SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853-3801

TECHNICAL REPORT NO. 959

February 1991

LINEAR MODELS WITH LONG-RANGE
DEPENDENCE AND WITH
FINITE OR INFINITE VARIANCE¹

by

G. Samorodnitsky and M.S. Taqqu²

¹Supported in part by the ONR Grant N00014-90-J-1287 at Boston University.

²Department of Mathematics, Boston University, Boston, MA 02215.

Linear models with long-range dependence and with finite or infinite variance ^{*†‡}

Gennady Samorodnitsky
School of OR/IE
Cornell University
Ithaca, NY 14853

Murad S. Taqqu
Department of Mathematics
Boston University
Boston, MA 02215

February 15, 1991

Abstract

We analyze a number of linear models that display long-range dependence and either finite or infinite variance.

1 Introduction

This paper discusses a number of linear models that display long-range dependence and either finite or infinite variance. Our goal is to show that these models, while different, share many common features. The models include the finite variance fractional ARIMA and fractional Gaussian noise, as well as processes with long-range dependence and infinite variance such as the linear fractional stable noise, the harmonizable stable noise, and the sub-Gaussian and sub-stable processes. All these processes are linear transformations of white noise. For a brief discussion of some non-linear processes, see Taqqu [Taq81], [Taq87] and [Taq88]. For a bibliographical guide on long-range dependence, see Taqqu [Taq86].

Finite variance time series exhibit long-range dependence if, as $t \rightarrow \infty$, their (auto)covariances $\text{Cov}(X_0, X_t)$ decrease to zero like a power function, slowly enough so that $\sum_{t=0}^{\infty} \text{Cov}(X_0, X_t) = \infty$. These conditions on the covariance function can be

^{*}Supported in part by the ONR Grant N00014-90-J-1287 at Boston University.

[†]*AMS 1980 subject classifications.* Primary 62M10; secondary 60G10, 60G15, 60E07

[‡]*Keyword and phrases:* Stable processes, stable distributions, self-similar processes, stable Lévy motion, ARIMA, fractional ARIMA, fractional Brownian motion, fractional Gaussian noise, linear fractional stable noise, harmonizable, sub-Gaussian, sub-stable.

expressed in terms of the spectral density as follows: at low frequencies, the spectral density behaves like a power function and it diverges as the frequencies tend to zero, a phenomenon sometimes referred to as “1/f noise.”

Mandelbrot and Van Ness [MV68] refer to long-range dependence as the *Joseph effect* because the biblical character Joseph was faced with long periods of plenty followed by long periods of famine. To model it, Mandelbrot introduced *fractional Gaussian noise* (FGN). FGN, however, is much more than an ad hoc construction devised to exhibit the Joseph effect. It is the 1-step increment of the Gaussian self-similar process fractional Brownian motion (FBM), and as such, it turns up naturally in the context of central limit theorems. Its special properties become apparent when it is regarded as a fixed point of the renormalization group connected with the central limit theorem (Taqqu [Taq81], Sinai [Sin76]).

Because FGN is a specific process, its covariances are determined at all lags. This lack of flexibility can be a problem if one wants to use FGN to model a real-life phenomenon. This problem can be easily remedied, however. Since long-range dependence is only a feature of the long-run behavior of the covariances, one can always modify the covariances of FGN at small lags in order to fit a desired short-run dependence structure.

Fractional ARIMA offers a convenient way to achieve such a modification. The fractional ARIMA described in the following section was introduced by Granger and Joyeux [GJ80] and developed by Hosking [Hos81]. Our exposition is based on Brockwell and Davis [BD87] and Samarov and Taqqu [ST88].

Although fractional ARIMA is typically obtained from the usual ARIMA (p, d, q) by taking d fractional, the basic construction of FGN and fractional ARIMA is the same: start with white noise (or stationary noise with weak dependence) and “convolve” it with a power function. Integral representations best capture the centrality of such a construction. Integral representations, moreover, can also be used to generate models with infinite variance if the underlying noise is allowed to have infinite variance.

In contrast to Gaussian noises, noises with infinite variance are more likely to take large values. Mandelbrot and Van Ness [MV68] refer to this as the *Noah effect*, an allusion to the biblical character Noah who lived through an exceptionally heavy flood.

This paper will use symmetric stable distributions to model noises with infinite variance. The stable distributions are parametrized by $\alpha \in (0, 2]$ and are Gaussian when $\alpha = 2$. There is an important distinction between the Gaussian distribution and stable distributions with $0 < \alpha < 2$. Whereas the tail of a Gaussian distribution is exponential, the tail of a stable distribution with $0 < \alpha < 2$ obeys a power law, and decreases to zero so slowly that the variance is infinite. Note that both long-range dependence (Joseph effect) and high variability (Noah effect) involve power laws, but in different contexts. In the case of long-range dependence, the covariance function behaves like a power law at large lags. In the case of high variability, it is

the distribution tail which behaves like a power law.

There are several other advantages to choosing stable distributions with $0 < \alpha < 2$. First, a symmetric stable random vector is stable if and only if any linear combination of its components is stable. (See Samorodnitsky and Taqqu [ST91b] for a discussion of this property.) Secondly, stable distributions appear as limits in central limit theorems. Finally, these distributions are quite well understood. For properties of marginal stable distributions, see Feller [Fel71] and Zolotarev [Zol86]. Properties of finite-dimensional stable distributions are scattered throughout the literature and will be presented in detail in a monograph on stable random processes (Samorodnitsky and Taqqu [ST93]).

Stable distributions with $0 < \alpha < 2$, like the Gaussian distribution, serve as archetypes. In practice, one can use any distribution which is in the domain of (normal) attraction of the stable, i.e. the Pareto distribution. The only requirement is that the tail obey a power law. By restricting only the tail distribution we obtain modeling flexibility.

In order to construct models that exhibit both the Joseph and the Noah effects, we shall replace the Gaussian noises in the integral representations by stable noises. Various Gaussian models have multiple extensions to the stable case $0 < \alpha < 2$. For example, the linear fractional stable noise and the harmonizable fractional stable noise are two different processes that display the Joseph and the Noah effects. They both reduce to FGN when α is set equal to 2.

Because the variance is infinite, covariances do not exist in the stable case $0 < \alpha < 2$. One can focus on a function $r(t)$, however, defined as the difference between the joint characteristic function of (X_0, X_t) and the product of the marginal characteristic functions. Obviously, $r(t) \equiv 0$ if the time series X_t , $t = 0, 1, 2, \dots$ is i.i.d.. Because $r(t)$ is asymptotic to the covariance in the Gaussian case (provided the covariances tend to zero at large lags), one can try to use $r(t)$ as a measure of *asymptotic* dependence, even in the stable case $0 < \alpha < 2$. For example, Astrauskas [Ast83], Astrauskas, Levy and Taqqu [ALT90] and Levy and Taqqu [LT91] show that $r(t)$ for the linear fractional stable noise decreases like a power law at large lags. There is, however, still much to be learned about the significance of $r(t)$ in the stable case $0 < \alpha < 2$.

In the following section we focus on fractional ARIMA. Then, in Section 3, we describe how to generate long-range dependence by using self-similar processes and define fractional Gaussian noise (FGN). Section 4 discusses the various integral representations for FGN. In Section 5, we concentrate on processes that display both the Joseph and the Noah effects and present several extensions of FGN to the stable case $0 < \alpha < 2$. We conclude by describing the behavior of $r(t)$ in Section 6.

