Fractals and Scaling in Finance: a comparison of two models

Blaž Krese¹, Sarah Marzen², Cornelia Metzig³, Zhi Qiao⁴, and Vipin P. Veetil⁵

¹Laboratory of Synergetics, Faculty of Mechanical Engineering, University of Ljubljana, Slovenia
²Department of Physics, University of California at Berkeley, USA
³Department of Infectious Disease Epidemiology, Imperial College London, UK
⁴NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore
⁵Department of Economics, George Mason University, USA

September 17, 2014

Abstract
Mandelbrot found that financial time series had fat tails and long dependence. Over the years a variety of empirical work has found these two properties to hold in many different financial time series. There are two alternative ways to model fat tails and long dependence. One, a stochastic process that is generated by the subordination of a levy process to a Fractal Brownian Motion. This was Mandelbrot’s approach. Two, a “fractal levy process” in which shocks from a levy-distribution are correlated. Model selection is a particularly difficult problem with financial time series that have fat tails and long dependence. In this paper we argue that the generalized Hurst exponent can be used as a simple measure for model selection. We provide mathematical preliminaries, review Mandelbrot’s analysis of financial time series and then compare the two stochastic processes that can be used to model fat tails and long dependence.

1 Introduction
Financial time series have long-range correlations, as do many other natural processes. However, the standard financial time series model of asset pricing (geometric Brownian motion) which yields the celebrated Black-Scholes equation does not exhibit such property. Modified stochastic models with Gaussian noise and non-stationary diffusion coefficients can be altered to be consistent with both behaviors on longer time scales.
In this paper, we study statistical signatures of alternative hypotheses and ideally, we would like to understand which of these stochastic models are best fits to empirical data from the financial market. However, the answer is a prerequisite to answering more fundamental social science questions—namely, which kinds of social interactions and entities are consistent with the best-fit stochastic model?

In theory, one could list all stochastic models, fit parameters of each of those models to the time series, and then do a maximum a posteriori (MAP) estimate to choose between these various models, though see Ref. [25] for a critique of this approach. There are several technical difficulties involved in a MAP selection procedure. First, correctly estimating parameters for these more complicated stochastic models is very difficult. Second, MAP estimates are usually much harder than the parameter estimation itself, since MAP estimates involve integrating the model likelihood across all possible parameters. Therefore, we would like to find a less computationally demanding way of selecting between these various stochastic models for best model of financial times series which will extremely crucial for both researchers and practitioners.

We take some inspiration from the large literature of using informational and complexity measures to distinguish between different types of randomness. The $\epsilon, \tau$-entropy rate is finite for chaotic time series and diverges for inherently stochastic time series [7]; the multi-scale entropy [3] can distinguish between white and pink noise; one can also use intensive complexity measures to distinguish between chaotic and inherently stochastic time series [21]. In other words, such coarse measures can be useful model selection tools.

We study a measure designed to quantify the long-rangedness of a process’ correlation structure—the Generalized Hurst exponent (GHE) $H(q)$ and its variation with $q$. This measure is a generalization of the more popular Hurst exponent as it captures the degree of long-rangedness in the $q^{th}$-moments process for any $q$, not just long-rangedness in the autocorrelation function or power spectrum. And this measure has been used previously by many econophysicists to quantify the presence or absence of long-range memory in financial markets, e.g. see Refs. [9, 8, 2, 24, 4]. However, none of them have ever calculated GHEs for fractional Levy processes or the subordination process proposed by Mandelbrot [14] and made comparison for their usage in real financial time series, in this paper we manage to fill in this gap by simulating both processes and presenting our results afterwards, which indicates that both processes are good approximation of market price change.

2 A Review of Mandelbrot’s Work on Financial Time Series

In a 1971 paper published in the Review of Economics and Statistics Benoit Mandelbrot wrote “It is widely agreed among economists that is does not suffice to affirm that equilibrium must be realized and to study its properties; one must also show how equilibrium is either achieved or approached” [13, p 227]. These words have a tragedy about them: the 1970’s marked the beginning of an era in which economists ‘widely agreed’ on precisely the opposite, interest in $\alpha$-stable distributions declined and “by the
mid-1970s, almost all references to the Mandelbrot program in the econometrics literature and in the neoclassical theory literature disappeared” [23, p 446]. During his brief stay in economics from 1960 to 1972 Mandelbrot unearthed time-invariant structures in data which equilibrium-rational expectations economics of the post-1970s would find difficult to explain. Perhaps the most significant aspect of Mandelbrot’s work is that his criticism of standard economic theory did not consist of an accumulation of anomalies (like ‘behavioral economies’) but in the discovery of rich structures far more general than the stylized facts which were the radar of standard theory. The ongoing European sovereign debt crisis and the US financial crisis of 2008 mean that we live today in one of those fat-tails which motivated so much of Mandelbrot’s foray in economics, his work is more relevant now than ever before.

