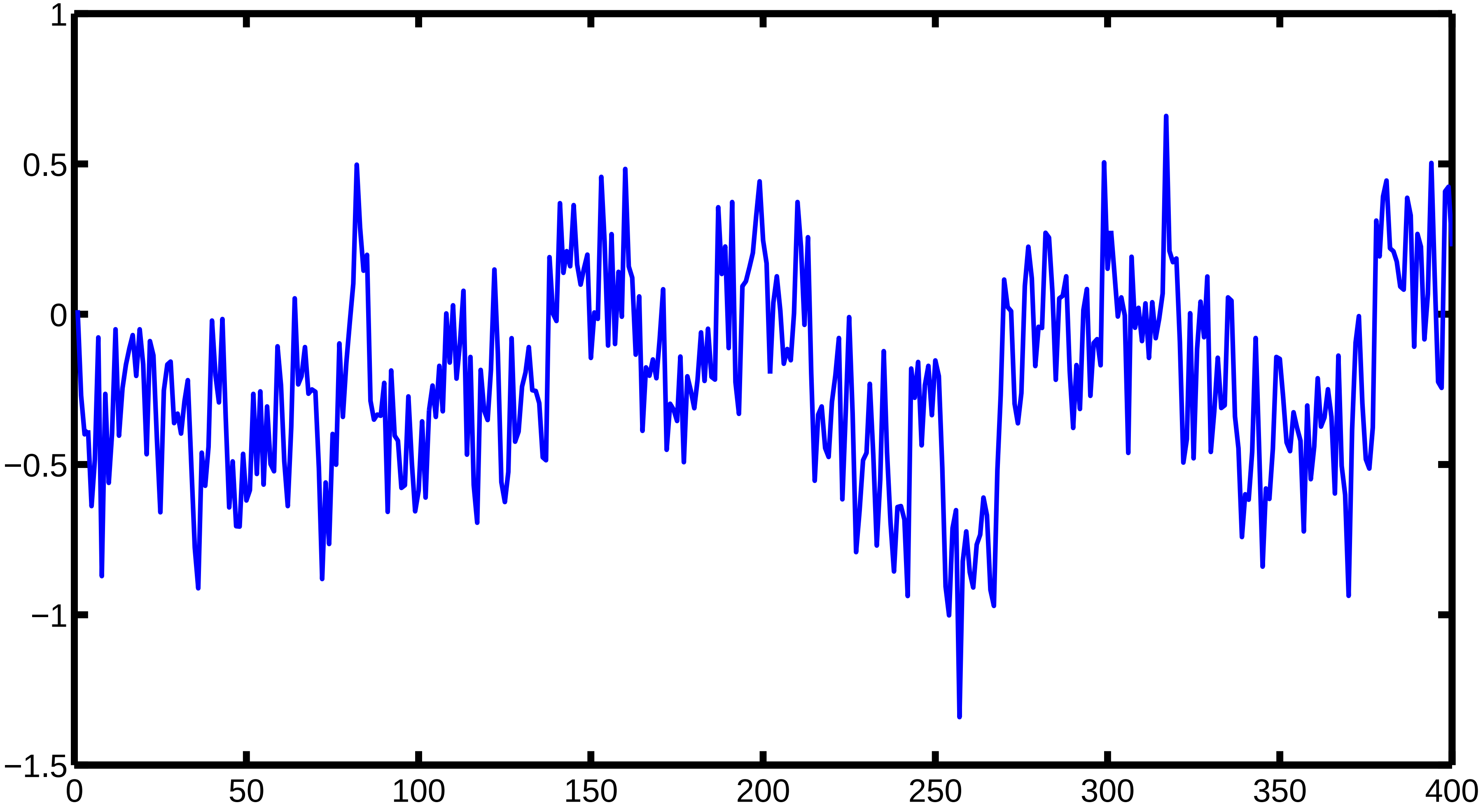
Example fGN: ($H = 0.5$)



Example fGN: $(H = 0.75)$



Example fGN: ($H = 0.99$)



Properties of Self-Similar Process

- Stationary so $\mathbb{E}X_i = 0$, Var$X_i = \sigma^2$ (constant).
- Cov(X_i, X_{i+k}) depends only on the lag k and is given by

$$\gamma(k) = \frac{1}{2}\sigma^2\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right).$$

- Cov$(X_i^{(m)}, X_{i+k}^{(m)})$ is given by

$$\gamma(k) = \frac{1}{2}m^{2(H-1)}\sigma^2\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right).$$

- Asymptotic behaviour of the autocorrelation

$$\lim_{k\to\infty}\frac{\rho_k}{k^{2(H-1)}} = H(2H-1).$$

- The variance varies with the aggregation level as

$$\text{Var}X^{(m)} = m^{2(H-1)}\sigma^2,$$

Asymptotic Statistical Self-similarit

Alternative definitional property

$$\rho^{(m)}(k) = \rho(k)$$

Useful because we can define **asymptotic self-similarity** by the limit as $m \to \infty$

$$\begin{aligned} \rho^{(m)}(1) &\to 2^{2H-1} - 1 \\ \rho^{(m)}(k) &\to \frac{1}{2}\delta^2\left(k^{2H}\right) \end{aligned}$$