A word about notation. A continuous time process is denoted $X(t)$, $-\infty < t < \infty$, while a discrete time process (time series, sequence) is denoted X_t , $t = \dots, -1, 0, 1, \dots$

2 Generating long-range dependence via fractional ARIMA

The ARIMA (p, d, q) model for a time series X_t , $t = \dots, -1, 0, 1, 1, \dots$, is symbolically denoted

$$\Phi_p(B)\Delta^d X_t = \Theta_q(B)\epsilon_t, \quad t = \dots, -1, 0, 1, \dots, \quad (2.1)$$

where p, d, q are non-negative integers, B is the backward operator (i.e. $BX_t = X_{t-1}$), $\Delta = I - B$ is the difference operator (i.e. $\Delta X_t = X_t - X_{t-1}$), Φ_p and Θ_q are polynomials of order p and q respectively, and $\{\epsilon_t, t = \dots, -1, 0, 1, \dots\}$ is a mean zero i.i.d. sequence assumed to be Gaussian or at least to have a finite variance σ^2 . Relation (2.1) forms a system of equations for $\{X_t\}$. When there is no differencing ($d = 0$), the ARIMA $(p, 0, q)$ is called ARMA (p, q) .

Assume, as is often done in practice, that the polynomials $\Phi_p(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ and $\Theta_q(z) = 1 + \theta_1 z + \dots + \theta_q z^q$, viewed a function of a complex number z have no zero in the unit circle $|z| \leq 1$. This ensures that when $d = 0$, the time series $\{X_t\}$ is stationary, causal (depends only on the past ϵ 's) and invertible (the ϵ_t 's can be expressed in terms of the X_t 's). Details can be found in any textbook on time series, e.g. Brockwell and Davis [BD87].

When $d = 0$, the covariances $R_t = EX_t X_0$, are a mixture of damped exponentials. This behavior is often referred to as “weak dependence.” Note however that by choosing the coefficients of Φ_p and Θ_q suitably, it is possible to ensure that the exponentials decrease very slowly. For example, the covariance of the AR(1)=ARMA(1,0) process $(1 - aB)X_t = \epsilon_t$ with $|a| < 1$ is $R(t) = \frac{\sigma^2}{1-a^2} a^t$, $t = 0, 1, 2, \dots$, so, if $a = 0.9$ for example, the rate of decrease of $R(t)$ as t increases is very slow. It is nevertheless exponential and hence falls in the category of “weak dependence.” In general, one has weak dependence if $\sum_{t=0}^{\infty} |R(t)| < \infty$.

Weak dependence can also be characterized by using the spectral density $f(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} e^{-it\lambda} R(t)$. The spectral density of a stationary ARMA (p, q) is

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\Theta_q(e^{-i\lambda})|^2}{|\Phi_p(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi.$$

It is the ratio of two polynomials in $\cos \lambda$ and $\sin \lambda$ and, in our setup, it is bounded at all frequencies.

We now turn to the fractional ARIMA time series. It is based on the ARIMA (p, d, q) model, but the parameter d , instead of being a non-negative integer is allowed to take fractional values, either positive or negative. To simplify the discussion, set $p = q = 0$. If d is a non-negative integer then $\Delta^d X_t = \epsilon_t$ describes a model where X_t , differenced d times, yields a sequence of i.i.d. random variables (the random walk, for example, is ARIMA $(0, 1, 0)$).

Now define $\Delta^d X_t = \epsilon_t$ for d fractional as $X_t = \Delta^{-d} \epsilon_t$ and interpret $\Delta^{-d} = (I - B)^{-d}$ by using the formal power series expansion $(1 - z)^{-d} = \sum_{j=0}^{\infty} b_j(-d) z^j$ as follows:

$$\Delta^{-d} = (I - B)^{-d} = \sum_{j=0}^{\infty} b_j(-d) B^j,$$

where B^j denotes the backward operator B iterated j times, $b_0(d) = 1$,

$$b_j(-d) = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} = \prod_{k=1}^j \frac{k-1+d}{k}, \quad j = 1, 2, \dots \quad (2.2)$$

and Γ is the Gamma function ($\Gamma(x+1) = x\Gamma(x)$) with the convention $1/\Gamma(x) = 0$ when x is a non-positive integer.

The formal definition of fractional ARIMA $(0, d, 0)$ is then as follows: it is the moving average

$$X_t = \sum_{j=0}^{\infty} b_j(-d) \epsilon_{t-j}, \quad t = \dots, -1, 0, 1, \dots, \quad (2.3)$$

where the b_j 's are given in (2.2). Note that when $d = 0, -1, -2, \dots$, Δ^{-d} is merely the difference operator Δ iterated $|d|$ times, and in that case, $b_j(-d) = 0$ when $j > |d|$. For example $X_t = \epsilon_t$ when $d = 0$, $X_t = \epsilon_t - \epsilon_{t-1}$ when $d = -1$ and $X_t = \epsilon_t - 2\epsilon_{t-1} + \epsilon_{t-2}$ when $d = -2$.

For (2.3) to be well-defined, it is necessary that $\sum_{j=0}^{\infty} b_j^2(-d) < \infty$, that is

$$d < \frac{1}{2},$$

since by using Sterling's formula, $b_j(-d) \sim \frac{1}{\Gamma(d)} j^{d-1}$ as $j \rightarrow \infty$ and $d \neq 0, -1, -2, \dots$. (When $d = 0, -1, -2, \dots$, $b_j(-d) = 0$ for $j \geq |d|$.) Note that $\sum_{j=0}^{\infty} |b_j(-d)| = \infty$ when $d > 0$, so the case $0 < d < \frac{1}{2}$ falls outside the classical theory of Box and Jenkins.

Since $\Delta^{-d} = \Delta^n \Delta^{-d'}$ with $n \geq 0$ and $d' \in [\frac{1}{2}, \frac{1}{2})$, the operator Δ^{-d} with $d < \frac{1}{2}$ can be interpreted as $\Delta^{-d'}$ followed by a number of full differencing.

In sharp contrast to the usual ARMA (p, q) , the covariances $R(k) = EX_0 X_k$ of the fractional ARIMA satisfy

$$R(k) \sim C_d k^{2d-1} \text{ as } k \rightarrow \infty \quad (2.4)$$

where $C_d = \frac{1}{\pi} \Gamma(1-2d) \sin \pi d$. Actually, $R(k) = 0$ for large enough k when $C_d = 0$. Note that C_d is 0 for $d = 0, -1, -2, \dots$, is positive for $d \in (0, \frac{1}{2})$ or $d \in (-(j+1), -j)$, $j = 1, 3, \dots$, and is negative otherwise.

The spectral density is particularly illuminating. In view of (2.3), it equals

$$f(\lambda) = \frac{\sigma^2}{2\pi} |1 - e^{-\lambda}|^{-2d} = \frac{\sigma^2}{2\pi} (2 \sin \frac{\lambda}{2})^{-2d}$$

and satisfies

$$f(\lambda) \sim \frac{\sigma^2}{2\pi} |\lambda|^{-2d} \text{ as } \lambda \rightarrow 0. \quad (2.5)$$

Since $\{X_t\}$ is well defined for $d < \frac{1}{2}$, we have $R(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda < \infty$ when $d < \frac{1}{2}$, a fact which can be directly verified by using (2.5).

