Fast, Exact Synthesis of Gaussian and nonGaussian Long-Range-Dependent Processes

Matthew S. Crouse and Richard G. Baraniuk[*]

Corresponding Author:
Richard G. Baraniuk
Department of Electrical and Computer Engineering
Rice University
6100 South Main Street
Houston, TX 77005, USA

Ph: (713) 285-5132
Fax: (713) 524-5237
Email: richb@rice.edu
URL: www.dsp.rice.edu

Submitted to:
IEEE Transactions on Information Theory
August 1999

Abstract

$1/f$ noise and statistically self-similar processes such as fractional Brownian motion (fBm) are vital for modeling numerous real-world phenomena, from network traffic to DNA to the stock market. Although several algorithms exist for synthesizing discrete-time samples of a $1/f$ process, these algorithms are *inexact*, meaning that the covariance of the synthesized processes can deviate significantly from that of a true $1/f$ process. However, the Fast Fourier Transform (FFT) can be used to exactly and efficiently synthesize such processes in $O(N \log N)$ operations for a length-N signal. Strangely enough, the key is to apply the FFT to match the target process's covariance structure, not its frequency spectrum. In this paper, we prove that this FFT-based synthesis is exact not only for $1/f$ processes such as fBm, but also for a wide class of *long-range dependent* processes. Leveraging the flexibility of the FFT approach, we develop new models for processes that exhibit one type of fBm scaling behavior over fine resolutions and a distinct scaling behavior over coarse resolutions. We also generalize the method in order to exactly synthesize various *nonGaussian* $1/f$ processes. Our nonGaussian $1/f$ synthesis is fast and simple. Used in simulations, our synthesis techniques could lead to new insights into areas such as computer networking, where the traffic processes exhibit nonGaussianity and a richer covariance than that of a strict fBm process.

[*]This work was supported by the National Science Foundation, grant no. MIP–9457438, the Office of Naval Research, grant no. N00014–99–1–0813, and by DARPA/AFOSR, grant no. F49620-97-1-0513. We would also like to thank Dr. Rudolf H. Riedi for providing insightful comments and criticism of this work.

1 Introduction

Long-range dependent (LRD), $1/f$,[1] and self-similar processes occur frequently in numerous disciplines, including physics, chemistry, astronomy, economics, and biology [1–3]. In electrical engineering these processes have served as extremely useful models for characterizing textures in images [4–6] and noise in analog circuits [2]. In finance, these processes have been recognized as important for characterizing price volatility and market risk [7, 8]. They have even impacted DNA research [9].

Data networking conjures up many of the salient features of $1/f$ processes. Recently, traffic loads and interarrival times in data networks have been shown to exhibit LRD and self-similar behavior [10]. This behavior has proven to be a key factor driving network performance, yet it is ill-described using classical Poisson or Markov models. For a given mean traffic load, these classical models predict levels of network performance that are much more optimistic than the performance observed in practice [11].

The inadequacy of classical models has been a driving force behind the application of Gaussian $1/f$ models such as the fractional Brownian motion (fBm) process (power spectrum $1/|f|^{2H+1}$) and its increments process the fractional Gaussian noise (fGn) (power spectrum $1/|f|^{2H-1}$), $0 < H < 1$ [12]. These models have led to exciting new insights regarding the impact of LRD and self-similarity on network performance [10, 11].

1.1 Synthesis of Gaussian $1/f$ processes

In scenarios where closed-form analysis is intractable, fBm and fGn synthesis algorithms become extremely useful for generating synthetic data that can be used in Monte Carlo simulations [13]. Synthesis provides two key advantages over simulations using real traces. First, real traces can be quite cumbersome to measure and store. Second, the few, easy-to-understand parameters of these models allow us to generalize to hypothetical scenarios for which empirical data would be difficult or impossible to obtain. The parameters consist of the variance, an optional mean, and the Hurst parameter H, which controls the covariance.

Unfortunately, accurate discrete-time synthesis of fBm and fGn has been fraught with difficulty, particularly for the huge data sizes often required for accurate simulations. As we will examine

[1] The description $1/f$ is used for a process with a power spectrum that decays like $1/|f|^{\alpha}$, $-1 < \alpha < 3$.

in Section 2, methods such as Cholesky factorization via Levinson Durbin are *exact* (matching the second-order statistics without error), but computationally expensive ($O(N^2)$ for a length-N trace). Other approaches such as wavelet-based synthesis are fast ($O(N)$), but only approximate.

In this paper, we apply an algorithm based on the Fast Fourier Transform (FFT) to exactly synthesize discrete-time $1/f$ processes such as sampled fBm and fGn in only $O(N \log N)$ operations for a length-N trace. The FFT algorithm, described in detail in Section 3, operates as follows. To synthesize a stationary, length-N, jointly Gaussian signal, we embed its covariance matrix into a circulant covariance matrix of approximately twice the size. Since the FFT provides an eigenvalue decomposition for circulant matrices, we can synthesize the *embedding* signal efficiently via the FFT, with our desired signal obtained as the first N samples. In Figure 1(a) and (c), we provide visual examples of our fBm synthesis for $H = 0.6$ and $H = 0.9$.

This FFT approach and conditions for its exactness were developed in a general context in [14]. However, [14] focused on the less interesting case of Gaussian processes with quickly-decaying covariances (short-range dependence). In 1-d, efficient synthesis algorithms already exist for such processes. The work in [4] insightfully recognized the utility of the FFT for synthesizing fBm, but focused on 2-d fBm, a case for which the algorithm is not exact. As a result, this FFT synthesis approach has been little-used for synthesis of 1-d fBm and fGn despite its computational efficiency. Here, we prove that the FFT synthesis is exact not only for 1-d fBm and fGn, but for a host of Gaussian LRD processes.

1.2 Synthesis of general LRD processes

FBm and fGn models are limited in their ability to characterize many real-world signals. For example, fBm and fGn are Gaussian, while processes such as financial data and network traffic are inherently positive-valued and, hence, nonGaussian. Moreover, real-world data exhibit a more general dependency structure than can be modeled with a strict $1/f$ spectrum. For one, fBm and fGn have a fixed covariance structure whose shape is entirely determined by one parameter. Although real-world processes may exhibit the same asymptotic covariance decay as these models, their short-term covariance structure may differ substantially. With our FFT approach, these generalizations are relatively straightforward.

Leveraging the core FFT synthesis algorithm, we will develop a framework for modeling and

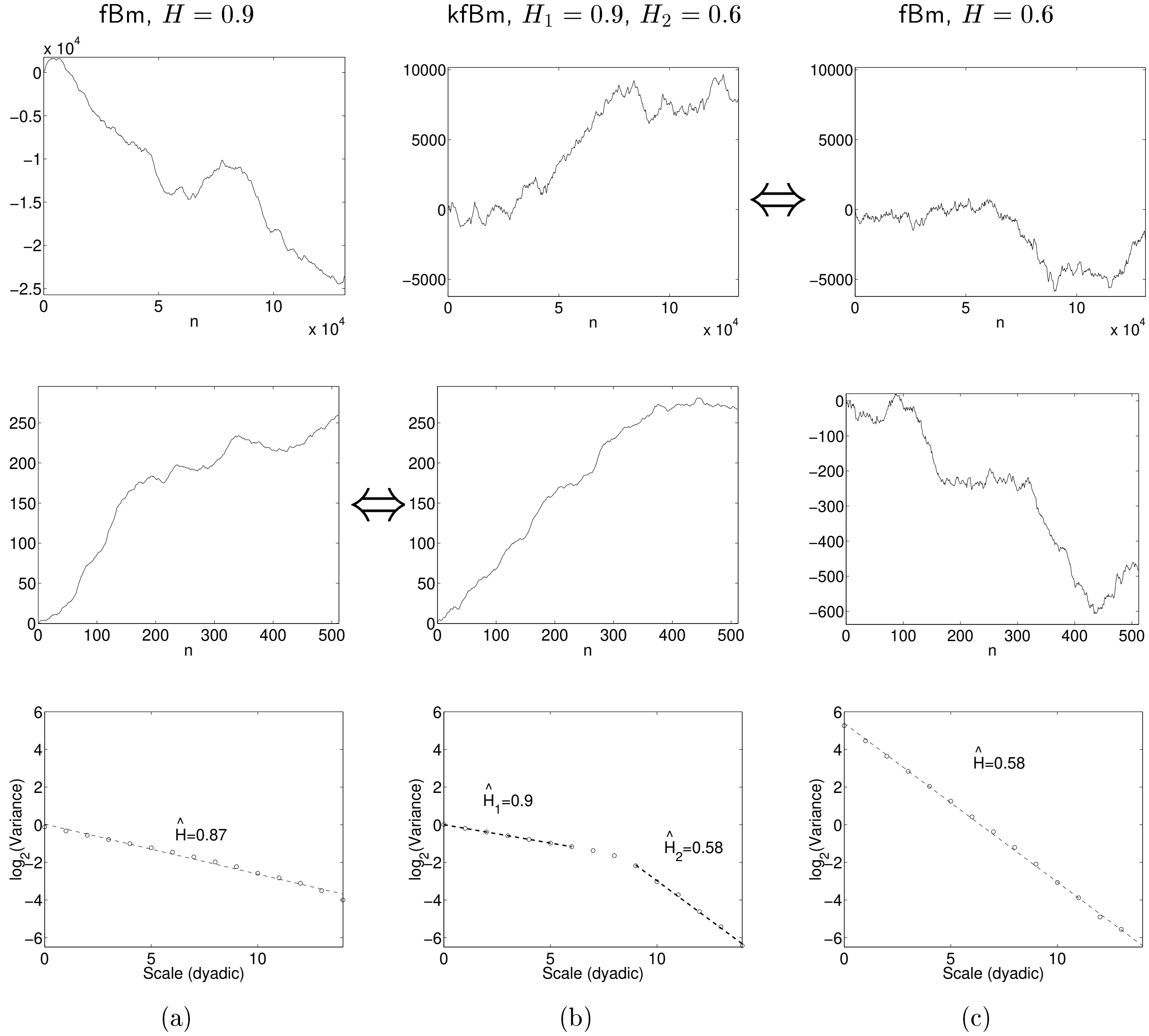


Figure 1: *Realizations of (a) Sampled fractional Brownian motion (fBm) with $H = 0.9$ and $\sigma_X^2 = 1$, (b) sampled kinked fractional Brownian motion (kfBm) with $H_1 = 0.9$, $H_2 = 0.6$, $\sigma_X^2 = 1$, and transition region from $n_1 = 64$ to $n_2 = 340$, and (c) sampled fBm with $H = 0.6$ and $\sigma_X^2 = 38.63$. All data sets were synthesized using the FFT method. Top: Plots of the entire length-$2^{17}+1$ traces. Over large time scales, the kfBm exhibits a smoothness and persistency similar to fBm with $H = 0.6$ (to the right). Middle: Plots zoomed-in to the first 512 samples. Over small time scales, the kfBm exhibits a smoothness and persistency similar to fBm with $H = 0.9$ (to the left). Bottom: Variance-time plots. The $\circ$'s are empirical variance-time plot measurements, and the dotted lines correspond to least-squares linear fits to estimate H. For all traces, the empirical variance-time plots and H estimates closely match the desired theoretical behavior. The kfBm trace demonstrates a linear scaling behavior corresponding to $H = 0.9$ over fine resolution and $H = 0.6$ over coarse resolutions, joined by a "kink" in the middle. See Section 4.3 for more details.*