where for some function $f(k)$

$$\delta^2(k) = f(k-1) - 2f(k) + f(k+1)$$

for example $f(k) = |k|$

$$\delta^2(k^{2H}) = |k-1|^{2H} - 2|k|^{2H} + |k-1|^{2H}$$

Long-range dependence

Long-range dependence (LRD) for stationary process

- LRD = slow (power-law) decay in the autocovariance

$$\gamma_X(k) \sim c_\gamma |k|^{-(1-\alpha)}$$

  as $k \to \infty$, for some $\alpha \in (0,1)$

- can define LRD by regularly varying decay

$$\gamma_X(k) = L(k)|k|^{-(1-\alpha)}$$

  for some $\alpha \in (0,1)$ and slowly varying function L

- both results imply for all N

$$\sum_{k=N}^{\infty} \gamma_X(k) \to \infty$$

  this is sometimes used as an alternative definition

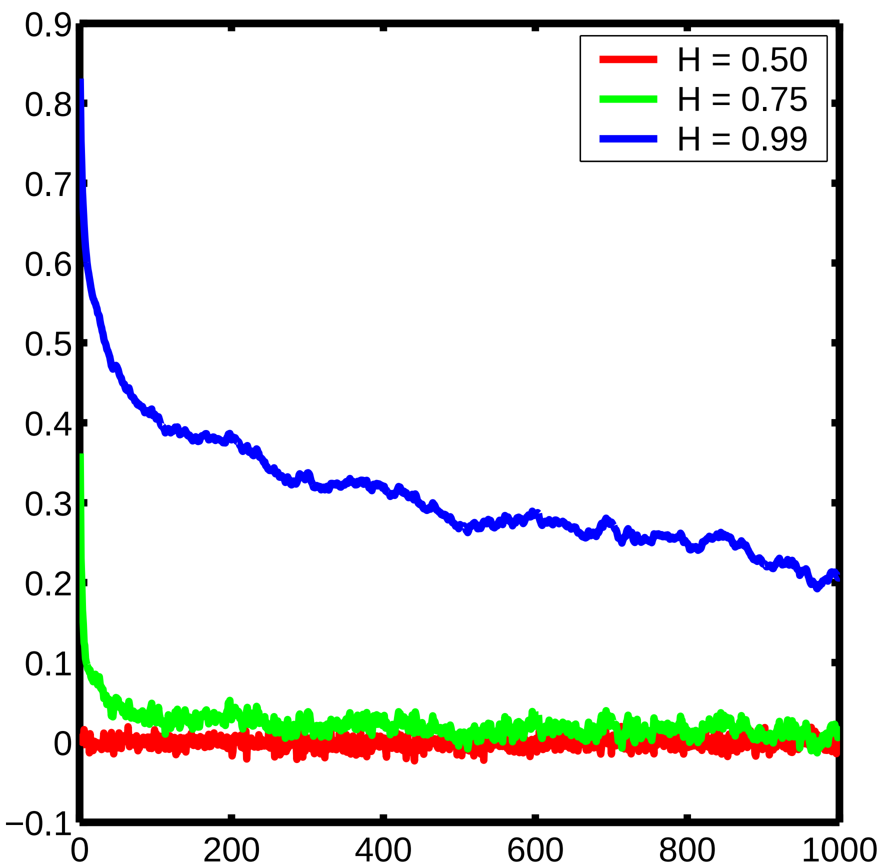
- also called **long-memory process**

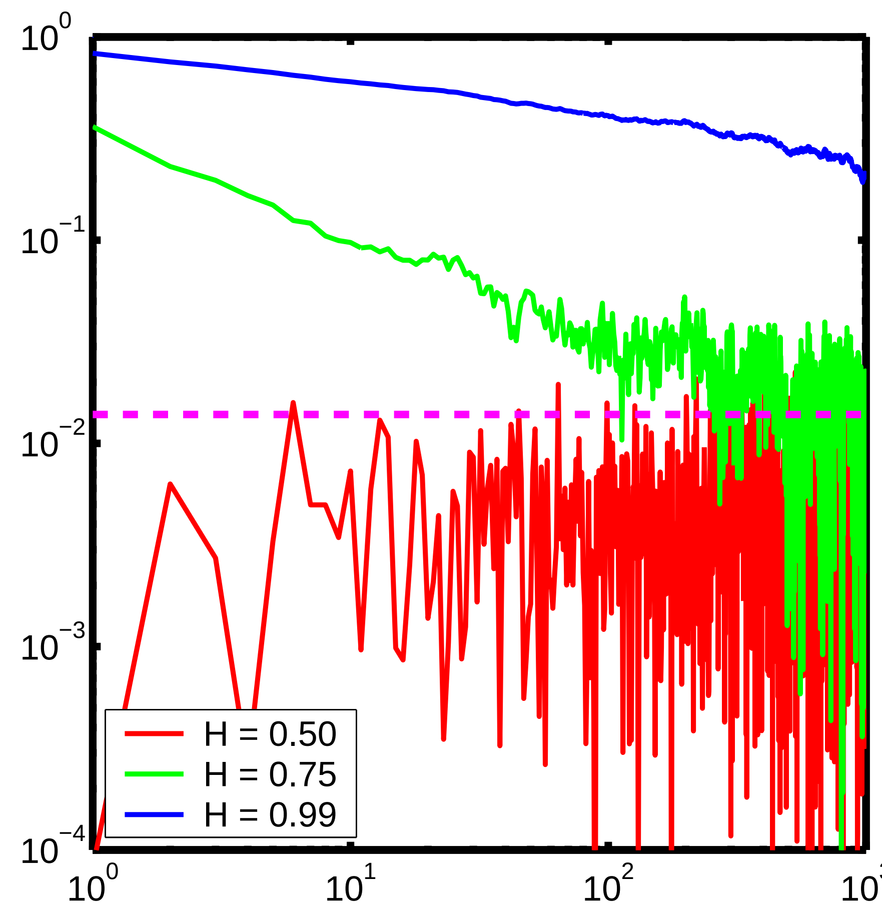
LRD and SS

Notice that self-similarity implies LRD with

$$\alpha = 2H - 1$$

for $0.5 \le H < 1$, and $0 \le \alpha < 1$





LRD in the frequency domain

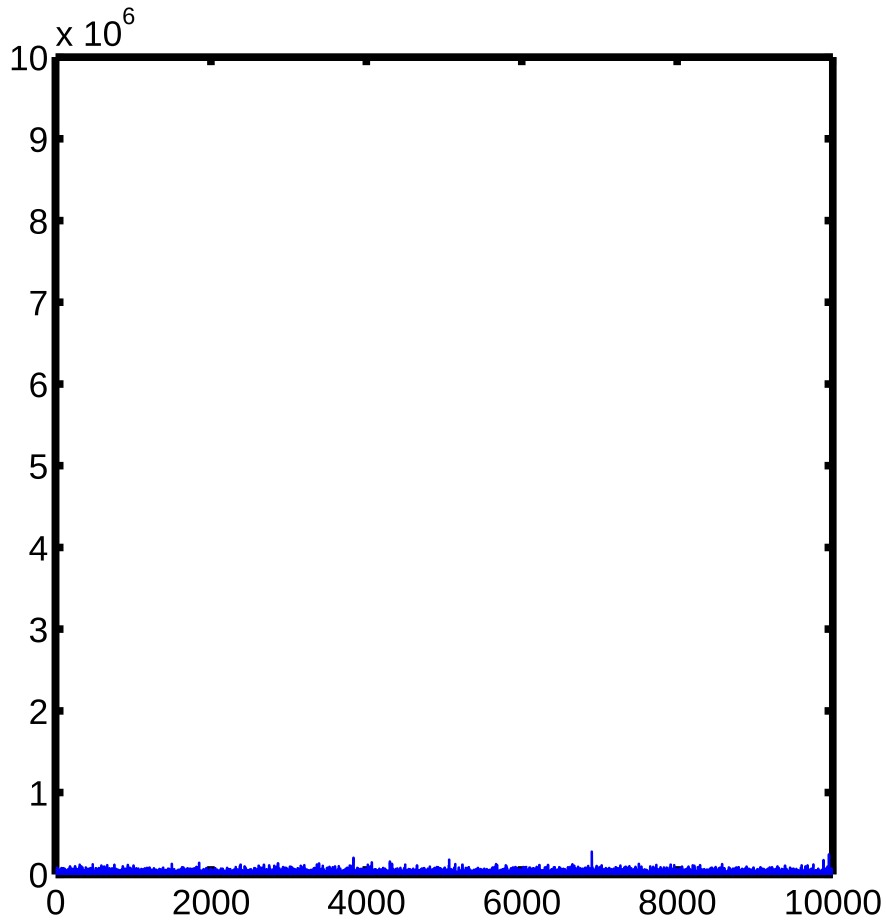
Long-range dependence (LRD) can also be defined in the frequency domain using the Fourier transform of the autocovariance

$$f_x(s) \sim c_f |s|^{-\alpha}, |s| \to 0$$

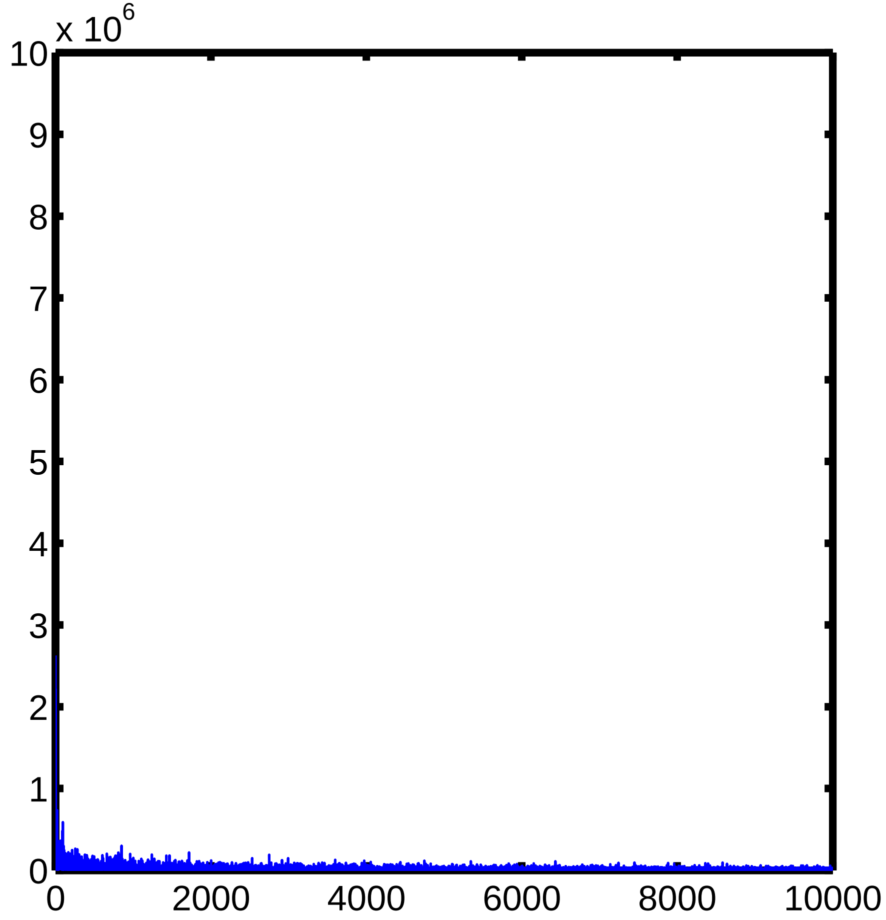
When $\alpha = 1$ we get **1/f noise**, but the term is often applied to the range of values of $\alpha = 2H - 1$.