The “autoregressive representation” $\Delta^d X_t = \epsilon_t$, i.e.

$$\sum_{j=0}^{\infty} b_j(d) X(t-j) = \epsilon_t \quad (2.6)$$

is always defined for $-\frac{1}{2} < d < \frac{1}{2}$, because the $b_j(d)$ ’s exactly compensate for the dependence structure of the time series $\{X_t\}$. Since $b_j(d) \sim \frac{1}{\Gamma(-d)} j^{-d-1}$ as $j \rightarrow \infty$, we have $\sum_{j=0}^{\infty} |b_j(d)| = \infty$ for $d < 0$, so this case falls again outside the classical theory of Box and Jenkins. It is best to view (2.6) as the effect of a linear filter which unravels the dependence structure of $\{X_t\}$.

We have been supposing $p = q = 0$. It is no problem now to define a general fractional ARIMA (p, d, q) with $d < \frac{1}{2}$ as

$$\Phi_p(B) X_t = \Theta_q(B) \Delta^{-d} \epsilon_t,$$

or

$$X_t = \Phi_p(B)^{-1} \Theta_q(B) \Delta^{-d} \epsilon_t$$

under our assumptions on Φ_p . The effect of the operator $\Phi_p^{-1}(B) \Theta_q(B)$ is to provide modelling flexibility. It modifies the covariances and spectral density without modifying their asymptotic behavior (as $k \rightarrow \infty$ for $R(k)$; as $\lambda \rightarrow 0$ for $f(\lambda)$). Thus

$$R(k) \sim C k^{2d-1} \quad \text{as } k \rightarrow \infty, \quad (2.7)$$

where C is a constant, and

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \frac{\Theta_q(e^{-i\lambda})}{\Phi_p(e^{-i\lambda})} \right|^2 |1 - e^{-i\lambda}|^{-2d} \sim \frac{\sigma^2}{2\pi} \left(\frac{\Theta_q(1)}{\Phi_p(1)} \right)^2 |\lambda|^{-2d} \text{ as } \lambda \rightarrow 0.$$

Remark. Suppose that the time series is $Y_t = \mu + X_t$ where X_t is fractional ARIMA with $d < \frac{1}{2}$ and μ is the unknown mean. How to estimate μ ? Samarov and Taqqu [ST88] compare the ordinary least square (OLS) estimator, $\bar{Y}_n = \frac{1}{n} \sum_{t=1}^n Y_t$, with the best linear unbiased estimator (BLUE). They show that when $-\frac{1}{2} < d < \frac{1}{2}$ the OLS estimator is almost as efficient as BLUE. The paper includes a discussion that provides some insight on the structure of fractional ARIMA for both $-\frac{1}{2} < d < \frac{1}{2}$ and $d \leq -\frac{1}{2}$. Yajima [Yaj88], [Yaj89] has extended the results of Samarov and Taqqu [ST88] to more complex regression (See also Beran [Ber90]).

3 Generating long-range dependence via self-similar processes

A continuous time process $Z(t)$, $-\infty < t < \infty$ is *self-similar* with exponent $H > 0$ if for all $a > 0$, $\{Z(at), -\infty < t < \infty\}$ and $\{a^H Z(t), -\infty < t < \infty\}$ have the same finite-dimensional distributions, that is, for all $a > 0$, $n = 1, 2, \dots$, and real $t_1, \dots, t_n$,

$$P\{Z(at_1) \leq z_1, \dots, Z(at_n) \leq z_n\} = P\{a^H Z(t_1) \leq z_1, \dots, a^H Z(t_n) \leq z_n\}.$$

We say that $\{Z(t), -\infty < t < \infty\}$ is *H-sssi* if it is self-similar with exponent $H > 0$ and if, in addition, it possesses stationary increments, that is, if the finite-dimensional distributions of the processes $\{Z(t+s) - Z(s), -\infty < t < \infty\}$ do not depend on s . Brownian motion $B(t)$, for example, is $\frac{1}{2}$ -sssi. Assuming stationary increments is convenient because by differencing $\{Z(t)\}$, one obtains stationary processes.

We want to describe properties of $\{Z(t)\}$ and of its increments. We will not suppose in this section that $\{Z(t)\}$ is Gaussian, only that it has a finite variance.

Self-similarity and the stationarity of the increments yield $Z(0) = 0$, $EZ^2(-1) = EZ^2(1)$ and $EZ(t) = 0$ for all t . Moreover, by self-similarity,

$$EZ^2(t) = |t|^{2H} EZ^2(1)$$

for any fixed t , and consequently

$$\begin{aligned} EZ(t_1)Z(t_2) &= \frac{1}{2}\{EZ^2(t_1) + EZ^2(t_2) - E(Z(t_1) - Z(t_2))^2\} \\ &= \frac{1}{2}\{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\}EZ^2(1). \end{aligned} \quad (3.1)$$

The stationarity of the increments entails $H \leq 1$, because, by the triangle inequality,

$$\begin{aligned} (EZ^2(t+t))^{1/2} &= (E(Z(t+t) - Z(t) + Z(t))^2)^{1/2} \\ &\leq (E(Z(t+t) - Z(t))^2)^{1/2} + (E(Z(t))^2)^{1/2} = 2(EZ^2(t))^{1/2}. \end{aligned}$$

Divided by $|t|^H$, this becomes $2^H \leq 2$ which can only hold when $H \leq 1$. The case $H = 1$ corresponds to the degenerate process $Z(t) = tZ(1)$. We exclude this case by requiring

$$0 < H < 1.$$

We will be interested in the 1-step increments

$$X_t = Z(t+1) - Z(t), \quad t = \dots, -1, 0, 1, \dots$$

of the process Z . They form a stationary sequence with covariances

$$R(k) = EX_t X_{t+k} = \frac{EX_1^2}{2} \{(k+1)^{2H} - 2k^{2H} + |k-1|^{2H}\} \quad (3.2)$$

for $k \geq 0$ and hence satisfy

$$R(k) \sim (EX_1^2)H(2H-1)k^{2H-2} \text{ as } k \rightarrow \infty. \quad (3.3)$$

Notice that this asymptotic behavior is similar to the asymptotic behavior of the covariances of fractional ARIMA (p, d, q) given in (2.7). The relation between d and H is:

$$d = H - \frac{1}{2}. \quad (3.4)$$

Observe that $0 < H < 1$ corresponds to $-\frac{1}{2} < d < \frac{1}{2}$. Thus we have long-range dependence when $\frac{1}{2} < H < 1$ ($\sum_{k=0}^{\infty} R(k) = \infty$). When $H = \frac{1}{2}$, the X_t are uncorrelated (white noise). When $0 < H < \frac{1}{2}$, we have weak and negative dependence because $R(k) < 0$ for $k \neq 0$. The spectral density of $\{X_k\}$ is

$$f(\lambda) = C|e^{i\lambda} - 1|^2 \sum_{j=-\infty}^{+\infty} \frac{1}{|\lambda + 2\pi j|^{2H+1}}, \quad -\pi \leq \lambda \leq \pi, \quad (3.5)$$

where C is some positive constant. The behavior of $f(\lambda)$ at low frequencies is

$$f(\lambda) \sim C|\lambda|^{1-2H} \text{ as } \lambda \rightarrow 0. \quad (3.6)$$

It is similar to that of fractional ARIMA with $d = H - 1/2$.