Much of economics models randomness as “fluctuations around a normal state that represents equilibrium” [14, p 15]. A pictorial analogy of this is the edge of a razor blade, which when enlarged has “many irregularities” but from the users point of view is a straight line. The irregularities can therefore be thought of as deviations around a trend line. Contrast this with the coast line of United States, taking “into account an increasingly long portion will average out the small irregularities, but at the same time inject larger ones. A straight trend is never reached and interesting structures exist at every stage” [14, p 15]. Economic data is much more like the coast line of United States than the edge of a razor blade \(^1\).

Financial time series is riddled with such structures. Mandelbrot found two (perhaps) different but related structures: non-Gaussian tails and long-dependence. He christened these the Noah and Joseph effects reflecting “two stories in the Bible, those of the Flood and of the Seven Fat and Seven Lean Cows” [14, p 27] and proposed ways of measuring them. These structures in data may be of interest to market process scholars because equilibrium models fail to reproduce them. However model validation using data which exhibits Mandelbrotian patterns is challenging because much of commonly used methods of statistical inference assume the existence of moments. All is not dismal though, a whole host of results are available for parameter estimation, assessing goodness of fit, hypothesis testing et al [20].

2.1 Noah Effect and Levy-stable processes

Noah effect or non-Gaussian tails refers to the idea that ‘extreme’ price movements tend to be far more frequent than would be predicted by a Gaussian distribution. “For example there have been 35 falls greater than 6% in the daily Dow Jones Industrial Average since its inception in 1896, about 110 years ago. If the changes in the (logarithm) of the index are normally distributed one would expect that 35 falls of this magnitude would take place about once every 600 million years” [6, p 20]. Similarly using a Gaussian distribution “we expect a loss of greater than 4 standard deviations to occur once every

\(^1\)There is however an important difference between the coastline of United States and stock data. The coastline of United States is self-similar whereas price data are self-affine. “Self-similar constructions make free use of angles, and distances can be taken along arbitrary directions in the plane...The scale of each coordinate can be changed with no regard for the other”. Self-affine structures are derived using “linear operations that applies different reduction ratios along the time and price axes” [14, p 150]. And while self-similar fractals have unique fractal dimension, self-affine fractals have several [14, p 161].
126 years", however 11 such losses occurred between 22 October 1987 and 21 January 2008 on the FTSE-100 total return index\textsuperscript{2} [6, p 4]. All this excluding the crash of 2008!

One of the advantages of Gaussian distribution is that it is invariant under aggregation, which means that the distribution of sum of random variables is the same as the distribution of the random variables themselves up to scale and location. The idea of invariance under aggregation is important in economics because aggregate data is often a mixture distribution, i.e. elements of the aggregate data may - in fact are likely to - come from different distributions. Consider price data which “often refer to grades of a commodity that are not precisely known” [10, p 424] or income data which aggregates incomes from different types of activities like commerce, agriculture et al; in so far as each grade of commodity or each type of income follows a different distribution aggregate data will be mixture of different distributions. If the proposition of invariance under aggregation does not hold, then we may be explaining very little about disaggregate level data by explaining the distribution of aggregate level data.

However the Gaussian is only a special case of larger class of distributions known as $\alpha$-stable distributions which are invariant under aggregation\textsuperscript{3}. A random variable $X$ is said to have an $\alpha$-stable “distribution if there are parameters $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$, and $\mu$ real such that its characteristic function has the following form:

$$
E \exp i\theta X = \begin{cases} 
\exp \left( -\sigma |\theta|^\alpha \left( 1 - i\beta (\text{sign } \theta) \tan \frac{\pi \alpha}{2} \right) + i\mu \theta \right), & \text{if } \alpha \neq 1 \\
\exp \left( -\sigma |\theta| (1 - i\beta \frac{2}{\pi} (\text{sign } \theta) \ln |\theta|) + i\mu \theta \right), & \text{otherwise}
\end{cases}
$$