synthesizing a wide class of Gaussian LRD processes. We will derive the covariance structure necessary to obtain self-similar scaling over two different sets of time scales. This includes processes that exhibit one type of fBm-like scaling over short time intervals and an entirely different scaling over long time intervals. This scaling can be characterized through the *variance-time plot* (see Section 2.3.3), a plot of the variance of the increments of a process as a function of aggregation size. We call the resulting model *kinked fBm* (kfBm) because of its variance-time plot behavior (see Figure 1(b)). This behavior has been discovered in network data [15] and may have important consequences for network management [16]. In addition, the kfBm model could provide a useful link to the processes that occur in science, where some claim that most "fractal" data sets exhibit a power-law scaling behavior only over a limited scaling range [17].

We will also extend our framework to model and synthesize nonGaussian LRD processes. To create a nonGaussian process with a target LRD covariance, we simply pass a Gaussian process through an appropriate nonlinear transformation, with the transformation chosen to convert the Gaussian marginal to the desired nonGaussian marginal [18, 19]. Transformation in hand, we need only "prewarp" the covariance of the input Gaussian process to account for the effect of the transformation on the covariance of the nonGaussian output process. In theory, this approach can be used for any finite-variance nonGaussian process and a wide class of covariances functions. Moreover, the prewarping can be calculated analytically for many nonGaussian processes, including processes with lognormal, exponential, uniform, and finite-variance Pareto marginal densities. For these densities, we prove that our approach can exactly synthesize nonGaussian processes with the same $1/f$-type spectrum and second-order scaling behavior as fBm.

1.3 Paper Organization

In Section 2, we introduce our notation, define self-similarity, the $1/f$ property, and LRD, and show how they are inter-related. We also review several current synthesis techniques for discrete-time $1/f$ noise. In Section 3, we develop the FFT-based synthesis approach for Gaussian processes. We present conditions for when this approach is exact, and (with help from Appendix A) prove that fGn and fBm satisfy these conditions. In Section 4, we look at two generalizations of fBm and demonstrate in which cases synthesis is exact for these generalizations. In Section 5, we develop a method for synthesizing nonGaussian processes and enumerate several nonGaussian processes that can be treated easily in closed form. In Appendix B we derive conditions for the nonGaussian $1/f$ synthesis to be exact and demonstrate that the enumerated nonGaussian densities jump through these hoops. In Section 6, we discuss potential extensions and further applications of this work.

2 Characterization and Synthesis of $1/f$ Processes

2.1 Notation

We will use $\equiv$ to define an expression, $\propto$ to mean proportional to, and $\simeq$ to denote equivalent asymptoptic behavior. The expression $X(t),\ t \in \mathbb{R}$ will denote a real stationary Gaussian process in continuous time; the expression $X[n],\ n \in \mathbb{Z}$ will denote a separate real stationary Gaussian process in discrete time. $\Delta X[n] \equiv X[n+1] - X[n]$ and $\Delta^2 X[n] \equiv \Delta X(n+1) - \Delta X[n]$ represent the *first* and *second differences* of $X[n]$, respectively. In some cases, we will identify $X[n]$ with the increments process $X(nT) \equiv B((n+1)T) - B(nT)$ of a nonstationary Gaussian process $B(t)$. (Generally, X will correspond to an fGn-like stationary process and B an fBm-like nonstationary process.) When dealing specifically with fBm or fGn, we will add the subscript H to signify their dependence on the Hurst parameter.

We will reserve $Y[n]$ for the stationary nonGaussian LRD process formed via a nonlinear transformation of $X[n]$ (see Section 5). For the first-order distribution function and probability density function (pdf) of $Y[n]$, we write $F_Y(y) \equiv \Pr[Y[n] \leq y]$ and $f_Y(y) \equiv \frac{d}{dy} F_Y(y)$, respectively. For the distribution function of the standard Gaussian random variable, we write $\Phi(x)$. The notation "$\stackrel{fd}{=}$" represents equality in finite-dimensional distributions.

Finally, we will find matrix and vector notation useful. We will use boldface notation for vectors; for instance, $\mathbf{X} \equiv [X[0], X[1], \ldots, X[N-1]]^T$ and similarly for the covariance vector $\mathbf{r}_X \equiv [r_X[0], r_X[1], \ldots, r_X[N-1]]$. The term $\mathbf{0}$ will represent the zero vector. We will denote the identity matrix by I, the covariance matrix for $\mathbf{X}$ by R_X, and the matrices corresponding to the FFT and IFFT operations [40] by F and F^*, respectively, with F^* the conjugate transpose of F.

We now provide a brief review of some of the properties of continuous-time and discrete-time $1/f$ processes, with a focus on fBm and fGn. Although a continuous-time treatment is key to understanding and analyzing the basic properties of $1/f$ processes, in the end a discrete-time treatment is needed to accurately analyze and synthesize these processes using the digital computer.

2.2 Continuous-time $1/f$ processes

Consider a second-order stationary Gaussian process $X(t)$ with covariance function

$$r_X(\tau) \equiv \mathbb{E}[X(t)X(t-\tau)] - \mathbb{E}[X(t)]^2 \tag{1}$$

and Fourier power spectrum

$$S_X(f) \equiv \int_{-\infty}^{\infty} r_X(\tau) e^{-j2\pi f\tau} d\tau, \ -\infty < f < \infty. \tag{2}$$

Such a process is called a continuous-time $1/f$ *process* if its Fourier power spectrum takes the form [20]

$$S_X(f) \propto \frac{1}{|f|^\alpha}, \ -\infty < f < \infty, \ -1 < \alpha < 3. \tag{3}$$

Right at the outset we encounter some technical difficulties. For the case $1 < \alpha < 3$, $S_X(f)$ is not integrable near the origin and hence, strictly speaking, not a valid power spectrum. This is usually attributed to an inherent nonstationarity in the process, in which case (1) and (2) are not valid definitions. In this case, (3) must be interpreted as a *time-averaged* power spectrum [21].

When $-1 < \alpha < 1$, $S(f)$ is not integrable as $f \to \infty$, since the high-frequency energy decays too slowly. Thus, $1/f$ processes for $-1 < \alpha < 1$ have infinite power and are well defined only in a generalized sense.[2] Nevertheless, these processes can still provide extremely useful models. For instance, continuous-time white Gaussian noise, $1/f$ noise with $\alpha = 0$, is indispensable in communication theory as a model for channel noise. Let us meet the Gaussian $1/f$ processes for $\alpha \neq 0$.

2.2.1 Fractional Brownian motion (fBm): $1 < \alpha < 3$

The *Gaussian* process used for modeling $1/f$ noise with $1 < \alpha < 3$ is known as *fractional Brownian motion (fBm)* [12]. The fBm process $B_H(t)$ is a continuous, jointly Gaussian process defined via the *self-similarity* property

$$B_H(at) \stackrel{fd}{=} a^H B_H(t), \tag{4}$$

with $H = 2\alpha + 1$, $0 < H < 1$, denoting the *Hurst parameter*. In words, a zoom of fBm by a factor of a leads to an fBm that, after scaling by a^H, has statistical properties identical to the original fBm.

[2] See [20] for one such definition that, in addition, treats the case of $1/f$ processes for $\alpha \leq -1$ and $\alpha \geq 3$.

For fBm the time-averaged power spectrum is given by (3) with $\alpha = 2H+1$ and the nonstationary covariance function by [12]

$$r_{B_H}(s,t) \equiv \mathbb{E}\left[B_H(t)B_H(s)\right] = \frac{\sigma^2}{2}\left[|s|^{2H} + |t|^{2H} - |s-t|^{2H}\right]. \tag{5}$$

2.2.2 Fractional Gaussian noise (fGn): $-1 < \alpha < 1$

The *Gaussian* $1/f$ process with $-1 < \alpha < 1$ can be viewed as the derivative of an fBm with $1 < \alpha < 3$, since a derivative provides a multiplicative factor of $1/f^2$ in the power spectrum. In this case, the decay parameter α relates to the Hurst parameter H via $\alpha = 2H - 1$. When $0 < \alpha < 1$, using the generalized Fourier Transform inverse for $1/|f|^\alpha$, we obtain the stationary covariance function for the idealized derivative of fBm:

$$r(\tau) \propto \tau^{\alpha-1},\ 0 < \alpha < 1. \tag{6}$$

As noted previously, for $-1 < \alpha < 1$, a process with strict power-law decay can be defined only in a generalized sense. Therefore, we find it useful to construct an *approximation* to $1/f$ noise, $-1 < \alpha < 1$, using the continuous-time increments process for fBm,

$$X_H(t) \equiv B_H(t+T) - B_H(t), \tag{7}$$

which is a stationary Gaussian process known as *fractional Gaussian noise* (fGn) [12]. It has a spectrum of the form [21]

$$S_{X_H}(f) = c\frac{\sin^2(\pi f T)}{|f|^{2H+1}}, \tag{8}$$

with c a proportionality constant. For $f \ll \frac{1}{T}$, we have $S_{X_H}(f) \propto 1/|f|^{2H-1}$. The accompanying covariance function is given by

$$r_{X_H}(\tau) = \frac{\sigma_X^2 T^{2H}}{2}\left[|\tau+1|^{2H} + |\tau-1|^{2H} - 2|\tau|^{2H}\right],\ 0 < H < 1. \tag{9}$$

Note that $r_{X_H}(\tau) \simeq \tau^{2H-2}$ as $\tau \to \infty$, meaning fGn has the same asymptotic covariance decay as the idealized derivative of fBm.