What are the advantages of using the increments of a self-similar process, as compared to fractional ARIMA? Because they scale, self-similar processes have particularly nice properties. Their increments are fixed points of renormalization group transformations (see Taqqu [Taq81]). Self-similar processes also appear as limits in central limit theorems. For example, if U_t is fractional ARIMA (p, d, q) with $d \in (-\frac{1}{2}, \frac{1}{2})$, then the finite-dimensional distributions of $\frac{1}{N^{2d+1}} \sum_{j=1}^{[Nt]} U_j$ converges as $N \rightarrow \infty$ to those of fractional Brownian motion, a Gaussian H-sssi process with $H = d + \frac{1}{2}$. This is why self-similar processes and their increments serve as prototypical models.

Because they have a specific covariance, namely (3.2), the increments of a H -sssi process form rigid models. Fractional ARIMA, with $p = q = 0$, shares the same problem. When modeling a phenomena with long-range dependence, one can use in fact any covariance structure, as long as (3.2) or (3.6) are satisfied.

Note that there are many finite variance *non-Gaussian* H-sssi processes with $0 < H < 1$. (See for example Taqqu [Taq87], [Taq88].) While they all possess the same covariance structure (3.2), their finite dimensional distributions is different. Since we focus here only on second order properties, we may as well assume that the process is Gaussian.

The Gaussian H-sssi process $\{Z(t)\}$ is called *fractional Brownian motion* (FBM) and its increment $\{X_t\}$ is called *fractional Gaussian noise* (FGN). Formally, FBM is defined as the Gaussian process with mean 0 and covariance (3.1). Its 1-step increment $\{X_t\}$ is defined as the Gaussian sequence $X_t = Z(t+1) - Z(t)$, $t = \dots, -1, 0, 1, \dots$, with mean 0 and covariance (3.2).

Both FBM and FGN possess an integral representation. Integral representations are useful because they demonstrate the structure of the process and because they also provide a means of combining long-range dependence and infinite variance. The integral representation for FGN is introduced in the next section.

4 Integral representation for fractional Gaussian noise

Fractional Gaussian noise, the 1-step increment of fractional Brownian motion can be represented, up to a multiplicative constant, as

$$X_t = \int_{-\infty}^{+\infty} g_t(x) M_2(dx), \quad t = \dots, -1, 0, 1, \dots \quad (4.1)$$

where

$$g_t(x) = a[(t+1-x)_+^{H-\frac{1}{2}} - (t-x)_+^{H-\frac{1}{2}}] + b[(t+1-x)_-^{H-\frac{1}{2}} - (t-x)_-^{H-\frac{1}{2}}] \quad (4.2)$$

and a and b are real numbers. $(\)_+$ and $(\)_-$ denotes respectively the positive and negative part. $M_2(\)$ is a white noise Gaussian measure. This means that for any fixed Borel subset Δ of the real line with finite Lebesgue measure, the random variable $M_2(\Delta)$ is Gaussian with mean 0 and variance $|\Delta|$. (Here $|\Delta|$ denotes the Lebesgue measure or “length” of the set Δ .) The measure M_2 , moreover, is independently scattered, that is, if $\Delta_1, \dots, \Delta_n$ are disjoint Borel subsets of the line with finite Lebesgue measure, then the random variables $M_2(\Delta_1), \dots, M_2(\Delta_n)$ are independent. Hence $EM_2(\Delta_1)M_2(\Delta_2) = |\Delta_1 \cap \Delta_2|$. The integral representation (4.1) can be viewed heuristically as a “sum” of white noises $M_2(dx)$ weighted by $g_t(x)$.

By taking $\Delta = [0, t]$ one can regard $M_2([0, t])$ as Brownian motion $B(t)$, the Gaussian process with stationary independent increments ($B(t) = -M_2([-t, 0])$ if $t < 0$). Thus $M_2(dx)$ can also be thought informally as $dB(x)$, the increment of Brownian motion at x .

When $H = \frac{1}{2}$, interpret (4.2) as $(a-b)1_{(t, t+1)}(x)$, so (4.1) becomes in this case $X_t = (a-b)M_2([t, t+1]) = (a-b)(B(t+1) - B(t))$, that is, a sequence of i.i.d. mean zero Gaussian random variables.

For the integral (4.1) to be defined, one must have

$$\int_{-\infty}^{+\infty} g_t^2(x) dx < \infty, \quad t = \dots, -1, 0, 1, \dots \quad (4.3)$$

The assumption $0 < H < 1$ ensures that (4.3) is satisfied. One then has $EX_t = 0$ and

$$EX_t X_{t+k} = \int_{-\infty}^{+\infty} g_t(x) g_{t+k}(x) dx = C(a, b) \frac{1}{2} \{ (k+1)^{2H} - 2k^{2H} + |k-1|^{2H} \} \quad (4.4)$$

for $k \geq 0$. Comparing this to (3.2), we conclude that $\{X_t\}$ is (up to scale constant) the fractional Gaussian noise defined in the preceding section. We also see that the values of a and b play no important role. All they do is change the scale constant $C(a, b)$. It is traditional to set $b = 0$ in (4.1) so as to get a causal representation. Taking $a = b = 1$ yields the well-balanced representation

$$X_t = \int_{-\infty}^{+\infty} (|t+1-x|^{H-1/2} - |t-x|^{H-1/2}) M_2(dx).$$

The representation (4.1) with kernel g_t given in (4.2) is a *moving-average representation*. We shall now consider the *harmonizable* (or spectral) representation obtained heuristically by taking the Fourier transforms $\tilde{g}_t$ and $\tilde{M}_2$ of g_t and M_2 and replacing $\int_{-\infty}^{+\infty} g_t(x) M_2(dx)$ by $\int_{-\infty}^{+\infty} \tilde{g}_t(\lambda) \tilde{M}_2(d\lambda)$ (think of Parseval's formula). These heuristic ideas can be made precise (Taqqu [Taq79]). Formally, let

$$X_t = C \int_{-\infty}^{+\infty} e^{i\lambda t} \frac{e^{i\lambda} - 1}{i\lambda} \frac{1}{|\lambda|^{H-1/2}} \tilde{M}_2(d\lambda) \quad (4.5)$$

C is a constant and $\tilde{M}_2(d\lambda)$ is, this time, a complex-valued Gaussian white noise measure. This means that for any Borel sets Δ_1 and Δ_2 with finite Lebesgue measure, $E \tilde{M}_2(\Delta_1) \tilde{M}_2(\Delta_2) = |\Delta_1 \cap \Delta_2| \tilde{M}_2(\Delta) = \tilde{M}_2(-\Delta)$, $\text{Re } \tilde{M}_2(\Delta)$ and $\text{Im } \tilde{M}_2(\Delta)$ are i.i.d. $N(0, |\Delta|/2)$.¹ The integrand (call it $\tilde{g}_t(\lambda)$) must satisfy $\int_{-\infty}^{+\infty} |\tilde{g}_t(\lambda)|^2 d\lambda < \infty$ for the integral (4.5) to be defined. Again, this is the case here, because $0 < H < 1$. X_t in (4.5) is also fractional Gaussian noise because its covariance equals (4.4) up to a multiplicative constant. Since its covariance can also be written

$$\begin{aligned} R(k) &= EX_{t+k}X_t = C^2 \int_{-\infty}^{+\infty} e^{i\lambda k} \left| \frac{e^{i\lambda} - 1}{i\lambda} \right|^2 \frac{1}{|\lambda|^{2H-1}} d\lambda \\ &= C^2 \sum_{j=-\infty}^{+\infty} \int_{-\pi+2\pi j}^{\pi+2\pi j} e^{i\lambda k} |e^{i\lambda} - 1|^2 \frac{1}{|\lambda|^{2H+1}} d\lambda \\ &= \int_{-\pi}^{\pi} e^{i\lambda k} \left(C^2 |e^{i\lambda} - 1|^2 \sum_{j=-\infty}^{+\infty} \frac{1}{|\lambda + 2\pi j|^{2H+1}} \right) d\lambda, \end{aligned}$$

we see that $f(\lambda)$ given in (3.5) is, indeed, up to a multiplicative constant, the spectral density of fractional Gaussian noise.