(1)

The parameter $\alpha$ is the index of stability and

$$
\text{sign } \theta = \begin{cases} 
1, & \text{if } \theta > 0 \\
0, & \text{if } \theta = 0 \\
-1, & \text{otherwise}
\end{cases}
$$

(2)

The parameters $\sigma, \beta, \mu$ are unique ($\beta$ is irrelevant when $\alpha = 2$)” [22, p 5]. $\alpha = 1$ gives the Cauchy distribution, $\alpha = 2$ Gaussian distribution and $\alpha = \frac{1}{2}$ Levy distribution. These are the only three cases in which we can write the density function in explicit form, for all other values of $\alpha$ we have to work with the characteristic function. Also, for $1 < \alpha < 2$ second and higher moments do not exist, and for $0 < \alpha \leq 1$ first and higher moments do not. The Gaussian has the special property of being the only $\alpha$-stable distribution for which first and second moments exist.

What does it mean for a moment not to exist? In case of the population distribution it simply means that the moment generating function does not exit, certain integrals of the density function does not exist. The idea is more intuitive when we speak of sample distribution. Imagine the following procedure. Randomly select two US firms, compute first and second moments. Now add another randomly selected US firm to the sample, and recompute the moment. We say that the sample moment does not exist if

\textsuperscript{2}Tests of fit overwhelmingly “reject the normal distribution” for six daily total return indices: ISEQ, CAC40, DAX30, FTSE100, Dow Jones Composite and S&P500 [6].

\textsuperscript{3}$\alpha$-stable distributions are “the only possible non-Gaussian limits of linearly weighted sums of random variables” [14, p 86].
as we add more firms the moments do not appear to converge to any particular values. The Gaussian therefore has the advantage of an analytical form and finite moments, however this comes at a cost. The existence of moments of Gaussian, especially second and higher moments, is a serious weakness when it comes understanding economic phenomenon like firm size distribution⁴, stock returns⁵ and income distribution for which second and higher sample moments are erratic⁶.

Needless to say we do not ‘know’ the distribution of the population from which data comes, this said we have three ways to model our ignorance when sample moments appear to be erratic. The first approach is to ignore the non-existence of sample moments and use Gaussian models, much of the work on asset pricing follows this approach. The second approach accepts that sample moments appear erratic but views it as a short-run phenomenon. The natural consequence of this belief is to model economic processes using a distribution for which population moments exist but may behave erratically for sample data, i.e. lognormal distribution⁷. And the third approach is to say that the erratic behavior of sample moments is a reflection of the fact that population moments do not exist and use the general class of α-stable distributions.

At a deeper level the Gaussian, lognormal and nonGaussian α-stable distributions reflect three kinds of randomness: mild, slow and wild. Mild randomness in firm size distribution would mean that if we were to randomly select some US firms and compute the contribution of the largest firm to the total employment or output - call this concentration - as we increase the sample size concentration declines. Wild randomness means concentration does not decline. Slow randomness resembles wild randomness in small samples (short-run) and mild randomness is large samples (long-run). Wild randomness is like the coastline of United States, interesting structures do not go away as we enlarge data size. Mild randomness is like the edge of a razor blade.

Mandelbrot proposes the use of nonGaussian α-stable stochastic processes to model economic data. The Gaussian is no good for reasons discussed earlier. And the lognormal is problematic for three reasons. One, economic structures like concentration are not short-run phenomena, they persist over long periods of time. Two, there is no reason in economics to assume that population moments are finite (this may come as a bit of a surprise to many economists!). “In physics, moments of lower order have a clear theoretical interpretation. For example, the population variance is often an energy that must be finite. In economics, to the contrary population variance is nothing but a tool for statistical analysis” [14, p 257]. And three “population moments of lognormal are not at all robust with respect to small deviations from absolutely precise lognormality” [14, p 258]. Sample estimates become meaningless if the data generating process is

⁴See [1]

⁵"If the price increase over a long period of time happens a posteriori to have been exceptionally large, one should expect, in a α-stable market, to find that most of this change occurred during only a few periods of especially high activity" [14, p 399]

⁶"My principle thesis is that to achieve a workable description of price change, of the distribution of income, firm sizes, etc., it is necessary to use random variables that have an infinite population variance” (Mandelbrot 1997, 79).