2.2.3 Long-range dependence

For $1/2 < H < 1$, the slow power-law decay of $r_{X_H}(\tau)$ in (6) and (9) implies that the integral of the covariance function $\int_0^{T_0} r_{X_H}(\tau)\,\mathrm{d}\tau$ will diverge to ∞ as $T_0 \to \infty$. For any stationary process X, this divergence is known as *long-range dependence* (LRD) [22]. LRD is equivalent to a singularity in the DC component of the spectrum, $S(0) = \infty$, and is a key feature of fGn with $1/2 < H < 1$.

Since the covariance function of an LRD process decays extremely slowly, LRD processes have a great deal of "memory." This is a key property in scenarios where LRD occurs, such as in data networking and the financial markets. In networking, LRD in traffic loads can lead to poor network performance, since the packets arrive in bursts over long time scales [11]. In financial markets, the memory aspect of LRD is potentially useful for predicting volatility and potential risk [7].

LRD is difficult to characterize using classical models such as AR or Markov models, since classical models exhibit a much faster exponential covariance decay, and robust model parameter estimation for highly correlated data is difficult. For instance, approximating an LRD $1/f$ process using an autoregressive moving average (ARMA) model is equivalent to fitting a slowly-decaying power-law covariance function using quickly decaying exponentials — an impossible task for a finite-order model. This is one reason why fGn and fBm have recently come to the fore as models for LRD processes.

2.3 Discrete-time properties

A discrete-time treatment is needed to accurately analyze and synthesize $1/f$ processes using the digital computer. To begin, we will examine the second-order behavior of two discrete-time processes: the sampled fBm process $B_H(nT)$ and the fGn process (7) sampled with period T known as *discrete fractional Gaussian noise* (dfGn):

$$X_H^{(T)}[n] \equiv X(nT) = B_H((n+1)T) - B_H(nT), \; n = 0, 1, \ldots \tag{10}$$

For clarity of presentation, we will replace $X_H^{(T)}[n]$ with $X_H[n]$ when clear from context.

To model these processes, we will develop three equivalent second-order characterizations: the power spectrum, the covariance function, and the variance-time plot. An understanding of these different characterizations provides insight into the various approaches to modeling and synthesizing LRD processes.

2.3.1 The power spectrum and why we should avoid it

In practice, we can only work with discrete-time, finite-length data, so the continuous Fourier power spectrum is of little practical use. A direct frequency-domain analysis or synthesis requires the Discrete Fourier Transform (DFT), which we will call the FFT for convenience.[3] The forward FFT can be written as

$$\lambda[k] \equiv \sum_{n=0}^{N-1} r[n] \exp\left(-j\frac{2\pi}{N}nk\right), \; k = 0, 1, \ldots, N-1, \tag{11}$$

and the inverse FFT (IFFT) written as

$$r[n] \equiv \frac{1}{N}\sum_{k=0}^{N-1} \lambda[k] \exp\left(j\frac{2\pi}{N}nk\right), \; n = 0, 1, \ldots, N-1. \tag{12}$$

In vector from, (11) and (12) are written as $\boldsymbol{\lambda} = F\mathbf{r}$ and $\mathbf{r} = F^*\boldsymbol{\lambda}$, respectively.

We define the *FFT power spectrum* (periodogram) of $X[n]$ as [23]

$$\widetilde{S}_X(f_k) \equiv \mathbb{E}\left[\left|\frac{1}{N}\sum_{n=0}^{N-1} X[n]\exp(-j2\pi n f_k)\right|^2\right], \; f_k = \frac{k}{N}, \; k = 0, 1, \ldots, N-1, \tag{13}$$

the expected value of the squared magnitude of the FFT of a length-N realization of the process $X[n]$. The FFT power spectrum provides approximate values of the Discrete Time Fourier Transform (DTFT) power spectrum at the frequencies $f = f_k$.

The FFT power spectrum of a length-N dfGn sequence can be expressed in terms of the Fourier spectrum of fGn (8) [23]

$$\begin{aligned}\widetilde{S}_{X_H}(f_k) \;&=\; \left(\frac{1}{N}\frac{\sin^2(\pi f N)}{\sin^2(\pi f)} \circledast \frac{c}{T^2}\sum_{m=-\infty}^{\infty} \sin^2(\pi(f-m))\left|\frac{f-m}{T}\right|^{-(2H+1)}\right)\Bigg|_{f=f_k}, \\ & \qquad\qquad f_k = \frac{k}{N}, \; k = 0, 1, \ldots, N-1,\end{aligned} \tag{14}$$

with $\circledast$ the circular convolution operator. The infinite sum in (14) is caused by aliasing of higher frequencies onto lower frequencies (due to sampling), and the convolutional smearing is caused by the effects of finite-length data. This complicated expression deviates noticeably from a strict $1/f$ behavior and is difficult to calculate. Even worse, it is a nontrivial function of the data length N.

Of course, the same types of sampling effects and smearing occur in the FFT power spectrum of sampled fBm, except with an added interpretational difficulty — fBm is nonstationary, and the notion of power spectrum applies only to stationary processes. Hence, the inherent difficulties of using and interpreting the FFT power spectrum suggest that we should look to alternative second-order characterizations of dfGn and sampled fBm.

[3]The FFT is simply an efficient $O(N \log N)$ algorithm for computing an N-point DFT.

2.3.2 The covariance function and why we should use it

Unlike the spectrum, which is warped and smeared by sampling, the covariance function for sampled fBm is obtained simply by plugging $s = nT$ and $t = mT$ into (5). Similarly, for dfGn the covariance function is found by substituting nT for τ in (9). Moreover, we can be confident that if we exactly match the covariance, then we have exactly matched all second-order statistics of the process, including the spectrum. In this paper, we will focus on modeling the covariance behavior of dfGn, since dfGn is easier to synthesize than sampled fBm and can be converted directly to sampled fBm by cumulative summation.

Since the choice of the sampling period T effects only a constant scaling factor in the covariance of dfGn, without loss of generality we can set $T = 1$ in (9) to obtain

$$r_{X_H}[n] = \frac{\sigma_X^2}{2}\left[|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}\right], \ 0 < H < 1. \tag{15}$$

For $0 < H < 1/2$ and $n \neq 0$, the covariance function is negative. For $1/2 < H < 1$, the covariance function is positive and decreasing in $|n|$ with an asymptotic power-law decay n^{2H-2} as $n \to \infty$. This decay implies that dfGn exhibits LRD for $1/2 < H < 1$ since, in accordance with the continuous-time LRD definition, a discrete-time process X is LRD if $\sum_{n=0}^{N} r_X[n]$ diverges as $N \to \infty$.

Note that, strictly speaking, dfGn has a mean of zero. However, in networking and other applications, non-zero mean processes with the covariance structure of dfGn prove useful. In this paper, we will not discriminate between zero and non-zero mean dfGn.

2.3.3 Variance-time plot and generalizations of fBm and fGn

An alternative characterization of the second-order properties of fBm and fGn can be obtained using the self-similarity property of fBm (4)

$$\mathrm{Var}[B_H(t+s) - B_H(t)] = |s|^{2H}\mathrm{Var}[B_H(t+1)) - B_H(t)] = |s|^{2H}\sigma_X^2, \tag{16}$$

or equivalently for dfGn with $T = 1$

$$\mathrm{Var}\left[\sum_{n=n_0}^{n_0+m} X_H[n]\right] = m^{2H}\sigma_X^2 \tag{17}$$

for $m \in \mathbb{N}$.

The latter property can be measured by forming the aggregated processes[4]

$$X_H^{(mT)}[n] \equiv \frac{1}{m} \sum_{i=(n-1)m+1}^{nm} X_H[i] \tag{18}$$

and tracking how their variances decrease with the aggregation variable m. It follows from (17) that

$$\mathrm{Var}\left[X_H^{(mT)}[n]\right] = m^{2H-2}\sigma_X^2. \tag{19}$$

Hence, a log-log plot of the variance of $X_H^{(mT)}[n]$ as a function of m — known as a *variance-time plot* — will have a slope $2H-2$. The variance-time plot can be applied to nonGaussian and non-zero-mean data as well. Processes whose variance scales with a power of $2H-2$ have been termed *exactly second-order self-similar (ESS) processes* [24].[5] Since ESS processes have power spectra identical to dfGn (a Gaussian $1/f$ noise), ESS processes do correspond to a type of nonGaussian $1/f$ noise. However, considering the difficulties associated with the FFT power spectrum (e.g., it is not truly $1/f$), the terminology ESS provides a more precise characterization.

For more general processes, we can define the *structure function* g_B [25, 26] such that

$$\mathrm{Var}\left[B(t+s) - B(t)\right] = g_B(s)\sigma_X^2 \tag{20}$$

and for $T=1$ (now in a discrete-time setting)

$$\mathrm{Var}\left[X^{(m)}[n]\right] = \frac{g_B[m]}{m^2}\mathrm{Var}[X[n]]\,. \tag{21}$$

When $g_B(s) = \sigma_X^2|s|^{2H}$, we have the self-similarity condition for fBm and fGn. The structure function directly describes the scaling behavior of a given process and provides a more intuitive description of LRD behavior than the covariance function. The structure function converts directly to the covariance function via one-half the central second difference

$$r_X[n] = \frac{1}{2}\left(g_B[n+1] - 2g_B[n] + g_B[n-1]\right). \tag{22}$$

We will use the structure function in Section 4 to characterize more general types of LRD processes.

[4]Recall from (10) that we write $X_H[m]$ instead of $X_H^{(T)}[m]$, with T our dfGn sampling rate.

[5]Strictly speaking, the Hurst parameter H only applies to fBm and fGn. However, we will use it in the context of more general LRD data to characterize the covariance decay.

2.4 Discrete-time synthesis algorithms for Gaussian $1/f$ noise

We now describe several commonly used alternatives for synthesizing the Gaussian $1/f$ processes corresponding to sampled fBm and dfGn. It is simple to convert between these processes using (10), so a synthesis algorithm for one can synthesize the other as well.