¹This definition ensures that the imaginary parts in the right-hand side of (4.5) cancel, so that X_t is real. One can proceed alternatively as follows: Take $\tilde{M}_2 = W_1 + iW_2$, where W_1 and W_2 are independent real white noise Gaussian measures, such that for each Δ , $W_i(\Delta) \sim N(0, |\Delta|/2)$, $i = 1, 2$. Then define X_t as the real part of the right-hand side of (4.5).

5 Infinite variance processes

Both fractional ARIMA and fractional Gaussian noise can model long-range dependence, that is, situations where the covariance function decreases slowly to zero, like a power law. This is the Joseph effect, referred to in the Introduction.

But what about the (marginal) distributions? The tails of the cumulative distribution function of a Gaussian random variable decrease exponentially fast. To get a power law decrease for the distribution tail, we will replace the Gaussian distribution by the more general stable distribution. Noises with stable distribution take large values with high probability, thus displaying the so-called Noah effect. They also have infinite variance. This means that one cannot use covariances anymore. It is a simple matter, however, to replace the Gaussian noises (or measures) that appear in the integral representations by stable noises. We will do this in a moment.

We start with a formal definition of the stable distribution. For simplicity, we focus in this paper solely on symmetric distributions. The characteristic function of a symmetric α -stable ($S\alpha S$) random variable is

$$Ee^{i\theta X} = e^{-\sigma^\alpha |\theta|^\alpha}, \quad -\infty < \theta < \infty.$$

There are two parameters, $\sigma > 0$ and $\alpha \in (0, 2]$. The parameter σ is the scale parameter and the parameter α is called the index of stability. Notice that the distribution is Gaussian when $\alpha = 2$ and in that case σ is a multiple of the standard deviation. In what follows, we typically suppose $0 < \alpha < 2$. Then the probability tails of X satisfy

$$P\{|X| > x\} \sim C_\alpha \sigma^\alpha x^{-\alpha} \text{ as } x \rightarrow \infty,$$

that is, as x increases, they decrease like a power law with exponent α . The smaller the value of α , the slower the decrease. Consequently the absolute p -moment of X ,

$$E|X|^p = \int_0^\infty P\{|X|^p > x\} dx = \int_0^\infty P\{|X| > x^{1/p}\} dx$$

is finite only if $p < \alpha$, because $P\{|X| > x^{1/p}\} \sim C_\alpha \sigma^\alpha x^{-\alpha/p}$ as $x \rightarrow \infty$. This implies that the second moment is always infinite when $\alpha < 2$ and the first absolute moment is infinite as well when $\alpha \leq 1$. The probability density function is known in closed form only for $\alpha = 1$; in this case, it is Cauchy with probability density function $2\sigma\pi^{-1}(x+4\sigma^2)^{-1}$. The α -stable density can be tabulated (see Hardin, Samorodnitsky, Taqqu [HST91] for a computer program that evaluates the density function for any α – even in the non-symmetric case.) Tables for the tails of the α -stable cumulative distribution function can be found in Dumouchel [Dum71].

Let us turn now to the integral representation. The idea now is to replace the white noise Gaussian measure $M_2(dx)$ in (4.1) by a white noise $S\alpha S$ measure $M_\alpha(dx)$, $0 < \alpha < 2$, and thus consider the integral representation

$$X_t = \int_{-\infty}^{+\infty} g_t(x) M_\alpha(dx). \quad (5.1)$$

The family of random variables $M_\alpha(\Delta)$, where Δ is any Borel subset of the line with a finite Lebesgue measure, is formally defined as follows. For any $n \geq 1$, $M_\alpha(\Delta_1), \dots, M_\alpha(\Delta_n)$ are independent if $\Delta_1, \dots, \Delta_n$ are disjoint and, for fixed Δ , $M_\alpha(\Delta)$ has a $S\alpha S$ distribution with characteristic function $Ee^{-\theta M_\alpha(\Delta)} = e^{-|\Delta||\theta|^\alpha}$, $-\infty < \theta < \infty$. Again, by taking $\Delta = [0, t]$, we can think of $M_\alpha([0, t])$ as $L_\alpha(t)$, the Lévy α -stable motion, which is the stable counterpart of Brownian motion. The process $L_\alpha(t)$ has stationary independent increments, and for fixed t , $L_\alpha(t)$ is $S\alpha S$ with characteristic function $Ee^{i\theta L_\alpha(t)} = e^{-|t||\theta|^\alpha}$. The process $L_\alpha(t)$ is H -sssi with $H \in \frac{1}{\alpha} \in (2, \infty)$. Heuristically, $M_\alpha(dx)$ can be thought as $dL_\alpha(x)$, that is, as the increment of Lévy α -stable motion at x .

For (5.1) to be defined we need

$$\int_{-\infty}^{+\infty} |g_t(x)|^\alpha dx < \infty, \quad -\infty < t < \infty.$$

$\{X_t\}$ is then a $S\alpha S$ process with joint characteristic function

$$Ee^{i\sum_{j=1}^n \theta_j X_{t_j}} = \exp\left\{-\int_{-\infty}^{+\infty} \left|\sum_{j=1}^n \theta_j g_{t_j}(x)\right|^\alpha dx\right\}, \quad -\infty < \theta_1, \dots, \theta_n < \infty. \quad (5.2)$$

The integral representation (5.1) is the natural extension of the Gaussian case to the α -stable case. We shall now present a number of examples.

Consider the Ornstein-Uhlenbeck process

$$X(t) = \int_{-\infty}^t e^{-\lambda(t-x)} M_\alpha(dx), \quad -\infty < t < \infty, \quad (5.3)$$

where $\lambda > 0$. In the Gaussian case $\alpha = 2$, this is the usual Ornstein-Uhlenbeck process, which is the continuous time counterpart of a Gaussian AR(1) sequence. Suppose now $0 < \alpha < 2$. The process is well-defined because $\int_0^t |e^{-\lambda(t-x)}|^\alpha dx < \infty$. It is stationary and Markov (See Adler, Cambanis and Samorodnitsky [ACS89].) The exponential kernel suggests that the dependence is not very strong (we will return to this point later).

One can also define the reverse Ornstein-Uhlenbeck process

$$X(t) = \int_t^\infty e^{\lambda(t-x)} M_\alpha(dx), \quad -\infty < t < \infty.$$

In the Gaussian case $\alpha = 2$, the Ornstein-Uhlenbeck process and the reverse Ornstein-Uhlenbeck process are the same process. In the stable case $\alpha < 2$, they are different processes, as we will be shown below.