⁷"Macroscopic physics is concerned with assemblies of a colossal number of items and the long-run is meaningful and effective. But in finance or economics, such assemblies are not given time to develop, and (once again) references to them fail to convince” [14, p 49].
slightly different from the assumed lognormal\textsuperscript{8}.

In short, it would be wise to model financial time series as an $\alpha$-stable Levy motion rather than a Brownian motion. The difference between $\alpha$-stable Levy motion and Brownian motion is that in the former increments $P(t) - P(s)$ follow an $\alpha$-stable process with parameters $((t - s)^{\frac{1}{\alpha}}, \beta, 0)$ in the latter increments follow a Gaussian with parameters $(t - s, 0)$. Note that $(t - s)$ is the second moment of the Gaussian, but $(t - s)^{\frac{1}{\alpha}}$ is not the second moment of the $\alpha$-stable distribution, its merely a parameter. In fact for $\alpha < 2$ the second moment does not exist.

And, “statistically, the $\alpha$-stable distribution is a much better fit to the six total return equity indices”: ISEQ, CAC-40, DAX-30, FTSE-100, Dow Jones Composite and S&P-500 [6, p i]. However the “argument for an $\alpha$-stable distribution does not rest solely on the statistical fit of the distribution” [6, p 6]. There are alternate methods of modeling “fat tails”, however in so far as one is interested in developing a scientific understanding rather than merely curve fitting these are not as useful as the $\alpha$-stable distribution\textsuperscript{9}.

\textbf{2.2 Joseph Effect and Fractal-Brownian Motion}

Long-dependence or Joseph effects refers to the idea that “large changes tend to be followed by large changes of either sign and small changes tend to be followed by small changes” [11, p 418], in other words “risky times are not scattered randomly across quarterly or annual data. Instead, there is a degree of autocorrelation in the riskiness of financial returns” [5, p 158]. This is know as heteroskedasticity in the finance literature. GARCH/ARCH models are the usual approach to modeling second moments of financial time series. These models give an ARMA functional form to time-varying second moments. Mandelbrot however thought that modeling financial time series using GARCH-types models is a mistake. To get a sense of why, imagine making five plots of the S&P\textsuperscript{500} index, time on x-axis and the index value on y-axis, with each plot representing a day, month, year, decade and a century of data. Mandelbrot tells us that our naked eye will find it impossible to discern which plot represents a decade and which a day! He found this to be true of daily closing prices of cotton in New York 1900-1905, 1944-58 and 1880-1940 (pp 390-91), prices of wheat, railroad stocks and exchange rates [12]. Just like the scale-invariant structures in coastline of United States, there are time-invariant structures in financial data. “To represent such data by Markov, ARMA, ARCH, or any other short-term dependent process would be

\textsuperscript{8}The sensitivity of lognormal distribution is particularly problematic because when a distribution is characterized by few parameters many of its properties are entangled, therefore many properties may vanish together. The Gaussian does not have this problem because its moments are delocalized whereas the moments of lognormal are localized, in other words “different moments of the lognormal are determined by different portions of the density” (Mandelbrot 1997, 264). Ignoring this problem can lead to fallacious results if fitted distribution is used to make predictions.

\textsuperscript{9}An alternative method of modeling fat tails uses what is known as extreme value theory. Such procedures use the tails of the empirical distribution to make inferences about extreme values. This provides valuable results in many fields of application including insurance, hydrology, material and life sciences. Here we are more interested in the properties of the entire return series” [6, p 6]. “The t and Pareto distributions are examples of such fat tailed distributions. These do not have the scaling properties that we find desirable in return distributions”[6, p 22].
highly unadvisable\textsuperscript{10}, unless very strong independent reasons exist to believe that long duration phenomena are distinct and different from the short range phenomena to be modeled by ARMA or ARCH\textsuperscript{14}. In other words, to use three different models to represent the coastline of United States depending on whether its looked at from 100 meters above, 1000 meters above or 10,000 meters above is unadvisable unless there is reason to believe the forces that bring about small scale structures are different from those bring about similar structures at larger scale. In fact economic data points to precisely the opposite. Interestingly, “many of the progenitors of this [ARCH-type] technique had explicitly cited Mandelbrot’s hypothesis as one they wished to displace or render obsolete” \textsuperscript{11}. 