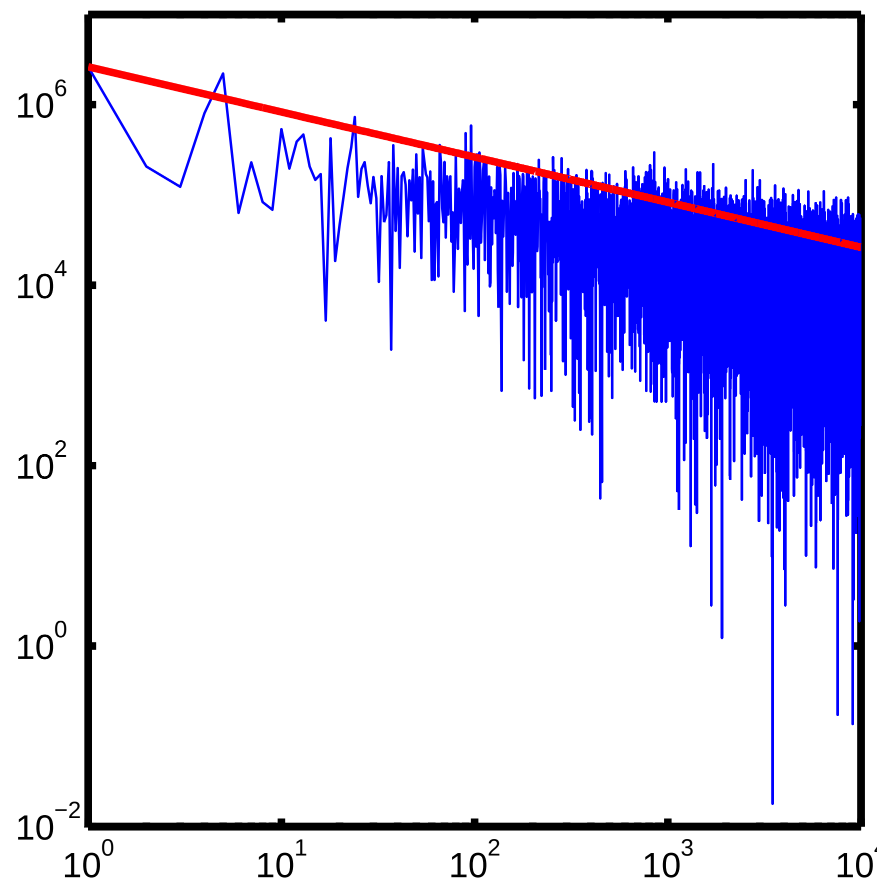
- frequency spectrum of white noise is flat
- frequency spectrum of Brownian motion is $1/f^2$
- frequency spectrum of "pink" noise is $1/f$

Example fGN spectrum ($H = 0.5$)

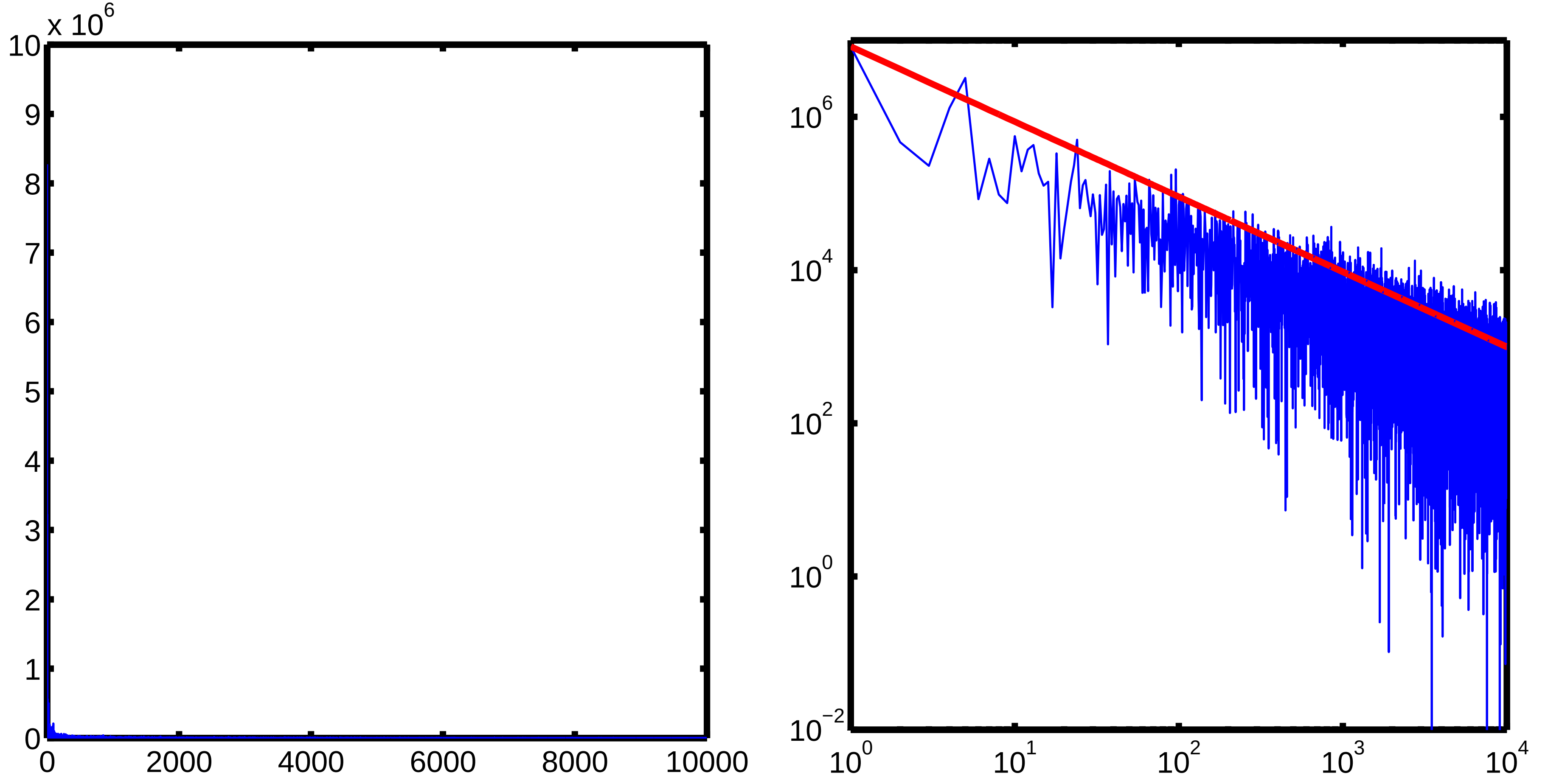


Example fGN spectrum ($H = 0.75$)





Example fGN spectrum ($H = 0.99$)



1/f noise

LRD and SS are also seen elsewhere

- cardiac rhythms (in healthy hearts)
- hydrological data (rainfall, and river flow)
  - Hurst's early work was actually in Nile river data
- music seems to have similar characteristics
- turbulence
- chaotic processes in general

LRDc vs LRDi

LRD for counts vs LRD for intervals

- Point process can look at correlations in
  - intervals (actual intervals between points)
  - counts (number of points in certain time intervals)
- we can look for LRD in either of these two
  - LRDi = LRD correlations in intervals
  - LRDc = LRD correlations in counts
- are they related
  - LRDi $\Rightarrow$ LRDc
  - LRDc $\not\Rightarrow$ LRDi
    - you can have a renewal process that is LRDc

Link between LRD and SS

Central Limit Theorem again – another condition of the CLT is that the data not be correlated.