The fact that the Gaussian Ornstein-Uhlenbeck process can be extended to the stable case $\alpha < 2$ in different ways, illustrates a typical phenomenon. Namely, in the stable case $\alpha < 2$, one often encounters different processes that become identical (up to a multiplicative constant) when α is set to 2.

We now turn to the *linear fractional stable noise* (LFSN) which is a stable counterpart of fractional Gaussian noise (FGN). It is defined for $0 < \alpha < 2$, as

$$X_t = \int_{-\infty}^{+\infty} g_t(x) M_\alpha(dx), \quad t = \dots, -1, 0, 1, \dots, \quad (5.4)$$

where

$$g_t(x) = a[(t+1-x)_+^{H-\frac{1}{\alpha}} - (t-x)_+^{H-\frac{1}{\alpha}}] + b[(t+1-x)_-^{H-\frac{1}{\alpha}} - (t-x)_-^{H-\frac{1}{\alpha}}], \quad (5.5)$$

and a and b are real numbers. Note that the kernel (5.5) is like the kernel (4.2) but with $1/2$ replaced by $1/\alpha$. To ensure that X_t is well defined, we must find the values of H that ensure $\int_{-\infty}^{+\infty} |g_t(x)|^\alpha dx < \infty$. A quick computation shows that as in the case $\alpha = 2$, we must have

$$0 < H < 1 \quad \text{if} \quad H \neq 1/\alpha.$$

When $H = 1/\alpha$, interpret (5.5) as $g_t(x) = (a-b)1_{(t,t+1]}$, so in that case X_t is $L_\alpha(t)$, a sequence of i.i.d. $S_\alpha S$ random variables (proportional to the 1-step increments of the Lévy stable motion $L_\alpha(t)$). Thus, when $H = 1/\alpha$, we can have $H \in (2, \infty)$.

As in the Gaussian case $\alpha = 2$, the behavior of the kernel (5.5) depends essentially on the sign of the exponent $H - \frac{1}{\alpha}$. Consider the three possibilities $H > \frac{1}{\alpha}$, $H < \frac{1}{\alpha}$ and $H = \frac{1}{\alpha}$. By analogy with the Gaussian case $\alpha = 2$, we say that we have long-range dependence when $H > \frac{1}{\alpha}$ and negative dependence when $H < \frac{1}{\alpha}$. (The case $H = \frac{1}{\alpha}$, as we noted, corresponds to independence). In the Gaussian case $\alpha = 2$, the permissible interval $(0,1)$ for H is divided in two subintervals of equal length $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$, the first corresponding to negative dependence, the second to long-range dependence. When $1 < \alpha < 2$, the subinterval $(0, \frac{1}{\alpha})$ for H where there is negative dependence is longer than the subinterval $(\frac{1}{\alpha}, 1)$ where there is long-range dependence. When $0 < \alpha < 1$, $\frac{1}{\alpha} > 1$, so the constraint $0 < H < 1$ effectively means that there can only be negative dependence. To understand this heuristically, note that when $0 < \alpha < 1$, there is large variability (the mean is infinite). This variability together with long-range dependence would make the integral (5.4) blow up.

Contrary to the Gaussian case $\alpha = 2$, in the stable case $0 < \alpha < 2$, different constants a and b in (5.5) yield different linear fractional stable noises (when $H \neq 1/\alpha$), that is, processes with different finite-dimensional distributions. More precisely, two LFSN's, one defined with (a, b) the other with (a', b') , are the same processes (up to a multiplicative constant) if and only if (i) $a = a' = 0$, or (ii) $b = b' = 0$, or (iii) $aa'bb' \neq 0$ and $a/a' = b/b'$. (See Cambanis and Maejima [CM89]). Thus, once again, we see that we have, in the stable case, a much richer family than in the Gaussian case.

The adjective “linear” in “linear fractional stable noise” is included in order to indicate that $X_t = \int_{-\infty}^{+\infty} g_t(x) M_\alpha(dx)$ with $g_t(x)$ given in (5.5) is a moving average. This specification is important, because contrary to the Gaussian case, the corresponding

harmonizable representation yields a totally different process. (See Cambanis and Soltani [CS84].)

The harmonizable representation is defined as

$$X_t = \int_{-\infty}^{+\infty} e^{i\lambda t} \frac{e^{i\lambda} - 1}{i\lambda} \frac{1}{|\lambda|^{H-1+\frac{1}{\alpha}}} \tilde{M}_\alpha(d\lambda) \quad (5.6)$$

where $0 < \alpha < 2$ and $0 < H < 1$. (Compare with (4.5)). Here $\tilde{M}_\alpha$ is a complex isotropic $S\alpha S$ random measure.

The formal definition of $\tilde{M}_\alpha$ is as follows: for each $n \geq 1$ and disjoint Borel sets with finite Lebesgue measure $\Delta_1, \dots, \Delta_n$, the random variables $\tilde{M}_\alpha(\Delta_1), \dots, \tilde{M}_\alpha(\Delta_n)$ are independent; moreover, $\tilde{M}_\alpha = M_\alpha^{(1)} + iM_\alpha^{(2)}$ where $M_\alpha^{(1)}$ and $M_\alpha^{(2)}$ are real-valued, and for any Δ , the joint characteristic function of the two-dimensional random vector $(M_\alpha^{(1)}(\Delta), M_\alpha^{(2)}(\Delta))$ is

$$E e^{i(\theta_1 M_\alpha^{(1)}(\Delta) + \theta_2 M_\alpha^{(2)}(\Delta))} = e^{-|\Delta| \int_{S_2} |\theta_1 s_1 + \theta_2 s_2|^\alpha dl} - \infty < \theta_1, \theta_2 < \infty,$$

where, as before, $|\Delta|$ denotes the Lebesgue measure (“length”) of Δ , S_2 is the unit circle centered at the origin and dl is the uniform probability measure on the circle S_2 .

The process (5.6) is called *harmonizable fractional stable noise* (HFSN). The reason that we have not defined X_t as

$$X_t = \operatorname{Re} \int_{-\infty}^{+\infty} e^{i\lambda t} \frac{e^{i\lambda} - 1}{i\lambda} \left(a \frac{1}{\lambda_+^{H-1+\frac{1}{\alpha}}} + b \frac{1}{\lambda_-^{H-1+\frac{1}{\alpha}}} \right) \tilde{M}_\alpha(d\lambda)$$

is that all values of a and b yield the same process (up to a multiplicative constant). See Cambanis and Maejima [CM89] for details.

The fact that the harmonizable and moving average representations yield different processes is not peculiar to fractional stable motions. It is a general fact, that besides trivial cases, an α -stable process, $0 < \alpha < 2$, that has a moving average representation cannot have a harmonizable representation and vice-versa. (Cambanis and Soltani [CS84]). This fact is in sharp contrast to the Gaussian case $\alpha = 2$.