Mandelbrot proposed Fractal Brownian Motion (FBM) as a way of modeling long-dependence. FBM is characterized by two properties: $\mathbb{E}[P_{t+T} - P_t] = 0$ and $\mathbb{E}[P_{t+T} - P_t]^2 = T^{2H}$, $H$ is the Hurst-Holder exponent. FBM is a generalization of the Wiener Brownian Motion (WBM) in which $H = \frac{1}{2}$, and is founded on the Gaussian distribution like WBM. This appearing simple generalization yields long-run dependence between variables, the correlation between past values going back T periods and future values going forward T periods is $2^{2H-1} - 1$, note that the correlation is independent of T and does not go to zero as T increases \textsuperscript{14}. However the existence of long-run dependence does not mean that knowledge of such dependence is “valuable or capable of producing a profit” \textsuperscript{14}. Interestingly, though the increments of the motion are “very strongly dependent... they are uncorrelated, therefore they remain spectrally white” \textsuperscript{14}. 

2.3 Noah and Joseph Effects Together

$\alpha$-stable Levy process is a good model of nonGaussian tails and FBM of long-dependence. However Mandelbrot was interested in combining the two because real world price series show both long-dependence and nonGaussian tails. Mandelbrot and Taylor introduced the idea of subordination to account for both the Noah and Joseph effects \textsuperscript{14}. The idea was to decompose price movements into two parts, with the first part being price movements as function of an \textit{artificial} time, and the second part being artificial time as a function of real time. They called the artificial time \textit{fractal time}. More specifically, model fractal time as a $\alpha$-stable Levy process of real time, and model price as a FBM of fractal time. The economic reason behind this decomposition is that large price movements are associated with many more trades, in other words price movements for a given number of trades tend be to Gaussian but price movement for a given period of time tend to be nonGaussian. Mandelbrot found an elegant and economic meaningful decomposition of the Noah and Joseph effects. “This was the first time in history that someone had developed a taxonomy for all of the possible\textsuperscript{10}

\textsuperscript{11}Mandelbrot does not accept usual statistical tests of how ‘good’ a GARCH-type model is, in fact he sees those tests as ‘tests of the tests themselves’. He goes on to challenge “the proponents of the Brownian motion and ARCH-like processes to examine how the graphic outputs of their algorithms compare with actual data” \textsuperscript{14}. 

\textsuperscript{10}One of the problems with GARCH-type models is that the estimated parameters are “near-invariably mutually contradictory and have no intrinsic meaning” and “A satisfactory statistical fit is of no use in science unless the fitting parameters are consistent in time and have an intrinsic meaning ” \textsuperscript{14}. “Scientific modeling is not primarily a matter of curve fitting” \textsuperscript{14}.
cases which would cover the typical spectral shape of an economic time series” [18, p 295].

3 Mathematical Preliminaries

3.1 Generalized Hurst exponent

The generalized Hurst exponent and the related singularity spectra have been used previously to characterize multifractality in financial time series, e.g. as in Refs. [9, 8]. Suppose we have some time series in which data is taken at time intervals of length \( \tau \). Properly estimating these exponents from finite, noisy data is quite difficult, and it is easy to incorrectly conclude that certain time series have long-range memory, e.g. as studied in Refs. [2, 24, 4].

3.2 \( \alpha \)-stable random variables

Most of the models we study here use \( \alpha \)-stable random variables to parametrize possible noise terms. These \( \alpha \)-stable random variables can be characterized by four different, equivalent definitions, but we will focus on the two that matter most to us. These two definitions/theorems are presented without proof, because this is not a fucking math paper.

The first characterization is the generalized central limit theorem: a random variable \( X \) is \( \alpha \)-stable if and only if it has the same distribution as \( a_n(Y_1 + Y_2 + \ldots + Y_n) + b_n \) for asymptotically large \( n \) where \( Y_i \) are i.i.d. random variables, \( a_n > 0 \), and \( b_n \) is a real number. Interestingly, in order for the limit (and not just the asymptotic form) to be well-defined, we need choose \( a_n = \frac{1}{n^{\alpha/2}} \) for some \( 0 < \alpha \leq 2 \), and exactly which \( \alpha \) is appropriate depends on the tail behavior of \( Y_i \). When \( Y_i \) are drawn from a distribution with finite mean and variance, then we choose \( a_n = \frac{1}{n} \) and recover the usual central limit theorem—so Gaussian distributions are a special type of \( \alpha \)-stable random variable with \( \alpha = 2 \). Cauchy and Levy distributions are also special types of \( \alpha \)-stable distributions, with \( \alpha = 1 \) and \( \alpha = \infty \), respectively. One additional useful fact is that the sum of \( \alpha \)-stable random variables is an \( \alpha \)-stable random variable.