2.4.1 Cholesky factorization

This method is exact [5, 27], but slow, requiring $O(N^2)$ computations to form a length-N dfGN vector $\mathbf{X}$. From (9), we form the Toeplitz covariance matrix R_{X_H} for dfGn and then factor it into $R_{X_H} = QQ^T$ via the Levinson-Durbin algorithm (direct Cholesky factorization is even slower — $O(N^3)$). We then form $\mathbf{X} = Q\mathbf{V}$, with $\mathbf{V}$ a white Gaussian noise vector, to obtain the desired dfGn process.

2.4.2 FFT synthesis based on the power spectrum

If one is to use the FFT power spectrum for synthesis, a reasonable approach is to ensure that the FFT power spectrum of the synthesized process has properties equivalent to those of the FFT power spectrum of the data to be modeled. Paxson [28] takes this approach, performing an FFT-based synthesis of dfGn under the assumption that for a single realization the FFT power spectrum at $S_{X_H}(f_k)$ is independent and exponentially distributed with mean given by the infinite sum in (14). He thus generates exponentially distributed magnitudes in each FFT frequency bin, multiplies each magnitude by a random phase, and then inverts to obtain the desired dfGn signal.

Although fast, Paxson's approach does not account for the effects of the aliasing in (14) nor for the fact that the power in different FFT frequency bins can be correlated for dfGn. Furthermore, the output process has an undesirable circulant covariance structure (see Section 3), which among other things implies that the end of the synthesized dfGn trace will often be highly correlated with the beginning.

In passing, we note that any direct FFT-based synthesis of sampled fBm through this approach is undesirable because the same sampling and smearing effects occur. Moreover, any process synthesized by this approach will be stationary, whereas fBm is nonstationary.

2.4.3 Random midpoint displacement

The next technique [29] synthesizes the values of a sampled fBm trace at a few dispersed points and then successively refines (or randomly interpolates) the values between the synthesized points.

Since it does not attempt to model all cross-covariances between all fBm samples, it is inexact and in practice can deviate noticeably from exact fBm.

2.4.4 Parametric approaches

Corsini and Saletti [30] apply to white noise a discretized set of analog filters with logarithmically-spaced poles and zeros in order to generate a process with an approximate $1/|f|^{\alpha}$ power spectrum for $-2 < \alpha < 2$. Aside from the aliasing and finite-length concerns mentioned earlier, this approach is inexact due to the discretization of analog filters.

The fractionally-differenced autoregressive integrated moving average (FARIMA) model combines an ARMA model for modeling short-range dependence with a fractional differencing d operation for generating LRD [31, 32]. Thus, the FARIMA model can approximate not only $1/f$ noise, but also a wide class of LRD processes. However, estimation of ARIMA parameters can be difficult, and exact synthesis of ARIMA data is $O(N^2)$ complexity.

2.4.5 Self-similar basis functions

The remaining two techniques construct $1/f$ noise from self-similar basis functions that are spaced logarithmically in frequency. The first uses wavelets and relies on the fact that the discrete wavelet transform (DWT), in $O(N)$ computations, provides an approximate Karhunen-Loève transform (KLT) for $1/f$ processes, including fBm and fGn [20, 33, 34].

Since the DWT is only an approximate KLT for $1/f$ processes, wavelet-based synthesis under the assumption of independent wavelet coefficients does not lead to exact $1/f$ behavior [20]. Moreover, the theoretical DWT synthesis results of [20] are derived only for continuous-time $1/f$ noise using an infinite number of wavelet scales. On a computer, only a finite number of wavelet transform scales can be used. Furthermore, recall from (14) that in discrete time the target power spectrum is not exactly $1/f$. These facts add to the potential synthesis inaccuracies resulting from the KLT approximation.

Wavelet-based synthesis of fBm and fGn involves further inaccuracies. FBm is a trended process — its variance increases with time, whereas wavelet-synthesized data is not trended. Moreover, fGn is stationary, whereas wavelet-synthesized data is nonstationary in covariance [20, 35]. The *time-averaged* covariance of the wavelet-synthesized process will approximately correspond to the covariance of fGn, but two neighboring samples of this process, depending on their location, could range from highly correlated to almost independent.

The randomized Weierstrass approach [36] synthesizes fBm by summing scaled sinusoidal functions at logarithmically-spaced frequencies with random height and phases. Again, the approach is inexact, although its deviation from true fBm has not been well quantified.

3 FFT-based Synthesis of Gaussian Processes

The FFT synthesis approach is based on the concept of embedding a Toeplitz covariance matrix into a circulant covariance matrix. This has been studied previously in several contexts [4,14,37,38]. The work [37,38] provides procedures for embedding an $N \times N$ positive-definite covariance matrix R into a circulant matrix of size $M \times M$, $M \geq 2N-2$. Unfortunately, M can become extremely large depending on the condition number of the matrix R.

The work [14] develops several conditions under which the embedding is exact for $M = 2N - 2$ and uses this embedding to derive an algorithm for general FFT-based synthesis of Gaussian processes. The work in [4], independently derived, specifically applies the approach to synthesize 2-d fBm, but does not rigorously analyze the exactness of the synthesis. Indeed, the algorithm is not always exact for synthesizing 2-d fBm. We will follow along the lines of [14], but will specialize to the case of $1/f$ and LRD data, demonstrating that the FFT-based synthesis is exact for many 1-d Gaussian and nonGaussian LRD processes of interest.

3.1 FFT properties and application

The FFT (11) has many magical and wondrous properties. Among them is the eigenvalue decomposition it provides for an $M \times M$ circulant matrix C [39]. If we denote the top row as $[c[0]\ c[1]\ \ldots\ c[M-1]]$, we can write the entire circulant matrix as

$$C = \begin{bmatrix} c[0] & c[1] & c[2] & \cdots & c[M-1] \\ c[M-1] & c[0] & c[1] & \cdots & c(M-2) \\ c[M-2] & c[M-1] & c[0] & \cdots & c[M-3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c[1] & c[2] & c(3) & \cdots & c[0] \end{bmatrix}. \tag{23}$$

C then has an eigenvalue decomposition $C = \frac{1}{M} F \Lambda F^*$, with Λ a diagonal matrix. The vector $\boldsymbol{\lambda} \equiv [\Lambda(1,1)\ \Lambda(2,2)\ \ldots\ \Lambda(M,M)]$ is the FFT of the vector $[c[0]\ c[1]\ \ldots\ c[M-1]]$.

This is exactly the decomposition we desire for the covariance matrix R_X of a vector $\mathbf{X} \sim N(\mathbf{0}, R_X)$ drawn from a stationary process with underlying covariance function $r_X[n]$. If R_X were equal to some circulant matrix C, we could synthesize $\mathbf{X}$ via $\mathbf{X} = \mathcal{F}\left(\frac{1}{M}\Lambda\right)^{1/2}\mathbf{W}$, with $\mathbf{W} \sim N(\mathbf{0}, I)$. We can easily check that $\mathbf{X}$ has the correct covariance in this case:

$$\mathbb{E}[\mathbf{X}\mathbf{X}^H] = \frac{1}{N}\mathcal{F}\Lambda^{1/2}\mathbb{E}[\mathbf{W}\mathbf{W}^H]\Lambda^{1/2}\mathcal{F}^* = C. \tag{24}$$

For a real second-order-stationary process, the covariance matrix R_X for a length-N vector $\mathbf{X}$ is formed from the covariance function $r_X[n]$ according to

$$R_X = \begin{bmatrix} r_X[0] & r_X[1] & r_X[2] & \cdots & r_X[N-1] \\ r_X[1] & r_X[0] & r_X[2] & \cdots & r_X[N-2] \\ r_X[2] & r_X[1] & r_X[0] & \cdots & r_X[N-3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_X[N-1] & r_X[N-2] & r_X[N-3] & \cdots & r_X[0] \end{bmatrix}. \tag{25}$$

This matrix is Toeplitz, but not necessarily circulant. In general, the FFT will diagonalize such a matrix only as $N \to \infty$.

However, we can *embed* R_X into a $2N-2 \times 2N-2$ circulant matrix R_Z with first row equal to

$$\mathbf{r}_Z \equiv [r_X[0]\ r_X[1]\ \ldots\ r_X[N-2]\ r_X[N-1]\ r_X[N-2]\ \ldots\ r_X[2]\ r_X[1]]^T. \tag{26}$$

The entry in the ith row and jth column of R_Z is given by

$$R_Z[i,j] = \begin{cases} r_X[|i-j|], & i,j \le N \text{ or } i,j > N \\ r_X[|2M-i-j|], & j \le N, i > N \text{ or } i \le N, j > N. \end{cases} \tag{27}$$

Note that $R_Z[i,j] = R_X[i,j]$ for $i,j \le N$. Thus, our plan will be to synthesize a vector $\mathbf{Z} \sim N(\mathbf{0}, R_Z)$ of length $2N-2$ and then form $\mathbf{X}$ as the first N elements of $\mathbf{Z}$.

In accordance with (24) we could construct $\mathbf{Z} = \mathcal{F}(\frac{1}{2N-2}\Lambda)^{1/2}\mathbf{U}$, with $\mathbf{U} \sim N(\mathbf{0}, I)$. Unfortunately, since $\mathcal{F}$ is a complex matrix, $\mathbf{Z}$ is in general complex. However, if we add a imaginary component $i\mathbf{V}$, $\mathbf{V} \sim N(\mathbf{0}, I)$, $\mathbb{E}[\mathbf{U}\mathbf{V}^T] = 0$, to the spectrum (corresponding to a random phase), we can generate *two* real sequences with the desired covariance: $\mathbf{X}_1 = \mathrm{Re}\Big(\mathcal{F}(\frac{1}{2N-2}\Lambda)^{1/2}(\mathbf{U} + i\mathbf{V})\Big)$ and $\mathbf{X}_2 = \mathrm{Im}\Big(\mathcal{F}(\frac{1}{2N-2}\Lambda)^{1/2}(\mathbf{U} + i\mathbf{V})\Big)$.

We have glossed over the fact that the embedding process Z *may not even exist.* In this case, X cannot be synthesized exactly through this approach. In general, R_Z may not be positive semi-definite and, hence, we can obtain eigenvalues that are negative, implying negative power at certain frequencies. In this case, the best we can hope to do is to set the negative eigenvalues to zero.