Linear fractional stable noises and harmonizable fractional stable noises are not the only extensions of fractional Gaussian noise to the stable case $0 < \alpha < 2$. Consider, for example, the *sub-Gaussian* process

$$X_t = \sqrt{A} G_t, \quad t = \dots, -1, 0, 1, \dots \quad (5.7)$$

where $0 < \alpha < 2$, and A is an $\frac{\alpha}{2}$ -stable random variable totally skewed to the right, that is, A is a positive random variable with Laplace transform $E e^{-\theta A} = e^{-\theta^{\alpha/2}}$, $\theta \geq 0$. Let $\{G_t\}$ be the fractional Gaussian noise process. This process was defined at the end of Section 2, and is also given by the integral representation (4.1). Assume that A and

$\{G_t\}$ are independent. It then turns out that $X_t = \sqrt{A}G_t$ is a symmetric α -stable process. Even though the process $\{X_t\}$ does not have covariances, its dependence structure is directly influenced by that of the fractional Gaussian noise process $\{G_t\}$.

It is possible to go even further and consider a *sub-stable* process

$$X_t = A^{1/\alpha'} S_t, \quad t = \dots, -1, 0, 1, \dots$$

where $0 < \alpha < 2$, $2 > \alpha' > \alpha$, A is an α/α' stable random variable totally skewed to the right, that is, a positive random variable with Laplace transform $Ee^{-\theta A} = e^{-\theta^{\alpha/\alpha'}}$, $\theta \geq 0$. Take $\{S_t\}$ to be either a linear fractional α' -stable noise (defined by (5.4) and (5.5) with the α replaced by α'), or a harmonizable fractional α' -stable noise (defined by (5.6) with the α replaced by α'). Assume A independent of $\{S_t\}$. It turns out that X_t is a symmetric α -stable process. Clearly, its dependence structure is influenced by that of S_t , and it will be quite different, depending on whether S_t is a linear fractional stable noise or a harmonizable fractional stable noise.

In conclusion, we see that compared to the Gaussian case, we have many more choices in the α -stable case for modelling long-range dependence. On the other hand, because $\alpha < 2$, we cannot use covariances to describe the dependence. We have used the integral representation both to define the processes and to somewhat describe the dependence structure. We will see in the next section, that in the stable case, one can introduce some counterparts to the covariance.

6 How to describe the dependence in the α -stable case?

Suppose that X_t is α -stable with $0 < \alpha \leq 2$. We include the Gaussian case $\alpha = 2$ because we want to introduce measures of dependence valid for $\alpha \leq 2$ and compare them to the covariance when $\alpha = 2$.

The *covariation* is a possible extension of the covariance to the case $1 < \alpha < 2$. Suppose that $\{X_t, t = \dots, -1, 0, 1, \dots\}$ has the integral representation $X_t = \int_{-\infty}^{\infty} g_t(x) M_{\alpha}(dx)$. The covariation $[X_{t_1}, X_{t_2}]_{\alpha}$ of the random variable X_{t_1} with X_{t_2} is defined as

$$[X_{t_1}, X_{t_2}]_{\alpha} = \int_{-\infty}^{\infty} g_{t_1}(x)(g_{t_2}(x))^{\langle \alpha-1 \rangle} dx$$

where $a^{\langle p \rangle} = |a|^p \text{sign}(a)$ is the “signed power”. The covariation reduces to the covariance $EX_{t_1}X_{t_2} = \int_{-\infty}^{\infty} g_{t_1}(x)g_{t_2}(x)dx$ when $\alpha = 2$. The covariation appears naturally in the context of the regression of symmetric α -stable bivariate vectors with $\alpha > 1$, where one has

$$E[X_{t_2}|X_{t_1} = x] = \frac{[X_{t_2}, X_{t_1}]_{\alpha}}{[X_{t_1}, X_{t_1}]_{\alpha}}. \quad (6.1)$$

(See Kanter [Kan72]. For extensions of (6.1) to $\alpha \leq 1$ see Samorodnitsky and Taqqu [ST91a].) Observe that (6.1) reduces to the classical regression formula $E[X_{t_2}|X_{t_1} = x] = \frac{EX_{t_2}X_{t_1}}{EX_{t_1}^2}x$ when $\alpha = 2$.

The covariation, however, suffers from a number of drawbacks. While it is always defined when $\alpha > 1$, it is not always defined when $0 < \alpha \leq 1$. It is, in general, not symmetric ($[X_{t_1}, X_{t_2}]_\alpha \neq [X_{t_2}, X_{t_1}]_\alpha$) and not linear in the second argument ($[X_{t_1}, aX_{t_2} + bX_{t_3}]_\alpha \neq a[X_{t_1}, X_{t_2}]_\alpha + b[X_{t_1}, X_{t_3}]_\alpha$).

We will focus instead on a measure of *asymptotic* dependence. Consider

$$r(t) = Ee^{i(\theta_1 X_t + \theta_2 X_0)} - Ee^{i\theta_1 X_t} Ee^{i\theta_2 X_0}, \quad (6.2)$$

where θ_1 and θ_2 are real numbers. The function $r(t)$ is the difference between the joint characteristic function of (X_t, X_0) and the product of the marginal characteristic functions. Because $\{X_t\}$ is stationary, we can write

$$r(t) = e^{-A_\alpha}(e^{-I(t)} - 1) \quad (6.3)$$

where $A_\alpha = -\ln Ee^{i\theta_1 X_t} - \ln Ee^{i\theta_2 X_0}$ and

$$I(t) = -\ln Ee^{i(\theta_1 X_t + \theta_2 X_0)} + \ln Ee^{i\theta_1 X_t} + \ln Ee^{i\theta_2 X_0}. \quad (6.4)$$

Clearly $r(t) \equiv 0$ if X_t is independent of X_0 . Moreover, if $\alpha = 2$ and $EX_t X_0 \rightarrow 0$ as $t \rightarrow 0$, then $r(t) \approx EX_t X_0$ as $t \rightarrow 0$, that is $r(t)$ is asymptotically proportional to the covariance.

Notation: $a_t \approx b_t$ as $t \rightarrow \infty$ with $b_t \neq 0$ means that there is a constant $C \neq 0$ such that $\lim_{t \rightarrow \infty} b_t / C a_t = 1$. We write $a_t \approx 0$ if $a_t = 0$ for all large enough t .

The function $r(t)$ does not seem to be useful in the harmonizable and sub-Gaussian cases because it does not typically tend to zero as $t \rightarrow \infty$. (See Levy and Taqqu [LT91].) We shall focus therefore on moving averages.

If X_t is the Ornstein-Uhlenbeck process (5.3), then as $t \rightarrow \infty$,

$$\begin{cases} r(t) \approx e^{-\lambda t} & \text{if } 1 \leq \alpha \leq 2, \\ r(t) \approx e^{-\alpha \lambda t} & \text{if } 0 < \alpha \leq 1. \end{cases} \quad (6.5)$$

(Levy and Taqqu [LT91]) We see that in the stable case $0 < \alpha < 2$, $r(t)$ decays exponentially fast as in the Gaussian case. The rate of decay is λ , the same rate as in the Gaussian case if $1 \leq \alpha < 2$, but it is $\alpha\lambda$, that is, it is smaller than λ , if $0 < \alpha < 1$. We have no heuristic explanation for this difference. The important point however is that the decay is exponential, like the covariance in the case $\alpha = 2$. The $r(t)$ for the reverse Ornstein-Uhlenbeck process also satisfies (6.5) (but the exponentials have different multiplicative factors if $\alpha < 2$ indicating that the reverse Ornstein-Uhlenbeck is different from the Ornstein-Uhlenbeck when $\alpha < 2$). See again Levy and Taqqu [LT91] for details.