This first characterization motivates the use of \( \alpha \)-stable random variables as noise terms in stochastic models. Oftentimes, what we observe in time series data, financial or otherwise, is some very coarse-grained observable of a system which has many, many interacting elements. These interacting elements have a net effect on the observable which is often surprisingly well-described as the sum of many i.i.d. random variables, e.g. when simulating particle diffusion on a potential surface in a heat bath. The generalized central limit theorem above explains which noise distributions we’d expect if the collective dynamical effect of these many other interacting entities was some weak, disordered impulse. Of course, and somewhat unfortunately, these kinds of stochastic models will fail precisely when the system exhibits interesting collective/critical behavior—so perhaps one can think of these types of stochastic models as good null models. Any data which is badly fit by these kinds of stochastic models then shows evidence of interesting, large-scale collective effects.
The second characterization is more useful for simulating stochastic processes with \( \alpha \)-stable random variable noise terms and for fitting models to empirical distributions. Every \( \alpha \)-stable random variable \( X \) has the same distribution as

\[
X \equiv \begin{cases} 
\gamma Z + \delta & \alpha \neq 1 \\
\gamma Z + \left( \delta + \frac{2\beta}{\pi} \gamma \log \gamma \right) & \alpha = 1
\end{cases}
\]  

(3)

where \( Z \) is a random variable with characteristic function

\[
E[e^{itZ}] = \begin{cases} 
\exp(-|t|^\alpha \left(1 - i\beta \tan \frac{\pi \beta}{2} t\right)) & \alpha \neq 1 \\
\exp(-|t| \left(1 + \frac{2i\beta}{\pi} (t \log |t|)\right)) & \alpha = 1
\end{cases}
\]  

(4)

The parameters \( \gamma, \delta \) can be any real number, but \( |\beta| \leq 1 \) and \( 0 < \alpha \leq 2 \). (An \( \alpha \)-stable random variable with \( \alpha = 2 \) is Gaussian, and based on the characterization above, there would be no reason to look at \( \alpha > 2 \).) There are many other parameterizations of the characteristic function, but this way of writing an \( \alpha \)-stable random variable gives some intuitive meaning to \( \alpha, \beta, \delta, \gamma \). Looking at Eqn. 3, we intuit that \( \gamma \) is a scaling parameter (stretching or shrinking the \( x \)-axis, changing our units) and that \( \delta \) is a shift in the mean. From Eqn. 4, we see that if \( \beta = 0 \), then the characteristic function is symmetric; so \( \beta \) is a factor describing the asymmetry in the p.d.f. of \( X \) about its mean. An additional fact is that the p.d.f. of \( X \) falls off as \( x^{-\alpha+1} \) when \( \alpha < 2 \), which allows us to interpret \( \alpha \) as the parameter governing the long-rangedness of the distribution.

4 Some minimal generative models

We are interested not just in fitting time series data well, but in studying to minimal generative models which have some reasonable mechanistic interpretation. Practically speaking, restricting to such models is a form of model selection and complexity control. More optimistically, we might hope to gain some intuition for the mechanisms of the process from the values of the maximum likelihood parameters.

In each of these sections, we will not be clear about what the variable \( X_t \) corresponds to in real data. Typically, \( X_t \) is the logarithm of the asset price or differences thereof, as we will discuss later.

4.1 \( \alpha \)-stable Levy process

A Levy process is one in which increments \( X_{t+s} - X_t \) are stationary, i.i.d random variables. Modulo mathematical details, these processes are Markovian, and have a corresponding generative model given by

\[
dX_t = d\mathcal{N}(t)
\]  

(5)

where \( d\mathcal{N}(t) \) is some noise term. When \( d\mathcal{N}(t) \) are drawn from an \( \alpha \)-stable distribution with \( \sigma = t^{1/\alpha} \) and \( \delta = 0 \), then we’ve got an \( \alpha \)-stable Levy process. If \( \alpha = 2 \), this is a Wiener process. For other \( \alpha \), it’s a little more complicated. Simulations proceed as follows: choose a time discretization \( \Delta t \), draw an \( \alpha \)-stable random variable with any
\(\alpha, \beta\) and \(\gamma = 1, \ \delta = 0\), and multiply this noise impulse by \(\Delta t^{1/\alpha}\). Then increment the position by this noise term.