- Often this is approximately true, and even more so when you block the data into groups.
- With LRD data it is never true, no matter how large a block size you choose.
- Hence CLT does not apply – hence as we aggregate, we can see SS properties
- Hence LRD leads to SS

Central Limit Theorem for LRD

The standard Central Limit Theorem does not apply with LRD correlations in the data, so we get an alternate, e.g. when we take the mean of a set of n LRD random variables $\bar{X}_{\mathrm{LRD}} = \frac{1}{n}\sum_{i=1}^{n} X_i$ the variance of the mean estimate

$$\mathrm{var}\left(\bar{X}_{\mathrm{LRD}}\right) \rightarrow \frac{c_\gamma n^{2H-2}}{H(2H-1)}.$$

as $n \rightarrow \infty$

- results, under aggregation, in SS
- estimates of the mean of LRD traffic will converge much slower than for IID traffic

LRD and SS

Also (almost) universally observed in data traffic

- **Ethernet:** W.E.Leland , M.S.Taqqu, W.Willinger, D.V.Wilson, On the self-similar nature of Ethernet traffic (extended version), IEEE/ACM Transactions on Networking, v.2 n.1, p.1-15, Feb. 1994
- **WAN:** V.Paxson, S.Floyd, Wide area traffic: the failure of Poisson modeling, IEEE/ACM Transactions on Networking, v.3 n.3, p.226-244, June 1995
- **CSS7:** D.E.Duffy and A.A.McIntosh and M.Rosenstein and W.Willinger, “Statistical Analysis of CCSN/SS7 Traffic Data from Working CCS Subnetworks”, IEEE Journal on Selected Areas in Communications, 12, 3, 1994,
- **ATM:** Judith L. Jerkins and Jonathan L. Wang, “A measurement analysis of ATM cell-level aggregate traffic”, IEEE GLOBECOM'97, 1997

LRD and SS

Also (almost) universally observed in data traffic

- **Web traffic:** M.E.Crovella, A.Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes", Proceedings of the 1996 ACM SIGMETRICS, p.160-169, May, 1996, Philadelphia, PA, USA
- **VBR video:** M.W. Garrett and W. Willinger, "Analysis, Modeling, and Generation of Self-Similar VBR Video Traffic", Proceedings of ACM Sigcomm, pp. 269–280, London, UK, 1994

Models

fractional Gaussian Noise

fGN (fractional Gaussian Noise) is stationary Gaussian process X_t with mean μ, variance σ^2 and autocorrelation function

$$\rho(k) = \frac{1}{2}\left(|k+1|^{2H} - |k|^{2H} + |k-1|^{2H}\right)$$

which asymptotically goes like

$$\rho(k) \sim H(2H-1)|k|^{2H-2}, \quad k \to \infty$$

so $c_\gamma = H(2H-1)$. In the frequency domain,

$$f_x(s) \sim c_f|s|^{1-2H}, \quad |s| \to 0$$

where now

$$c_f = \sigma_Z^2 \cdot 2(2\pi)^{1-2H} H(2H-1)\Gamma(2H-1)\sin(\pi(1-H)),$$

where $\Gamma(x)$ is the gamma function.

fractional Gaussian Noise

Synthesis of fGN:

- Durbin-Levinson: *generate white noise, and then impose exact correlation structure.* Slow $O(N^2)$ algorithm
- Spectral synthesis:
  - generate white noise
  - take FFT
  - multiply by desired spectrum
  - inverse FFT, to get back to time domain

Note that discrete version of continuous process is no longer exactly self-similar.

fractional Brownian Motion

The (non-stationary) Gaussian process with covariance function given by

$$\Gamma(s,t) = \frac{1}{2}\sigma^2 \left(s^{2H} - (t-s)^{2H} + t^{2H}\right),$$

variance σ^2 and expectation 0 is called fractional Brownian motion (fBM).

Note the increment process of fBM is fGN, just as the increments of BM are white noise.

FBM with $H = 0.7$ and $\sigma^2 = 1$.

fARIMA

fARIMA (fractional Auto-Regressive Moving Average)

- start with Gaussian white noise X_k
- run through an fARIMA(p,d,q) filter, e.g.

$$\Phi(B)\Delta^d Y_k = \Theta(B) X_k$$

where B is the backshift operator,

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \cdots - \phi_p B^p \\ \Theta(B) &= 1 - \theta_1 B - \cdots - \theta_q B^q \\ \Delta^d &= (1-B)^d = \sum_k \binom{d}{k} (-B)^k \end{aligned}$$

- Asymptotically SS with $H = d + 1/2$

advantages: more flexible correlations/Fourier spectrum
disadvantage: more parameters, less simple

Link between heavy-tails and LRD

Heavy-tails generate LRD, e.g. superposition of On/Off sources in the limit as $n \to \infty$.

- renormalize, i.e. reduce the rate of individual sources so that we keep $rn = const$
- take the Gaussian process that is the limit of the superposition of n sources as $n \to \infty$
- mean and variance will be fixed in the limit. The correlation structure is determined by the On/Off time distributions
  - if On or Off times have (infinite variance) power-law tail with decay parameter $c \in (1,2)$, then the superposition process will be asymptotically SS with $H = (3-c)/2$, and LRD with $\alpha = 2-c$.

Link between heavy-tails and LRD

Actually you can get a number of different processes depending on how you renormalize

- reduce rate $\rightarrow$ fGN
- reduce On time $\rightarrow$ α-stable process
- increase Off time $\rightarrow$ $M/G/\infty$

On/Off models become explanatory

- explain why we see SS and LRD in traffic

non-Gaussian processes

fGN and fARIMA are both Gaussian processes

- advantage: they can be characterized by 2nd order stats: mean, covariance and autocovariance.
- disadvantage: they all have a Gaussian marginal distribution

Other marginal distributions through the transform

$$Y_i = F^{-1}(G(X_i))$$

where $F(x)$ and $G(x)$ are the desired and Gaussian CDFs.

- gives a given marginal
- also distorts correlations
- process is no longer Gaussian, so we can't characterize completely by 2nd order stats.

Multi-fractals

See "Data networks as cascades: investigating the multifractal nature of Internet WAN traffic", A. Feldmann, A. C. Gilbert, and W. Willinger, Proceedings of ACM SIGCOMM '98, Vancouver, British Columbia, Canada, pp. 42–55, 1998.

- related to the fact that variability at small time scales is caused by different processes to long time scales, e.g. TCP congestion control, rather than offered traffic

Estimators

Estimators for Self-Similarity

- Regression on the autocovariance in log-log graph
- R/S statistic
- Aggregated Variance Method
- Whittle (approximate MLE)
- Wavelets

R/S statistic

R/S statistic

- X_i = number of bytes in time period i
- $Y_j = \sum_{i=1}^{j} X_i$

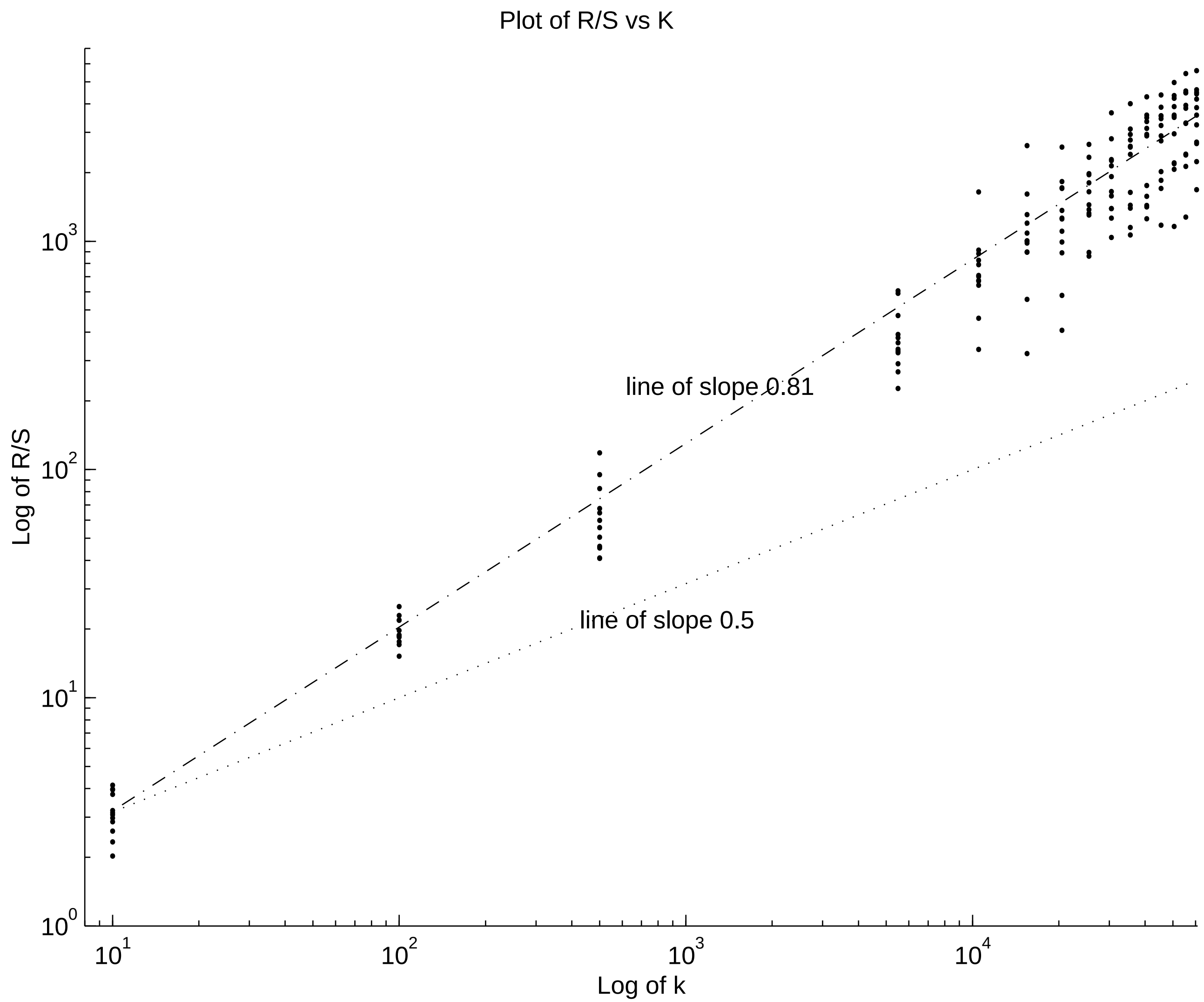
$$
\begin{aligned}
R(t,k) &= \max_{0\le i\le k}\left[Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t)\right] \\
&\quad - \min_{0\le i\le k}\left[Y_{t+i} - Y_t - \frac{i}{k}(Y_{t+k} - Y_t)\right] \\
S(t,k) &= \sqrt{k^{-1}\sum_{i=t+1}^{t+k}(X_i - \overline{X}_{t,k})^2}
\end{aligned}
$$