3.2 Algorithm

Putting these ideas together, we present an algorithm to generate two real, length-N, mutually independent Gaussian random vectors $\mathbf{X}_1, \mathbf{X}_2$ with mean μ_X and covariance function $r_X[n]$:

Gaussian LRD FFT Synthesis Algorithm

1. Form the covariance function of the embedding process
$$r_Z[n] = \begin{cases} r_X[n], & 0 < n \leq N-1 \\ r_X(2N-n+1), & N-1 < n \leq 2N-2 \end{cases}.$$
2. Calculate the FFT power spectrum of the embedding process, $\boldsymbol{\lambda} = \mathcal{F}\mathbf{r}_Z$.
3. Generate a length-$2N-2$ realization of complex white Gaussian noise $\mathbf{W} = \mathbf{U} + i\mathbf{V}$, $\mathbf{U},\ \mathbf{V} \sim N(\mathbf{0}, I),\ \mathbb{E}[\mathbf{U}\mathbf{V}^T] = 0$.
4. Set any negative eigenvalues to zero via $\widetilde{\lambda}[k] = \max(0, \lambda[k]),\ k = 0, 1, \ldots, 2N-2$.[6]
5. Compute $\mathbf{Z}$ as the IFFT of the sequence $\left(\frac{\widetilde{\lambda}[k]}{2N-2}\right)^{1/2} W[k],\ 0 \leq k \leq 2N-2$. In matrix form, $\mathbf{Z} = \mathcal{F}(\frac{1}{2N-2}\widetilde{\Lambda})^{1/2}\mathbf{W}$.
6. Set $X_1[n] = \mathrm{Re}(Z[n]) + \mu_X$ and $X_2[n] = \mathrm{Im}(Z[n]) + \mu_X,\ n = 0, 1, \ldots, N-1$.
7. To generate additional realizations, return to 3.

For software that utilizes power-of-two FFT algorithms, an efficient synthesis will occur if $N = 2^k+1,\ k \in \mathbb{N}$, leading to an $M = 2N-2$ that is a power of two. For other lengths, efficient algorithms such as the prime factor algorithm also exist [40]. If any of the eigenvalues $\lambda[k]$ are negative, then the mean-squared deviation of the obtained covariance $\widetilde{R}_Z[n]$ from the desired covariance will be

$$\sum_{n=0}^{2N-3} (R_Z[n] - \widetilde{R}_Z[n])^2 = \frac{1}{2N-2} \sum_{k \text{ s.t. } \lambda[k]<0} \lambda^2[k]. \tag{28}$$

3.3 Sufficient conditions for exactness

For many cases of practical interest, we can guarantee that R_Z is positive semi-definite [14], and hence that the algorithm is exact. We will develop several sufficient conditions based on the properties of the covariance function $r_X[n]$.

[6] This step is unnecessary when the algorithm is exact (see Section 3.3).

Theorem 1 *If the covariance vector* $\mathbf{r}_X = [r_X[0],\ r_X[1],\ \ldots,\ r_X[N-1]]^T$ *satisfies the condition*

$$r_X[0] \geq 2\left|\sum_{n=1}^{N-2} r_X[n]\right| + |r_X[N-1]| \tag{29}$$

then the matrix R_Z *in (27) is positive semi-definite, and the FFT-based synthesis of Section 3.2 is exact.*

Proof: Rewrite (11) with $M = 2N-2$ applied to $\mathbf{r} = \mathbf{r}_Z$ from (26) to obtain

$$S_Z[k] = r_X[0] + (-1)^k r_X[N-1] + 2\sum_{n=1}^{N-2} r_X[n]\ \cos\left(\frac{2\pi nk}{2N-2}\right), k = 0, 1, \ldots, 2N-2. \tag{30}$$

The theorem follows from the triangle inequality and the fact that $|\cos(x)| \leq 1,\ \forall x \in \mathbb{R}$. □

The next theorem will provide the primary theoretical justification behind the algorithm. A sequence $\{r[n],\ n = 0, 1, \ldots, N\}$ is *convex* if its second difference satisfies

$$\Delta^2 r[n] \geq 0, \qquad n = 0, 1, \ldots, N-2. \tag{31}$$

Theorem 2 *If the covariance vector* $\mathbf{r}_X = [r_X[0],\ r_X[1],\ \ldots,\ r_X[N-1]]$ *is convex, decreasing, and nonnegative, then* R_Z *in (27) is positive semi-definite, and the FFT-based synthesis of Section 3.2 is exact.*

Proof: See [14]. Under the above conditions, the FFT series in (30) can be written as a sum of nonnegative Fejér kernels with nonnegative boundary terms. □

We will apply both Theorem 1 and Theorem 2 to demonstrate the validity of our $1/f$ synthesis. It might appear that the conditions in Theorem 2 are overly conservative and can be broadened, since the proof involves writing the FFT spectrum in (30) as the sum of strictly nonnegative terms. In practice we have observed this to be the case — covariance functions deviating slightly from the conditions of Theorem 2 lead to positive-definite R_Z's as well. However, the complicated nature of the Fejér kernels makes it difficult to significantly improve the bound. Fortunately, the conditions of Theorem 2 are broad enough to include many LRD covariance functions of practical interest.

3.4 Proof of exactness for $1/f$ processes

We now establish the exactness of the FFT-based method for synthesizing dfGn for $0 < H < 1$. By cumulatively summing the dfGn, we obtain exact sampled fBm for $0 < H < 1$ as well. Although dfGn for $0 < H \leq 1/2$ is not LRD (hence not of primary interest), we include it for completeness.

Recall the covariance function r_{X_H} of a dfGn process $X_H[n]$ (15). We apply Theorem 1 to show for $0 < H < 1/2$, that (15) satisfies the desired conditions for any $N > 1$. We will apply Theorem 2 to show for $1/2 < H < 1$ that (15) satisfies the desired conditions for $1/2 < H < 1$ and any $N > 1$.

Case I: $0 < H \leq 1/2$

We verify the conditions of Theorem 1 through the following set of inequalities:

$$r_{X_H}[0] = \sigma_X^2 \geq \sigma_X^2 \left(1 - \frac{1}{2}\left(N^{2H} - (N-2)^{2H}\right)\right) \tag{32}$$

$$\geq \sigma_X^2 \left(1 + (N-2)^{2H} - (N-1)^{2H}\right) - \frac{\sigma_X^2}{2}\left((N-2)^{2H} - 2(N-1)^{2H} + N^{2H}\right). \tag{33}$$

By substituting in (15), cancelling common terms, and observing that $H < 1/2$, we obtain

$$2\left|\sum_{n=1}^{N-2} r_{X_H}[n]\right| = 2\left|\sigma_X^2\left(-1 - (N-2)^{2H} + (N-1)^{2H}\right)\right| = \sigma_X^2\left(1 + (N-2)^{2H} - (N-1)^{2H}\right). \tag{34}$$

Substituting (34) into (33) leads to

$$r_{X_H}[0] \geq 2\left|\sum_{n=1}^{N-2} r_{X_H}[n]\right| + |r_{X_H}[N-1]|. \tag{35}$$

Hence, from Theorem 1, the FFT synthesis is exact for $M > 1$ and $0 < H \leq 1/2$.

Case II: $1/2 < H < 1$

We leave the proof that r_{X_H} satisfies the conditions of Theorem 2 to Appendix A. However, it is easily verified for large n. Recall that the covariance function $r_{X_H}[n]$ is one-half the central second difference of the structure function $\sigma_X^2|n|^{2H}$. For large n, the second difference approximates the second derivative, so the claim that the sequence $r_{X_H}[n] = \sigma_X^2|n|^{2H}$ is positive, decreasing, and convex can be easily verified from the derivatives of $\sigma_X^2|t|^{2H}$. For the case $n = 0, 1$, the second difference differs considerably from the second derivative; this case must be treated separately.

4 Exactness for Two Classes of Gaussian LRD Processes

It may be unrealistic to model real-world processes using a strict fGn or fBm process, with an entire covariance structure determined through the choice of the single H parameter. Although real-world processes may exhibit the same LRD as fGn or fBm, the short-term covariances may behave differently, and for certain applications the behavior of these short-term covariances is of great importance. For instance, in our canonical networking example, we expect that both LRD and short-term covariances will play a major role in queuing behavior. We investigate two parametric models that overcome this limitation and prove that the FFT synthesis is exact for these two models.

4.1 Asymptotic discrete fractional Gaussian noise

The first class, proposed by Kaplan and Kuo [25], results from the following covariance

$$r_X[n] = \begin{cases} \sigma_X^2, & n = 0 \\ \frac{\sigma_X^2}{2}\left[(A-1)(1-|p|)|p|^{n-1} + A\left(|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}\right)\right], & |n| > 1 \end{cases} \tag{36}$$

with $A = \frac{2H+p(2-2H)}{2H-p(2-2H)}$. They call this process *asymptotic discrete fractional Gaussian noise* (adfGn) and design a wavelet-based algorithm for estimating its parameters from observed data. Just as for standard dfGn, $1/2 < H < 1$ models the long-range dependence of the process. However, with the additional term corresponding to the parameter p, we have more flexibility in modeling the short-range dependence of the process. Notice that the covariance corresponds to the sum of the covariance of an AR(1) process with the covariance of fGn.

We verify that (36) satisfies the conditions of Theorem 2 by observing that $|p|^{n-1}$ is convex, decreasing, and nonnegative for $n > 0$ and $|p| < 1$ and that the sum of two convex, decreasing, and nonnegative sequences is also convex, decreasing, and nonnegative. Hence, we can apply the FFT method of Section 3 to exactly synthesize adfGn processes modeled via (36).

4.2 Kinked fractional Brownian motion

Other processes of interest exhibit different self-similar scaling over different scaling ranges [15]. To model such behavior, let us consider a process with a structure function of the form:

$$g_B[n] = \begin{cases} \sigma_X^2 |n|^{2H_1}, & 0 \le n \le n_1, \\ \gamma |n|^{2H_2}, & n_2 \le n < \infty, \end{cases}$$

with $1/2 < H_1, H_2 < 1$, $\gamma > 0$, and $n_2 > n_1 > 0$. We call such a process *kinked fBm* (kfBm) and the corresponding increments process X *kinked discrete fraction Gaussian noise* (kdfGn). These are so named, because the variance-time plot exhibits two straight line segments connected with a "kink" (see Figure 1(b)). We focus on $H_1, H_2 > 1/2$, since this corresponds to a process with LRD behavior.[7]

[7] It is easily shown that the special case of $H_1 > 1/2$ (LRD power-law scaling behavior over fine scales) combined with $H_2 = 1/2$ (independence over coarse scales) can be obtained by selecting $r_X[n]$ to be equal to (15) for $0 \le n \le n_1$ and equal to 0 for $n > n_2$. This model could be useful for processes such as those described in [17].