There are many, many other types of noise distributions that one can draw from, such as Gamma, Inverse Gaussian, Poisson, etc. In fact, there is plenty of empirical evidence in favor of using non-\(\alpha\)-stable random variables as noise terms. So why focus on \(\alpha\)-stable random variables?

One answer is that they correspond to a particularly interpretable class of minimal generative models, which could be viewed as some sort of zeroth-order approximation in a Taylor series towards ground truth. The ideas here are very much drawn from equilibrium statistical mechanics dynamics models, and are probably already somewhere in the economics literature, but there’s no terrible thing in repeating something obvious.

When we model weakly interacting particles diffusing on a potential surface, our model must account for the forcing of the external potential surface and the collisions with other particles. It is basically impossible to model all of the \(10^{23}\) particles’ motion, so that we very precisely understand the effect of collisions. So instead, as a zeroth-order approximation, we can make two immensely simplifying assumptions:

- The timescale on which collisions occur is very small compared to the timescale on which the particle responds to the external forcing
- Many, many particles collide with the particular particle we’re watching, and the activity of these colliding particles is uncorrelated

These two assumptions (or perhaps other assumptions) provide some justification for modeling the effect of these collisions with one “drag term” (moving the particle away from its external forcing-based trajectory) and a cancelling random noise term (redirecting the particle’s energy into some random direction). Since all of these particles are weakly interacting, as long as we’re not near some sort of critical point, we can probably safely assume that this noise term is the sum of a large number of i.i.d. random variables. By the generalized central limit theorem, this noise term then has to be drawn from an \(\alpha\)-stable distribution. When the impulses from the colliding particles have finite mean and variance, then the noise term is normally distributed (central limit theorem), and we recover the usual Langevin equation modulo the diffusion and drag coefficients. When the impulses from the colliding particles do not have finite variance and/or mean, then the noise term is drawn from a distribution with heavier tails than that of a Gaussian.

The analogy to the economy is that corporations are like colliding interacting particles whose motion is governed by some external potential, e.g. GDP. Trying to model the fluctuations in stock price of a particle company is similar, then, to modeling the fluctuations in position of a particle in a heat bath.

### 4.2 Subordination

Another possible way of generating long-range memory and heavy tails of a distribution by a joint process consists of subordination [16]. Its idea is that, If \(S(t)_{t \in \mathbb{R}^+_0}\) is a stochastic process, and \(T(t)_{t \in \mathbb{R}^+_0}\) is a non-negative stochastic process, a catenated process \(Z(t) = S(T(t))_{t \in \mathbb{R}^+_0}\) can be defined. \(T(t)\) is often called intrinsic time process.
and is admitted to have only non-negative stationary independent increments. Applied to financial data, it may be seen as a process modelling the market time, which may not be linear with physical time, since trades tend to occur very heterogeneously in time.

In order to model the two empirical facts (long range memory and fat tails of a distribution), the process is chosen such that \( T(t) \) be \( \alpha/2 \)-stable Lévy process (\( \alpha \in [0, 2] \))

\[
T(t) - T(s) \propto S_{\alpha/2}(c|t - s|^{\alpha/2}, 1, 0)
\]  

The increments are non-negative, since \( \beta = 1, \mu = 0 \). We use this process in section 5 for the simulation, where for \( S(t) \) fractional Brownian motion is also used.

Despite its technical success to reproduce the desired features, a limitation remains the interpretation: Since \( T(t) \) can only have non-negative increments, it can only be seen as market activity that does not happen over a constant time, (not as a time series that is amplified/damped by the exchange rate). As described by [?], \( T(t) \) describes the “speed” at which the price process evolves, depends on the quantity of new information arriving: more news usually implies more intense trading and faster price movements.”