R/S Statistic

The R/S statistic is $Q(t,k) := \frac{R(t,k)}{S(t,k)}$

- construct a two-dimensional table of $Q(t,k)$ values for various starting times t and lags k
- plot $\log Q(t,k)$ against $\log k$ on a log-log scale
- H estimated as the slope of the line of best fit.

R/S Statistic



R/S method for the Ethernet data set

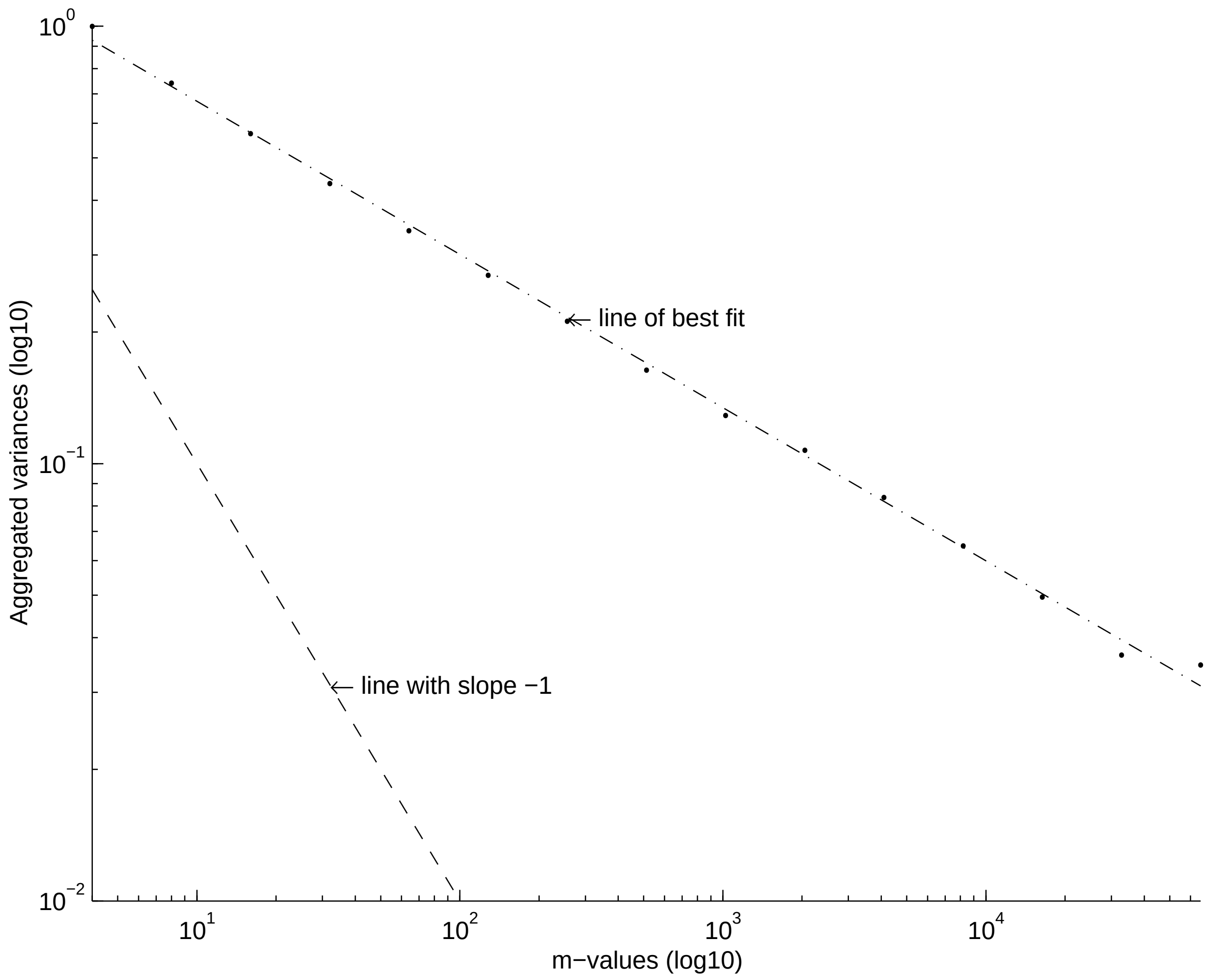
Aggregated Variance Method

Calculate the sample variance of $X^{(m)}$:

$$s^2(m) := \frac{1}{N/m}\sum_{k=1}^{N/m}\left(x_k^{(m)}\right)^2 - \left(\frac{1}{N/m}\sum_{k=1}^{N/n}x_k^{(m)}\right)^2 .$$

Plot $\log s^2(m)$ against $\log m$. Line of best fit has slope $2H-2$.