Over short time scales ($n < n_1$) the process behaves like sampled fBm with parameter H_1. Over longer time scales ($n > n_2$) the process behaves like fBm with parameter H_2. A variance-time plot of such a process has a slope of $2H_1 - 2$ for $n < n_1$ and $2H_2 - 2$ for $n > n_2$.

Although we define kfBm and kdfGn in terms of the structure function of kfBm, we will actually synthesize these processes by a generating the kdfGn increments process. Hence, we must derive the kdfGn covariance function $r_X[n]$. Observe that $g_B[n]$ is derived from $r_X[n]$ via the difference equation

$$g_B[n+1] = 2g_B[n] - g_B[n-1] + 2r_X[n], \tag{37}$$

with boundary conditions $g_B[0] = 0$ and $g_B[1] = \sigma_X^2$. We will invert this equation to solve for $r_X[n]$ in terms of $g_B[n]$.

4.2.1 Sampled kfBm without a transition region

Let us construct a kdfGn covariance function to model the kfBm scaling behavior. The simplest example results from $n_2 = n_1 + 1$. That is, we piece the kdfGn covariance function together as follows:

$$r_X[n] = \begin{cases} \frac{\sigma_X^2}{2}\left(|n+1|^{2H_1} - 2|n|^{2H_1} + |n-1|^{2H_1}\right), & 0 \le n \le n_1, \\ \frac{\gamma}{2}\left(|n+1|^{2H_2} - 2|n|^{2H_2} + |n-1|^{2H_2}\right), & n_1 + 1 \le n < \infty. \end{cases}$$

From (37) and (38), the structure function of kfBm is then given by

$$g_B[n] = \begin{cases} \sigma_X^2|n|^{2H_1}, & 0 \le n \le n_1 \\ \gamma|n_1+1|^{2H_2} + \alpha, & n = n+1 \\ \gamma|n_1+2|^{2H_2} + \beta, & n = n+2 \\ \gamma|n_2+2|^{2H_2} + \beta(n-n+1) - \alpha(n-n+2), & n > n+2 \end{cases} \tag{38}$$

with

$$\begin{aligned} \alpha &= 2\sigma_X^2|n_1|^{2H_1} - \sigma_X^2|n_1-1|^{2H_1} - 2\gamma|n_1|^{2H_2} + \gamma|n_1-1|^{2H_2} && (39) \\ \beta &= 3\sigma_X^2|n_1|^{2H_1} - 2\sigma_X^2|n_1-1|^{2H_1} + 2\gamma|n_1-1|^{2H_2} - 3|n_1|^{2H_2}. && (40) \end{aligned}$$

Since the error terms α and β are in general non-zero, with this approach it is impossible to exactly achieve the desired scaling behavior for $n > n_2$. Moreover, from (38), in general the deviation from strict self-similarity grows linearly with time for $n > n_2$. However, if γ is chosen as

$$\gamma = \sigma_X^2\left(\frac{|n_1|^{2H_1} - |n_1-1|^{2H_1}}{|n_1|^{2H_2} - |n_1-1|^{2H_2}}\right) \tag{41}$$

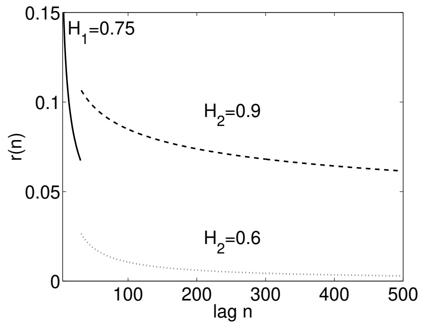


Figure 2: *Covariance function of kdfGn without a transition for parameter values $H_1 = 0.75$, $H_2 = 0.9$, $n_1 = 32$, and $n_2 = 33$ (top), and $H_1 = 0.75$, $H_2 = 0.6$, $n_1 = 32$, and $n_2 = 33$ (bottom). Each covariance function shares the same values for $n \le n_1 = 32$. At $n = n_1$, the kdfGn covariance functions exhibit a discontinuity — a jump up for $H_2 = 0.9$ and a drop for $H_2 = 0.6$. In some cases, this behavior may not correspond to a valid covariance function; however, it can be avoided by using a transition region (see Section 4.2.2).*

then it is easy to show that $\alpha = \beta$ and that

$$g_B[n] = \begin{cases} \sigma_X^2 |n|^{2H_1}, & 0 \le n \le n_1 \\ \gamma |n|^{2H_2} + \alpha, & n > n_1. \end{cases} \tag{42}$$

Thus, for γ as in (41), the kfBm process B will be in error by a fixed constant α and, hence, will exhibit the desired self-similar scaling asymptotically as $n \to \infty$. In practice, since $g_B[n]$ increases at a super-linear rate for $H_2 > 1/2$, the deviation for $n > n_1$ is often negligible.

With the γ of (41), the covariance function for $n > n_1$ is given by

$$r_X[n] = \begin{cases} \frac{\sigma_X^2}{2} \left(|n-1|^{2H_1} - 2|n|^{2H_1} + |n+1|^{2H_1} \right), & 0 \le n \le n_1 \\ \frac{|n_1|^{2H_1} - |n_1 - 1|^{2H_1}}{|n_1|^{2H_2} - |n_1 - 1|^{2H_2}} \frac{\sigma_X^2}{2} \left(|n+1|^{2H_2} - 2|n|^{2H_2} + |n-1|^{2H_2} \right), & n_1 + 1 \le n < \infty. \end{cases}$$

Note that the covariance function jumps substantially (up if $H_1 < H_2$, down if $H_2 > H_1$) in the transition from n_1 to $n_1 + 1$ (see Figure 2).

4.2.2 Sampled kfBm with linear transition

Aside from its artificial discontinuity, the covariance function of the aforementioned kdfGn may not even be valid (positive semi-definite). Even if valid, this covariance will not satisfy Theorem 2 and, hence, we may not be able to exactly synthesize the kdfGn process via our FFT approach. For this reason, we will utilize a *transition region* to ensure that the covariance remains smooth and convex, yet still leads to the desired scaling behavior in the structure function.

In the transition region ($n_1 < n < n_2$), we will choose the covariance function to be an affine

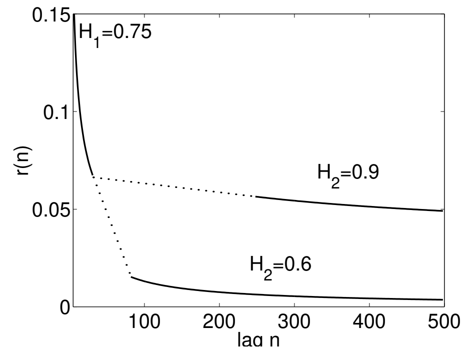


Figure 3: *Covariance function of kdfGn with a linear transition for parameter values $H_1 = 0.75$, $H_2 = 0.9$, $n_1 = 32$, and $n_2 = 248$ (top), and for kdfGn with $H_1 = 0.75$, $H_2 = 0.6$, $n_1 = 32$, and $n_2 = 82$ (bottom). In each case, n_2 is chosen to be the smallest integer for which the covariance function remains convex. For $n \leq n_1$, the covariance function corresponds to dfGn with $H = H_1$. Starting from $n = n_1$, the kdfGn covariance function is linear (dotted line) until n_2 is reached, after which the covariance function corresponds to dfGn with $H = H_2$.*

function $an + b$, leading to the overall expression:

$$r_X[n] = \begin{cases} \frac{\sigma_X^2}{2}\left[|n+1|^{2H_1} + |n-1|^{2H_1} - 2|n|^{2H_1}\right], & 0 \leq |n| \leq n_1 \\ an + b, & n_1 < |n| < n_2 \\ \frac{\gamma}{2}\left[|n+1|^{2H_2} + |n-1|^{2H_2} - 2|n|^{2H_2}\right] & n_2 \leq |n| \leq \infty. \end{cases} \tag{43}$$

We use a linear transition for two reasons. First, it is analytically tractable and always leads to a valid solution. Second, from the following heuristic argument we can demonstrate that a linear function provides an efficient (short) transition.

For $H_1 > H_2$, the value of the covariance function entering the transition is too large and must be reduced. Thus, in the transition we wish to decrease the covariance function as *quickly* as possible, while maintaining the convexity and monotonicity of the overall covariance function. In this case, a linear transition function with slope $a = \Delta r_X[n_1 - 1]$ leads to the shortest transition for a fixed decrease in the covariance function (see Figure 3).

For $H_2 > H_1$, the value of the covariance function entering the transition is too small. Hence, in the transition we wish the covariance function to decay as *slowly* as possible, yet still maintain convexity and monotonicity. In this case, a linear transition function with slope $a = \Delta r_X[n_2]$ leads to the smallest decrease in the covariance function for a given transition length (see Figure 3).

Having settled on a linear transition, we will take the following approach. Given the values σ_X^2, H_1, H_2, n_1, and n_2, we will solve for a, b, and γ to approximate the desired scaling behavior (37). As in (42), we will have to tolerate a slight error term in the structure function for $n > n_2$.

For n_2 large enough, we will demonstrate that the covariance function remains convex, decreasing, positive, and, if viewed as a continuous-time function, smooth.