We now turn to the linear fractional stable noise defined by (5.4) and (5.5). In this case, as $t \rightarrow \infty$,

$$r(t) \approx \begin{cases} t^{\alpha H - \alpha} & \text{if } 0 < \alpha \leq 2, 0 < H < 1, H \neq \frac{1}{\alpha}, (\alpha, H) \notin \Sigma \\ t^{H - \frac{1}{\alpha} - 1} & \text{if } (\alpha, H) \in \Sigma \\ 0 & \text{if } H = \frac{1}{\alpha} \end{cases} \quad (6.6)$$

where

$$\Sigma = \{(\alpha, H): 1 < \alpha < 2, 0 < H < \frac{1}{\alpha(\alpha - 1)}\} \quad (6.7)$$

For computational details, see Astrauskas, Levy and Taqqu [ALT90].

Let us analyze (6.6). In the case $H = \frac{1}{\alpha}$, the X_t are i.i.d. and hence $r(t) \equiv 0$. When $(\alpha, H) \notin \Sigma$ the asymptotic behavior of $r(t)$ is like the asymptotic behavior of the covariance for fractional Gaussian noise, but with 2 replaced by α as expected. We see once again that $r(t)$ seems to be a good choice for describing the asymptotic behavior in the stable case. When $(\alpha, H) \in \Sigma$, that is, when $1 < \alpha < 2$ and H is below the cutoff point $1/(\alpha(\alpha - 1))$, we encounter a situation which has no analogue in the Gaussian case: $r(t)$ still behaves asymptotically like a power law but the exponent instead of being as expected $\alpha H - \alpha$ is $H - \frac{1}{\alpha} - 1$.

Let us first analyze the cutoff point $H_0 = 1/(\alpha(\alpha - 1))$. Observe that $H_0 = 1$ when $\alpha = \alpha_0 \equiv \frac{1+\sqrt{5}}{2} = 1.6180$. Since we must have $0 < H < 1$, (α, H) always falls in Σ whenever $1 < \alpha < \alpha_0$, that is, in this case the asymptotic behavior of $r(t)$ is always $t^{H - \frac{1}{\alpha} - 1}$.

However, when $\alpha_0 < \alpha < 2$, we must distinguish between two regions, one corresponding to $0 < H < H_0$ where the asymptotic behavior of $r(t)$ is $t^{H - 1/\alpha - 1}$, and the other corresponding to $H_0 < H < 1$ where the asymptotic behavior is $t^{\alpha H - \alpha}$.

As a final observation, note that when $\alpha_0 < \alpha < 2$, one has

$$0 < \frac{1}{\alpha} < H_0 < 1,$$

that is, it is only when there is long-range dependence ($H > 1/\alpha$) that one must distinguish between the two different rates of $r(t)$.

References

- [ACS89] R. Adler, S. Cambanis, and G. Samorodnitsky. Stable Markov processes. *Stoch. Proc. Appl.*, 34:1–17, 1989.
- [ALT90] A. Astrauskas, J.B. Levy, and M. S. Taqqu. The asymptotic dependence structure of the linear fractional Lévy motion. To appear, *Lietuvos Matematikos Rinkinys (Lithuanian Mathematical Journal)*, 1990.

- [Ast83] A. Astrauskas. Limit theorems for sums of linearly generated random variables. *Lithuanian Mathematical Journal*, 23(2):127–134, 1983.
- [BD87] P.J. Brockwell and R.A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 1987.
- [Ber90] J. Beran. Recent developments in location estimation and regression for long-memory processes. Preprint, 1990.
- [CM89] S. Cambanis and M. Maejima. Two classes of self-similar stable processes with stationary increments. *Stoch. Proc. Appl.*, 32:305–329, 1989.
- [CS84] S. Cambanis and A.R. Soltani. Prediction of stable processes: Spectral and moving average representation. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 66:593–612, 1984.
- [Dum71] W.H. Dumouchel. *Stable distributions in statistical inference*. PhD thesis, Univ. Ann Arbor, Univ. Microfilms, Ann Arbor, Mich., 1971.
- [Fel71] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, 3rd edition, 1971.
- [GJ80] C.W.J. Granger and R. Joyeux. An introduction to long-memory time series and fractional differencing. *J. Time Series Anal.*, 1:15–30, 1980.
- [Hos81] J.R.M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–176, 1981.
- [HST91] C.D. Hardin Jr., G. Samorodnitsky, and M.S. Taqqu. Numerical computations of non-linear stable regression functions. In G. Samorodnitsky S. Cambanis and M.S. Taqqu, editors, *Recent Results in Stable Processes and Related Topics*, Boston, 1991. Birkhäuser. To appear.
- [Kan72] M. Kanter. Linear sample spaces and stable processes. *J. Funct. Anal.*, 9:441–459, 1972.
- [LT91] J. Levy and M.S. Taqqu. A characterization of the asymptotic behavior of stationary stable processes. In G. Samorodnitsky S. Cambanis and M.S. Taqqu, editors, *Recent Results in Stable Processes and Related Topics*, Boston, 1991. Birkhäuser. To appear.
- [MV68] B.B. Mandelbrot and J.W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.
- [Sin76] Ya. G. Sinai. Self-similar probability distributions. *Theory Probability Appl.*, 21:64–80, 1976.

- [ST88] A. Samarov and M.S. Taqqu. On the efficiency of the sample mean in long memory noise. *Journal of Time Series Analysis*, 9:191–200, 1988.
- [ST91a] G. Samorodnitsky and M.S. Taqqu. Conditional moments and linear regression for stable random variables. To appear in *Stoch. Proc. Appl.*, 1991.
- [ST91b] G. Samorodnitsky and M.S. Taqqu. Probability laws with 1-stable marginals are 1-stable. Preprint 1989. To appear in the *Annals of Probability*, 1991.
- [ST93] G. Samorodnitsky and M.S. Taqqu. Non-Gaussian Stable Processes. In preparation, 1993.
- [Taq79] M.S. Taqqu. Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 50:53–83, 1979.
- [Taq81] M.S. Taqqu. Self-similar processes and related ultraviolet and infrared catastrophes. In *Random Fields: Rigorous Results in Statistical Mechanics and Quantum Field Theory*, pages 1027–1096, Amsterdam, 1981. Colloquia Mathematica Societatis Janos Bolyai, North-Holland. Vol.27 Book 2.
- [Taq86] M.S. Taqqu. A bibliographical guide to self-similar processes and long-range dependence. In E. Eberlein and M.S. Taqqu, editors, *Dependence in Probability and Statistics*, pages 137–162, Boston, 1986. Birkäuser.
- [Taq87] M.S. Taqqu. Random processes with long-range dependence and high variability. *J. Geophys. Research*, 92:9683–9686, 1987.
- [Taq88] M.S. Taqqu. Self-similar processes. In S. Kotz and N. Johnson, editors, *Encyclopedia of Statistical Sciences*, pages 352–357. Wiley, New York, 1988. Volume 8.
- [Yaj88] Y. Yajima. On estimation of a regression model with long-memory stationary errors. *Ann. Statistics*, 16:791–807, 1988.
- [Yaj89] Y. Yajima. Asymptotic properties of LSE in a regression model with long-memory stationary errors. Preprint, 1989.
- [Zol86] V.M. Zolotarev. *One-dimensional Stable Distributions*, volume 65 of *Translations of mathematical monographs*. American Mathematical Society, 1986.