Aggregated Variance Method



Estimating H for the Ethernet data set

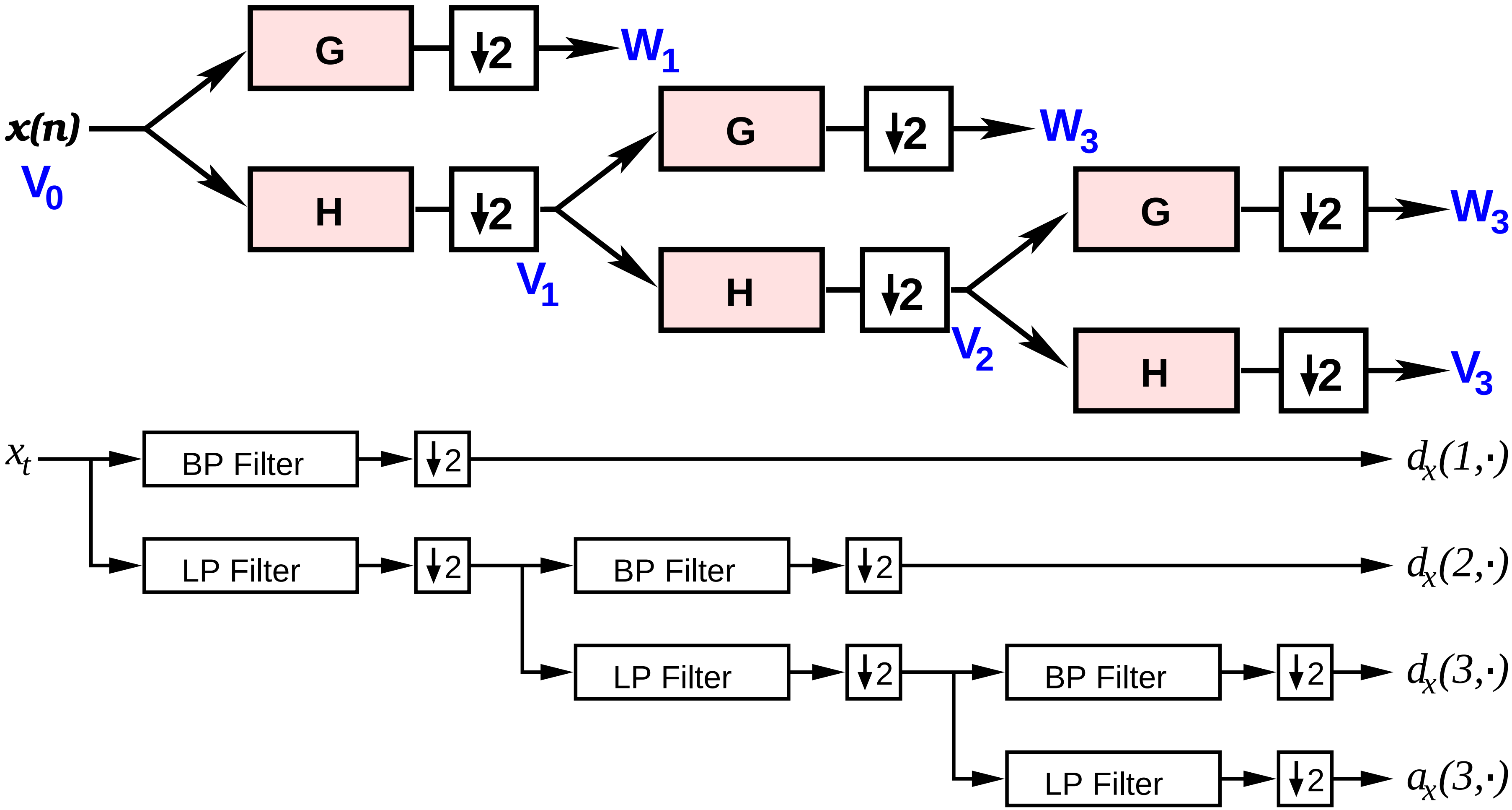
Whittle

- Assume fGN
- Gaussian process, with known autocorrelations
- for particular H we can compute the likelihood exactly
- complication is that the likehood involves inverting an $N \times N$ matrix, which is painful
- Whittle estimator uses some approximations to the likehood to make the computations tractable
- properties
  - probably most accurate estimator, if the data is actually fGN
  - assumes a model, so not robust

Wavelets: interpretation

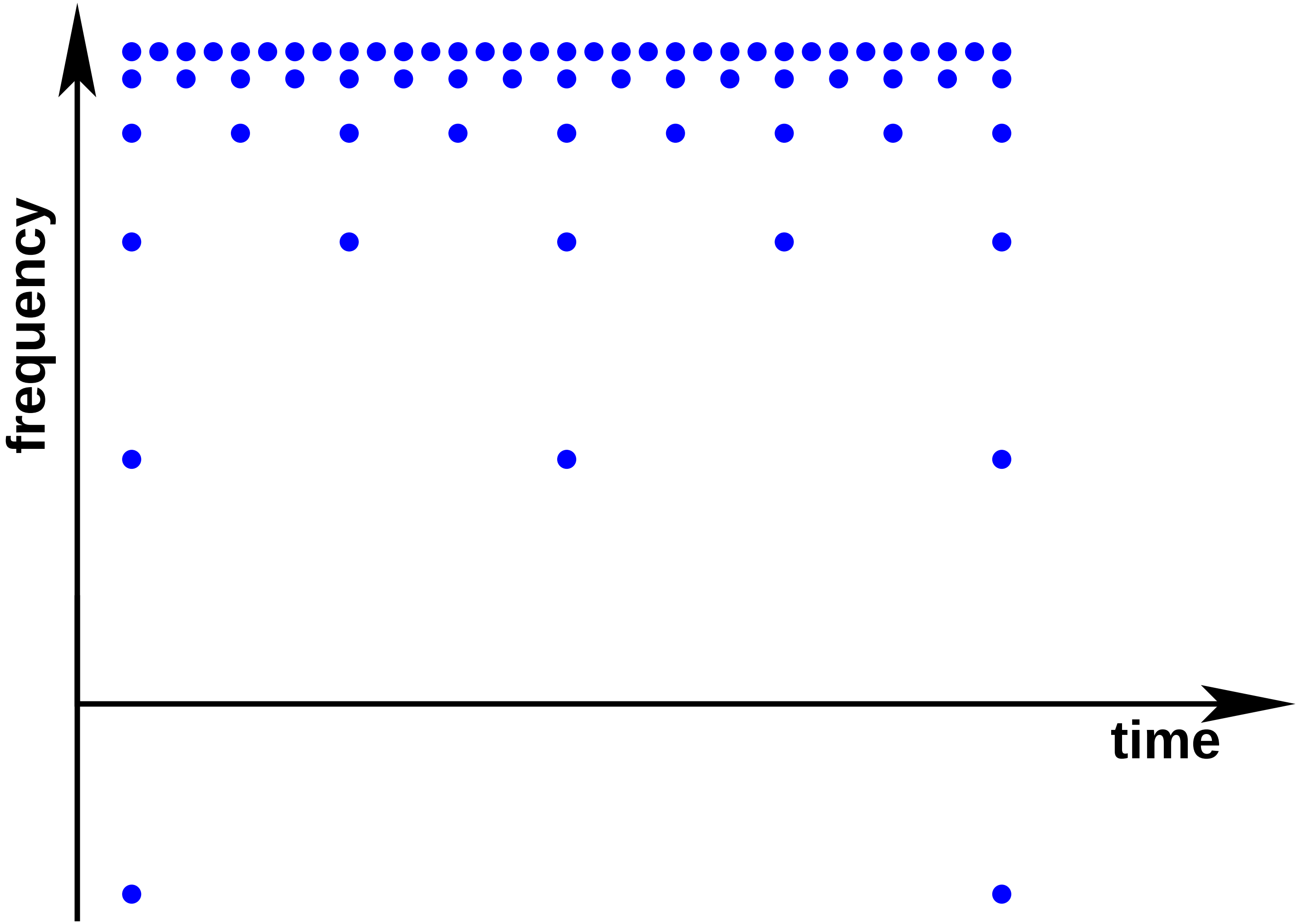
- extend aggregation idea (MRA)
  - aggregation at different scales is like approximating the data at different scales
  - data stats have known scaling properties
  - a more general way of doing multi-scale approximation is wavelets
    - has the advantage of **decorrelation** of wavelet coefficients
- sub-band filters (logarithmically placed)
  - logarithmically placed, so natural log scale arises in frequency domain.
  - sub-bands sampled at frequency appropriate to the bandwidth
  - analogy to FT, but retaining some temporal information