For the covariance function of (43), the structure function can be found through (37) to be

$$g_B[n] = \begin{cases} \sigma_X^2|n|^{2H_1}, & 0 \le n \le n_1 \\ 2\sigma_X^2 {n_1}^{2H_1} - \sigma_X^2(n_1-1)^{2H_1} + 2(an_1+b), & n = n_1+1 \\ h[n], & n_1+1 < n < n_2 \\ \gamma n_2^{2H_2} + \alpha, & n = n_2 \\ \gamma(n_2+1)^{2H_2} + \beta, & n = n_2+1 \\ \gamma(n+1)^{2H_2} + \beta(n-n_2) - \alpha(n-n_2-1), & n > n_2+1, \end{cases} \tag{44}$$

with

$$\begin{aligned} h[n] = \frac{a}{3}n^3 + bn^2 + \left(a(-n_1^2+n_1-\frac{1}{3}) + b(-2n_1+1) + \sigma_X^2 n_1^{2H_1} - \sigma_X^2(n_1-1)^{2H_1}\right) n \\ +a\left(\frac{2n_1^3}{3} - n_1^2 + \frac{n_1}{3}\right) + b\left(n_1^2 - n_1\right) + \sigma_X^2\left((-n_1+1)n_1^{2H_1} + n_1(n_1-1)^{2H_1}\right), \end{aligned} \tag{45}$$

$$\alpha = -\gamma(n_2)^{2H_2} + h(n_2), \tag{46}$$

and

$$\beta = 2h(n_2) - h(n_2-1) + \gamma(n_2-1)^{2H_2} - 2\gamma(n_2)^{2H_2}. \tag{47}$$

As in (42), we set $\alpha = \beta$ so that the structure function (44) will be in error by only a fixed constant for $n \ge n_2$. This leads to the equation

$$\gamma\left(n_2^{2H_2} - (n_2-1)^{2H_2}\right) = h(n_2) - h(n_2-1). \tag{48}$$

Although our covariance function is in discrete time, we will choose b and γ to ensure the "continuity" of the covariance function at n_1 and n_2, respectively via

$$b = r_X(n_1) - an_1 \tag{49}$$

$$\gamma = 2\frac{an_2+b}{(n_2-1)^{2H_2} - 2n_2^{2H_2} + (n_2+1)^{2H_2}}. \tag{50}$$

Substituting h into (48) and then plugging in (49) for b and (50) for γ leads to:

$$a = \frac{r_X(n_1)\left(2n_2 - \eta - 2n_1\right) + \sigma_X^2 n_1^{2H_1} - \sigma_X^2(n_1-1)^{2H_1}}{(n_2-n_1)(\eta+1) - (n_2-n_1)^2}, \tag{51}$$

with

$$\eta \;=\; 2\frac{n_2^{2H_2}-(n_2-1)^{2H_2}}{(n_2-1)^{2H_2}-2n_2^{2H_2}+(n_2+1)^{2H_2}}. \tag{52}$$

Having solved for the desired parameters, we will now verify that the resulting covariance function is in accordance with Theorem 2, which ensures that FFT-based synthesis of such a process is exact. For a fixed n_2, a sufficient condition for the overall covariance function to be convex, decreasing, and positive is that $\gamma > 0$ and

$$\Delta r_X[n_1-1] < a < \Delta r_X[n_2] < 0. \tag{53}$$

We will demonstrate that such an n_2 exists.

We will find it useful to approximate differences of a discrete function with the derivatives of the identical function viewed as a continuous function of time, for instance, $\Delta r_X[n_2] \approx r_X'(n_2)$. As $n_2 \to \infty$, these approximations become exact.

We start by substituting derivatives for the first and second differences in the η term of (51) and finding the limit of (51) as $n_2 \to \infty$. This leads to the result

$$\lim_{n_2\to\infty} a = 2\frac{2H_2-2}{3-2H_2}\frac{r_X[n_1]}{n_2} = 0, \tag{54}$$

Since $a \to 0$ and $\Delta r_X[n_1-1] < 0$, for n_2 large enough we find that $a > \Delta r_X[n_1-1]$.

Substituting (49) and (54) into (50), we observe that

$$\lim_{n_2\to\infty} \gamma = \frac{2H_2-1}{3-2H_2}\frac{r_X[n_1]}{(n_2-1)^{2H_2}-2n_2^{2H_2}+(n_2+1)^{2H_2}}. \tag{55}$$

Hence for n_2 large enough, $\gamma > 0$.

Finally, by substituting (55) into (43), approximating the second difference via the second derivative, and taking a derivative with respect to n, we obtain

$$\lim_{n_2\to\infty} \Delta r_X(n_2) = \lim_{n_2\to\infty} r_X'(n_2) = \frac{(2H_2-1)(2H_2-2)r_X[n_1]}{(3-2H_2)n_2} \cong \frac{2H_2-1}{2}a. \tag{56}$$

Since $a < 0$ and $\frac{2H_2-1}{2} < 1$, from (56) we see that for n_2 large enough $a < \Delta r_x[n_2]$. Hence, the conditions $\gamma > 0$ and (53) are satisfied, guaranteeing that an n_2 exists for which the covariance function is convex, decreasing, positive, and consistent with the desired scaling behavior of (37). Thus, to synthesize the desired kdfGn process exactly, we simply apply the FFT algorithm of Section 3.2 to the covariance function (43).

4.3 Gaussian LRD synthesis example

We can now provide more details on the synthesis experiment of Figure 1. Using the FFT method of Section 3.2, we synthesized length-$2^{17}+1$ sampled fBms for $H = 0.9$, $\mu_X = 0$, $\sigma_X^2 = 1$ and $H = 0.6$, $\mu_X = 0$, $\sigma_X^2 = 38.63$ and a sampled kfBm (37) with $H_1 = 0.9$, $H_2 = 0.6$, $\sigma_X^2 = 1$, $\mu_X = 0$, $n_1 = 64$, $n_2 = 340$, and $\gamma = 38.63$. The value $n_2 = 340$ is the minimum n_2 for which the covariance function remains convex, positive, and decreasing using a linear transition (as verified through (53)). As expected, over short time scales, the kfBm behaves like the fBm with $H = 0.9$ and $\sigma_X^2 = 1$. Over larger time scales, the kfBm behaves like the fBm with $H = 0.6$ and $\sigma_X^2 = \gamma = 38.63$.

At bottom in Figure 1 are the variance-time plots for the fBm and kfBm traces. Included within these plots are empirical estimates of H corresponding to linear fits of the log-log plots over the scales indicated. These H estimates agree quite well with the desired theoretical values, as they are within the standard error range of the variance-time plot [27]. To decrease the bias in these plots and H estimators, the variances in the variance-time plot were calculated using the knowledge that the processes were zero-mean. For H near one, reliable mean estimation is difficult and can lead to inaccuracy in the variance calculations, particularly at coarser resolutions.

5 FFT-based Synthesis of NonGaussian Processes

In many applications, data networking and the stock market to name two, the LRD processes of interest are strictly positive and can exhibit heavy-tailed, nonGaussian behavior. To model this behavior we consider the following task. We seek to synthesize a stationary, nonGaussian random process $Y[n]$ with distribution function F_Y and covariance function r_Y. As shown in Figure 4, our approach is to form Y as a memoryless nonlinear function T of a Gaussian process $X[n]$ with covariance function r_X. The question is whether (and how) we can choose T and r_X to obtain the desired F_Y and r_Y.

We will prove that through this approach we can synthesize *exactly second-order self-similar* (ESS) nonGaussian processes without error. Recall from Section 2.3.3 that a discrete-time process $Y[n]$ is ESS if it has a variance-time plot (hence covariance and power spectrum) identical to that of fGn. Thus ESS processes provide nonGaussian equivalents to the $1/f$ fGn process.[8]

[8]Note that we slightly prefer the use of ESS over nonGaussian $1/f$ noise, since as noted in Section 2, discrete-time realizations of ESS or fGn processes will not be strictly $1/f$.

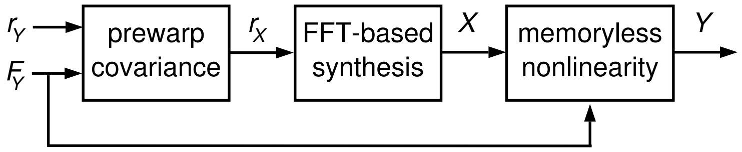


Figure 4: *Block diagram for the synthesis procedure for nonGaussian LRD processes. From the desired first-order distribution function F_Y and covariance function r_Y of the nonGaussian process Y, the prewarped correlation function r_X is calculated. The process X is synthesized efficiently via the FFT algorithm of Section 3, and passed through a nonlinearity to obtain the desired process $Y = T(X)$.*

In this formulation, we do not specify the higher-order covariances of the process. In addition, our approach applies only to finite-variance processes that have a well-defined covariance function, which would exclude, for example, the stable distributions. For a fast, but approximate, synthesis of positive-valued LRD data based on wavelets, see [35].

5.1 Creating a nonGaussian process from one Gaussian process

By definition, the random variable $\Phi(X[n])$ is distributed uniformly on $[0, 1]$, and we can apply the inversion principle [18, 19, 41] to synthesize a random process $Y[n]$ with distribution F_Y. Let the value $X[n]$ have a Gaussian first-order marginal distribution $N(\mu_X, \sigma_X^2)$, and let $F_Y(y)$ be a continuous distribution function on $\mathbb{R}$ with inverse F_Y^{-1} defined by [41]

$$F_Y^{-1}(u) = \inf\{y : F(y) = u,\ 0 < u < 1\}. \tag{57}$$

Then we can synthesize the nonGaussian process $Y[n]$ via

$$Y[n] = T(X[n]) = F_Y^{-1}\left(\Phi\left(\frac{X[n]-\mu_X}{\sigma_X}\right)\right). \tag{58}$$

If X has covariance function $r_X[n]$, then the covariance function for Y can be calculated as follows. Let

$$\rho_X[n] \equiv \frac{\mathbb{E}[X[m]X[m-n]] - \mu_X^2}{\sigma_X^2} = \frac{r_X[n]}{\sigma_X^2} \tag{59}$$

be the *normalized* covariance function for X. Then $r_Y[n]$ may be found via

$$r_Y[n] = \mathbb{E}[T(X[m])\,T(X[m-n])] - \mathbb{E}[T(X[n])]^2 \tag{60}$$

with

$$\mathbb{E}[T(X[m])\,T(X[m-n])] = \int_{\infty}^{\infty} \frac{T(x_1)T(x_2)}{2\pi\sigma_X^2\sqrt{1-\rho_X^2[n]}} \Upsilon(x_1, x_2, \mu_X, \sigma_x^2, \rho_X^2[n])dx_1dx_2, \tag{61}$$

$$\Upsilon(x_1, x_2, \mu_X, \sigma_x^2, \rho_X^2) \equiv \exp\left(\frac{-(x_1-\mu_X)^2 - 2\rho_X(x_1-\mu_X)(x_2-\mu_X) + (x_2-\mu_X)^2}{2\sigma_X^2(1-\rho_X^2)} \right). \tag{62}$$

Calculating (60) leads to the transformation $r_Y[n] = W(r_X[n])$, which we invert to obtain the desired "prewarped" covariance function for X via[9]

$$r_X[n] = W^{-1}(r_Y[n]). \tag{63}$$

5.2 Creating a nonGaussian process from two Gaussian processes

In the synthesis of nonGaussian processes, the distribution mapping Φ used in (58) can lead to difficulties in the calculation of T and W^{-1}. Closed-form analysis may be intractable, and hence numerical integration may be required to obtain these quantities. Here we investigate a second approach that is more amenable to closed-form analysis.