An often overlooked aspect is that the moments of stable distributions of a subordinated process depend on the definition of price fluctuations: if price fluctuations are defined as the increment between two successive trades, which is common for high-frequency trading, the timescale at which they occur does not matter, i.e. \( S(t) \) and \( Z(t) \) will have the same fat tails. If however price fluctuations are defined as price changes per unit time interval (e.g. daily returns, as is the case for our data), then the fat tails of its stable distribution are sensitive to a rescales ‘market time’ \( T(t) \). For instance, after a busy day, price change represents an average over a large number of trades which possibly average out well, whereas on a day with few trades, much fewer trades are averaged, resulting possibly in larger fluctuations. As a consequence, with the latter definition for price changes, using subordination can distort distributions and even account for fat tails in stable distributions, even if they are not present in \( S(t) \) and \( T(t) \). The concept of explaining fat tails via averaging over a variable number is also a generating mechanism for fat tails in financial and other economic time series [15], [17].

### 4.3 Comparison

In this scheme features of different stochastic processes are summarized and compared.

<table>
<thead>
<tr>
<th></th>
<th>Brownian Motion</th>
<th>fBM</th>
<th>Lévy</th>
<th>fractional Lévy</th>
<th>Subordination</th>
</tr>
</thead>
<tbody>
<tr>
<td>self-similarity</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Hurst Exponent</td>
<td>0.5</td>
<td>0</td>
<td>( 0 &lt; H &lt; 1 )</td>
<td>1</td>
<td>( H(q) )</td>
</tr>
<tr>
<td>( \alpha )-stable</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>other features</td>
<td></td>
<td></td>
<td></td>
<td>increments correlated</td>
<td>no infinite variance of moments</td>
</tr>
</tbody>
</table>

Table 1: Summary of assumptions and results.
5 Model Selection using the Generalized Hurst exponent (GHE)

Currently the usual approach in the financial industry is to use either Brownian motion or fractional Brownian motion to simulate financial time series, but they cannot capture both fractals and fat tails of the financial time series. In this paper, we firstly plot the simulation result of time series samples of various stochastic processes in Fig. 1 including Brownian motion, fractional Brownian motion, Levy process, fractional Levy process and subordination. Then we plot the generalized Hurst Exponent of different moments for each stochastic process in Fig. 2. In the end, generalized Hurst exponents for different moments of 8 stocks’ time series are plotted in Fig. 3. By comparing the plots for the real financial data in Fig. 3 with the simulation results in Fig. 1, we can clearly see that the generalized Hurst exponent of financial time series is actually qualitatively similar to the Fractional Levy process or the subordination process which highly depends on the specific stock time series. Another important finding is that GHE is actually a much easier and efficient way to distinguish different financial time series models and from the plots of the GHE against moments $q$ we not only find that Fractional Levy process and the subordination are two much better choices for modelling financial time series but also can see the differences between various stocks.
Figure 2: The corresponding generalized Hurst exponents for $q \in [0, 4]$ for the above stochastic processes. For the fractional Brownian motion, fractional Levy process and subordination the Hurst exponent is $H = 0.7$. For the Levy process $\alpha = 1.2$ and $\beta = 0$.

6 Discussion and Conclusion

In this paper, we summarize two key features of the financial time series highlighted by Mandelbrot [14] and try to combine these two features analytically and computationally through two methods: Fractional Levy process and subordination process where a levy process is subordinated to a Fractional Brownian Motion. By comparing both simulations and results from real data, we find that both methods are useful and much better than what have been commonly used in the financial industry. We also point out a good practical usage of the generalized Hurst Exponent (GHE), i.e. an easy and efficient tool for model selection. Further studies can be carried out on the usage of GHE to categorize different industries or markets. What’s more, the dynamics of the two methods in the paper still needs to be studied from agent behavior point of view so that our result will have a much solid foundation.

However, there are still some limitations for our usage of GHE. After the conceptual and numerical analysis we reflect on the question how much information is contained in the generalized Hurst exponent, in particular on its relation to the two processes we mentioned above, fractional Levy and subordination. On one hand, given $H(q)$ of a certain process across the moments, i.e. the scaling behaviour of the moments, is not sufficient to know its kind of distributions, since the moments might take different values for the same $H(q)$. On the other, two identical distributions may be generated
Figure 3: Generalized Hurst exponent for different moments of various financial time series. We use eight representative stocks from Dow Jones Industrial Average (DJIA) ranging from April 1990 to June 2007.

by processes that exhibit different $H(q)$, and from one point of time the time series cannot be inferred. Therefore, $H(q)$ represents additional information about a time series which empirically is an excellent method for model selection, but cannot be concluded as an alternative to have its all parameters of the distributions.

References