Pyramidal decomposition algorithm



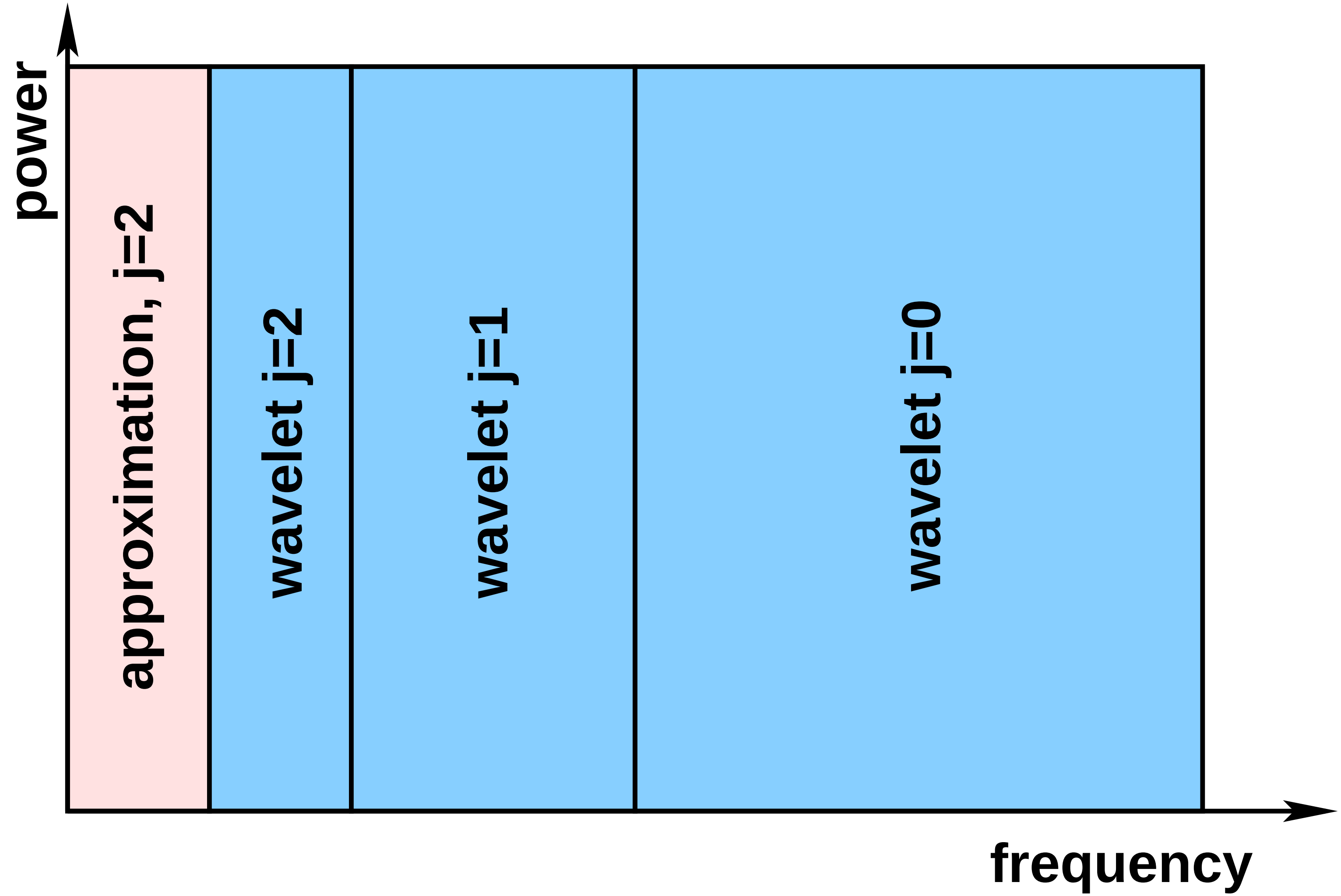
Dyadic grid

Downsampling automatically results in dyadic grid.



Ideal wavelet and scaling

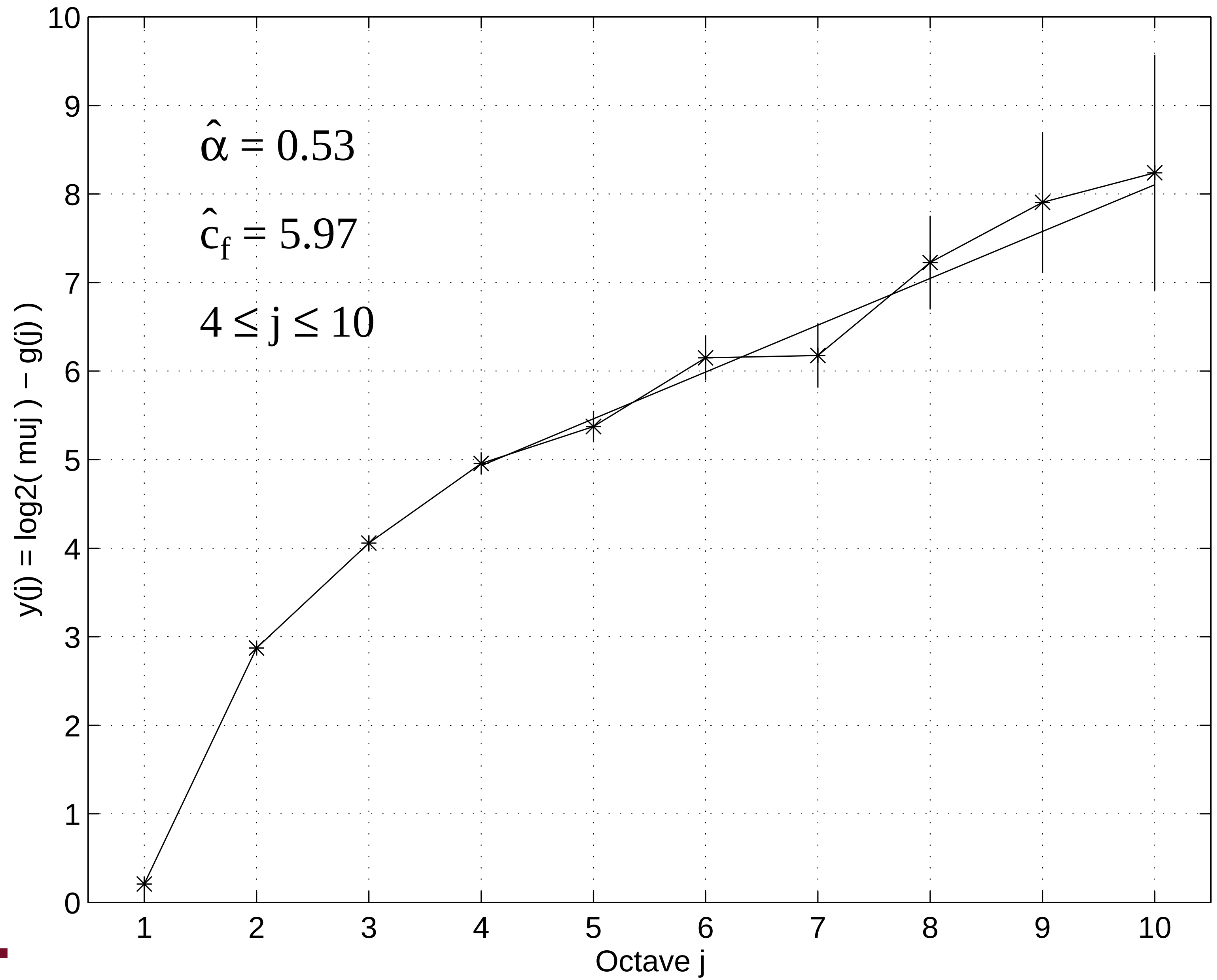
The idea (looking across frequencies or scales) is that the transform breaks frequency spectrum into bands.



Ideal wavelet and scaling

- we know the distribution of energy in each sub-band
- this translates to energy in each scale of wavelet coefficients $d_{j,k}$
- estimate the energy in each band, which will follow a power-law
- regress on a log-log graph (called a Logscale diagram)

Logscale diagram



Wavelet estimator properties

- asymptotically efficient and unbiased
  - almost as accurate as Whittle
- joint estimator of H and c_γ
- known variance of estimates
- robustness
  - non-Gaussianity
  - trends in the data
  - short-range correlative structure
  - much better than Whittle in these cases

`http://www.cubinlab.ee.mu.oz.au/~darryl/secondorder_code.html/`

Performance

Queueing with LRD

Assume the traffic follows the **Norris model** where the total traffic up to time t is

$$A(t) = mt + \sqrt{am}Z(t),$$

where $Z(t)$ is fBM, a is the traffic peakedness, and m is the mean traffic rate. The loss rate, for a system can be estimated using

$$\text{CLR} \sim \exp\left(-\frac{(c-m)^{2H}}{2am\kappa(H)^2} b^{2-2H}\right)$$

where

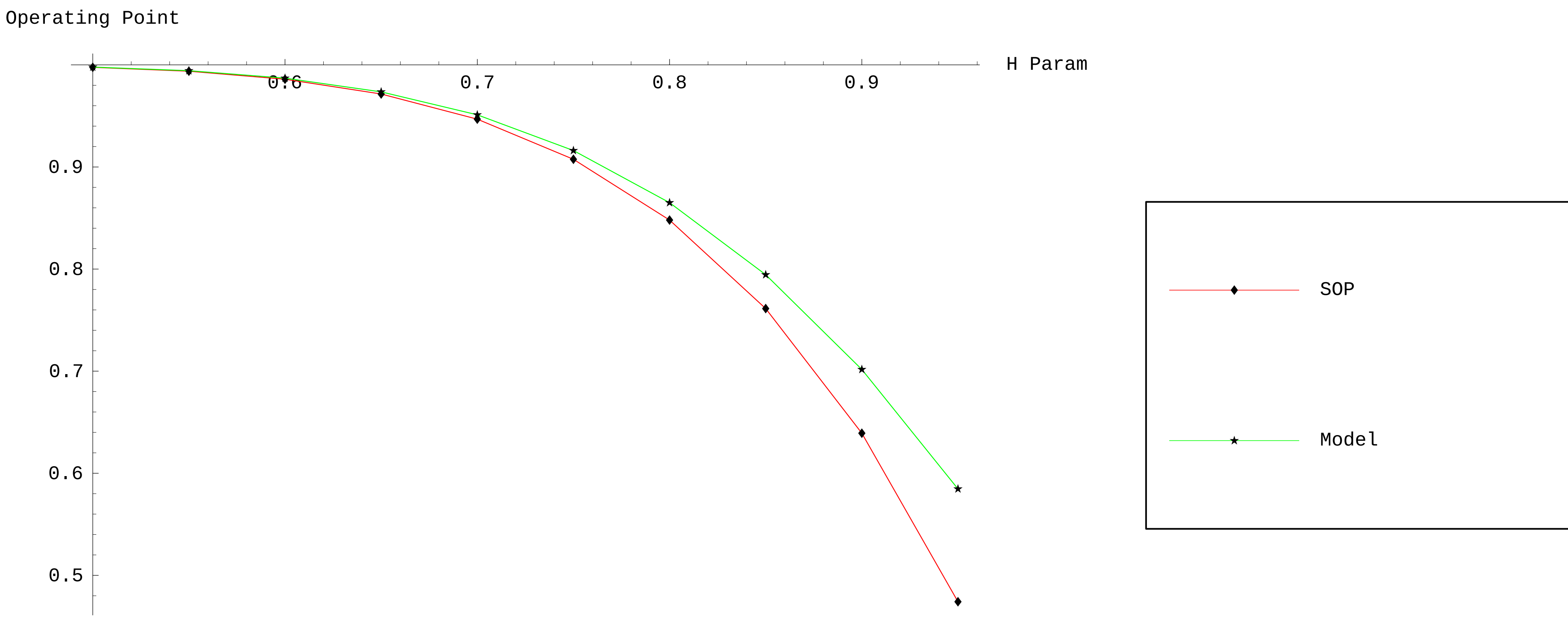
- $\kappa(H) = H^H(1-H)^{(1-H)}$
- c is the link capacity
- b is the buffer size

Example of Norris loss formula

- 155 Mpbs ATM link (capacity 353207 cells/sec)
- a of 10 cells-ms
- cell loss rate on this link below 10^{-8}
- b is 32000 cells
- Safe Operating Point, CLR below the target rate, given errors in the estimates of the traffic parameters, e.g.

$$m_{\text{model}} = m_{\text{SOP}} + 1.96\sqrt{a\, m_{\text{SOP}}\, n^{2H-2}},$$

Example of Norris loss formula



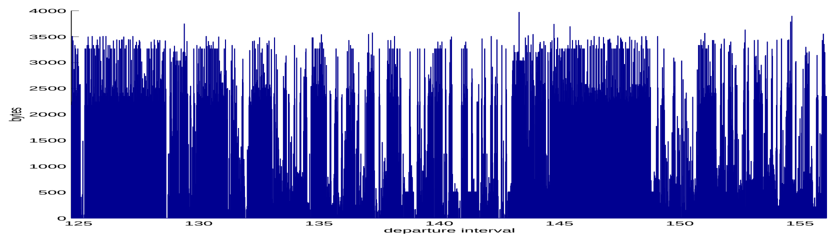
Self-similarity and Queueing

Simulation Example Model

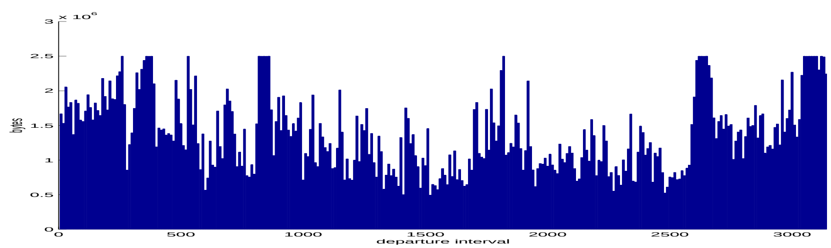
- single server queue
- deterministic service rates
- assumptions
  - only complete packets are accepted
  - no time spent in checking packets
- some of the statistics gathered
  - departure process
  - time in queue
  - queue utilization
  - queue lengths

Departure Process - 10 Mbits

Self Similar Data

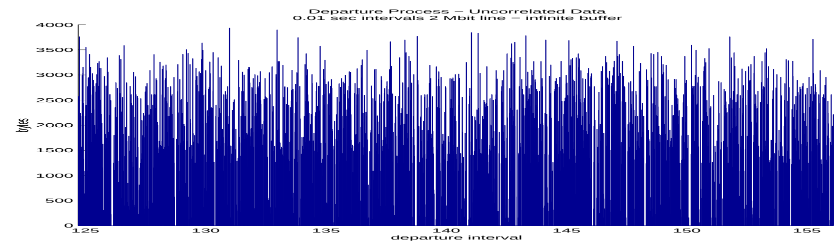


a) 0.01 second intervals

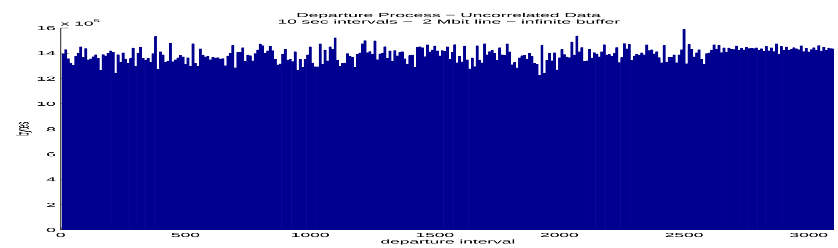


d) 10 second intervals

Uncorrelated Data



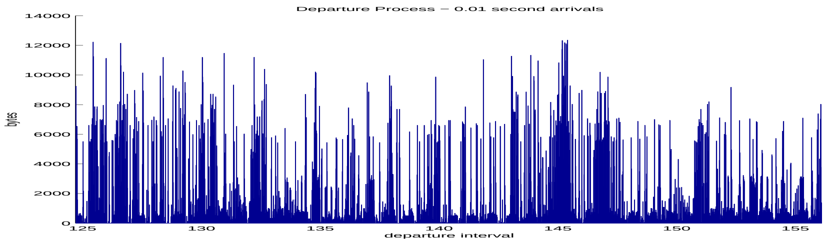
a') 0.01 second intervals



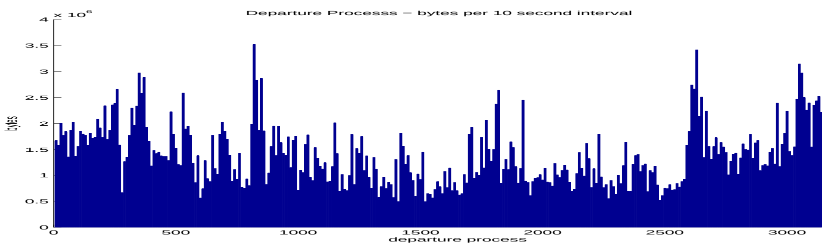
d') 10 second intervals

Departure Process 2 Mbits

Self Similar Data

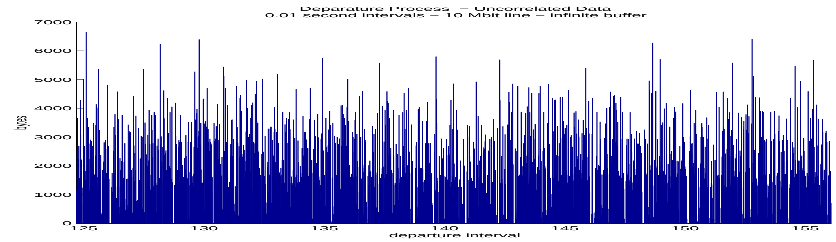


a) 0.01 second intervals

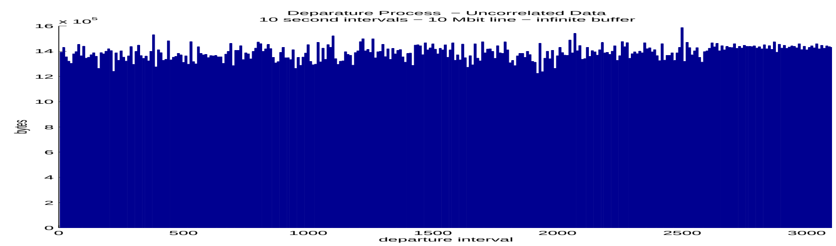


d) 10 second intervals

Uncorrelated Data



a') 0.01 second intervals



d') 10 second intervals

Departure Process

Results - 10 Mbits

	Self Similar Data	Uncorrelated Data
utilization	11.05%	11.14 %
average time in queue	0.000414	0.000387
max queue (packets)	6	6

Results - 2 Mbits

	Self Similar Data	Uncorrelated Data
utilization	55.27%	55.69 %
average time in queue	0.458592	0.004973
max queue (packets)	3604	24

General properties

- reduction in performance for heavier-correlations
- buffer insensitivity
- more difficulty in estimation of traffic parameters like the mean, so a wider margin of safety must be used.