For a fixed value of n, if $X_1[n],\ X_2[n] \sim N(0,1)$ and are independent, then the random process

$$Y[n] = \exp\left(-\frac{1}{2}\left(X_1^2[n] + X_2^2[n]\right)\right) \tag{64}$$

has a first order distribution that is uniform on $(0,1)$ [42]. This suggests an alternative synthesis approach for synthesizing a random process Y with first order distribution F_Y. We set

$$Y[n] = F_Y^{-1}\left(\exp\left(-\frac{1}{2}\left(X_1^2[n] + X_2^2[n]\right)\right)\right), \tag{65}$$

with X_1 and X_2 two independent zero-mean Gaussian random processes with common covariance function $r_X[n]$.

As before, to synthesize Y with a desired covariance structure, we must calculate the warping

$$W(r_X[n]) \equiv \mathbb{E}\left[T\left(X_1[m], X_2[m]\right)\, T\left(X_1[m-n], X_2[m-n]\right)\right] - \mathbb{E}^2\left[T\left(X_1[m], X_2[m]\right)\right] \tag{66}$$

and its inverse $r_X[n] = W^{-1}(r_Y[n])$. This leads to an integral similar to (61).

Note that since Y is a function of the square of Gaussian processes, this approach is only valid for synthesizing positively-correlated processes, that is, processes with $r_Y[n] \geq 0$.

[9] Aside from the covariance function r_Y, r_X may also depend on the parameters of the marginal density of Y (its mean for example). For clarity of presentation, we suppress the potential dependency of r_X on these auxiliary parameters, since they can be absorbed into W^{-1}.

5.3 Algorithm for synthesis of nonGaussian LRD processes

Putting all of these steps together, we obtain the algorithm for synthesizing nonGaussian LRD processes:

NonGaussian LRD FFT Synthesis Algorithm

1. Prewarp the desired LRD covariance function r_Y via $r_X = W^{-1}(r_Y)$, with W obtained from (60) if $\mathbf{Y}$ is synthesized from one process $\mathbf{X}_1$ or (66) if $\mathbf{Y}$ is synthesized from two processes $\mathbf{X}_1$ and $\mathbf{X}_2$. If necessary, calculate μ_X.
2. Using the FFT-based synthesis algorithm of Section 3, synthesize Gaussian random vectors $\mathbf{X}_1, \mathbf{X}_2$ with mean μ_X and covariance function r_X.
3. Transform the data via $\mathbf{Y} = T(\mathbf{X}_1)$ from (58) or $Y = T(\mathbf{X}_1, \mathbf{X}_2)$ from (65).
4. To generate additional realizations, return to 2.

As in the Gaussian case, there is no guarantee that we can exactly synthesize the target non-Gaussian covariance function r_Y. Indeed, the prewarped Gaussian process may not even exist! In some cases the covariance function r_Y would be valid if Y were Gaussian, but when prewarped back to $r_X = W^{-1}(r_Y)$ leads to an invalid covariance function for X [19].

Fortunately, for ESS processes with the correlation structure of dfGn, for $1/2 < H < 1$ we can use Theorem 2 demonstrate that in many cases the synthesis is exact for the underlying Gaussian covariance function r_X. We have already established that if Y has a covariance function corresponding to (15) then r_Y is convex, positive, and decreasing. Prewarping typically preserves monotonicity and positivity, but can destroy the convexity of the covariance function. In Appendix B, we provide general sufficient conditions for the nonGaussian ESS synthesis to be exact and demonstrate that ESS synthesis is exact for the lognormal, uniform, exponential and Pareto densities, densities that we now subject to closer scrutiny.

5.4 Example densities

5.4.1 Lognormal

If $X[n]$ has a first-order pdf $N(\mu_X, \sigma_X^2)$, then $Y[n] = \exp(X[n])$ has a first-order lognormal pdf [42]

$$f_Y(y) = \sigma y \sqrt{2\pi} \exp\left[-\frac{1}{2}\frac{(\log y - \mu_X)^2}{\sigma_X^2}\right], \ y > 0 \qquad (67)$$

with mean and variance

$$\begin{aligned}\mu_Y &= \exp\left(\mu_X + \frac{\sigma_X^2}{2}\right), &(68)\\ \sigma_Y^2 &= \exp\left(2\mu_X + 2\sigma_X^2\right) - \exp\left(2\mu_X + \sigma_X^2\right). &(69)\end{aligned}$$

In addition, if $X[n]$ has covariance $r_X[n]$ then from (61) it can be shown that $Y[n]$ has covariance

$$r_Y[n] = \sigma_Y^2 \frac{\exp\left(r_X[n]\right) - 1}{\exp\left(\sigma_X^2\right) - 1}. \tag{70}$$

The transformations (68–70) are inverted by

$$\begin{aligned}\mu_X &= \log\left(\frac{\mu_Y^2}{\sqrt{\mu_Y^2 + \sigma_Y^2}}\right), &(71)\\ \sigma_X^2 &= \log\left(1 + \frac{\sigma_Y^2}{\mu_Y^2}\right), &(72)\\ r_X[n] &= \log\left(1 + \frac{r_Y[n]}{\mu_Y^2}\right). &(73)\end{aligned}$$

5.4.2 Uniform

Although a random process whose values are uniformly-distributed on $[0, 1]$ can be obtained simply through the transformation $Y = \Phi(X)$, as noted earlier the transformation (65) provides a construction more amenable to closed-form analysis.

By completing the square and using the fact the Gaussian pdf integrates to one, it can be shown that

$$\mathbb{E}\left[\exp\left(-\frac{1}{2}\left(X[m]^2 + X[m-n]^2\right)\right)\right] = \frac{1}{\sqrt{4 - r_X^2[n]}}, \tag{74}$$

and hence

$$r_Y[n] = \frac{r_X^2[n]}{4(4 - r_X^2[n])}. \tag{75}$$

Inverting this, we obtain

$$r_X[n] = 4\sqrt{\frac{r_Y[n]}{1 + 4r_Y[n]}}. \tag{76}$$

To generate a uniform variable on the interval (a, b), we simply form

$$Y[n] = a + (b - a)\exp\left(-\frac{1}{2}\left(X_1^2[n] + X_2^2[n]\right)\right) \tag{77}$$

and obtain $r_X[n]$ as

$$r_X[n] = 4\sqrt{\frac{r_Y[n]}{(b-a)^2 + 4r_Y[n]}}. \tag{78}$$

5.4.3 Exponential

An exponential random variable Y with $f_Y(y) = \frac{1}{\mu_Y}e^{-y/\mu_Y},\ y > 0$ corresponds to $F_Y^{-1}(y) = -\mu_Y \log y$ [41]. Hence, to synthesize an exponential random process with mean μ_Y we form

$$Y[n] = \mu_Y\left(\frac{X_1^2[n] + X_2^2[n]}{2}\right) \tag{79}$$

We observe that

$$\mathbb{E}[Y[n]Y[m-n])] = \frac{\mu_Y^2}{4}\left(2 + 2\mathbb{E}\left[X^2[m]X^2[m-n]\right]\right) \tag{80}$$

and thus

$$r_Y[n] = \mu_Y^2 r_X^2[n], \tag{81}$$

which implies that

$$r_X[n] = \frac{1}{\mu_Y}\lambda\sqrt{r_Y[n]}. \tag{82}$$

In passing, we note that a Weibull random variable with parameters $(\mu,\ a)$ may be formed from the a-th root of an exponential random variable [41, 43]. Hence, a Weibull random process may be obtained via

$$Y[n] = \left[\mu\left(\frac{X_1^2[n] + X_2^2[n]}{2}\right)\right]^{1/a}. \tag{83}$$

An expression for the covariance warping in terms of the hypergeometric function is provided in [43]. However, a derivation for the inversion of this expression appears analytically intractable.

5.4.4 Pareto

A Pareto(a, b) random variable with $f_Y(y) = \frac{ab^a}{y^{a+1}},\ x \geq b > 0$ and

$$\mathbb{E}[Y] = \frac{ab}{a-1},\ (a > 1), \tag{84}$$

$$\mathbb{E}[Y^2] = \frac{ab^2}{a-2},\ (a > 2), \tag{85}$$

$$\sigma_Y^2 = \frac{ab^2}{(a-1)^2(a-2)},\ (a > 2), \tag{86}$$

has $F^{-1}(y) = \frac{b}{y^{1/a}}$. No moments of order $q \geq a$ exist.

To synthesize a Pareto random process, we apply [41]

$$Y[n] = b \exp\left(\frac{X_1^2[n] + X_2^2[n]}{2a}\right). \tag{87}$$

For $a > 2$, we find that

$$\mathbb{E}[Y[m]Y[m-n]] = \frac{a^2 b^2}{a^2 - 2a + (1 - r_X[n])}, \tag{88}$$

which leads to the following expression for $r_Y[n]$ in terms of $r_X[n]$

$$r_Y[n] = \frac{a^2 b^2 r_X[n]}{(a-1)^4 - r_X^2[n](a-1)^2}. \tag{89}$$

To find $r_X[n]$ in terms of $r_Y[n]$, we invert as follows

$$r_X[n] = \frac{(a-1)^2 \sqrt{r_Y[n]}}{\sqrt{a^2 b^2 + r_Y[n](a-1)^2}}. \tag{90}$$

By writing $r_X[n]$ in terms of $\rho_Y[n] = r_Y[n]/\sigma_Y^2$, we obtain the simpler expression

$$r_X[n] = \frac{(a-1)\sqrt{\rho_Y[n]}}{\sqrt{a^2 - 2a + \rho_Y[n]}}. \tag{91}$$

Care must be taken in the use of such a process for $2 < a \leq 4$, since the fourth moment does not exist. Hence, the process may not be covariance ergodic [44] and may vary significantly from realization to realization.

5.5 NonGaussian LRD synthesis examples

In Figure 5 we provide synthesis examples of Gaussian, uniform, exponential, lognormal, and Pareto ESS traces of length-65537, each synthesized using the FFT method of Section 5.3. The parameters of each process are chosen so that $H = 0.85$, $\mu_Y = 1$, and $\sigma_Y^2 = 1$. Although all of these processes have the same theoretical second-order statistics, they have different higher-order moments and are quite distinct visually. Hence, we could expect them to behave quite differently in data network simulations, for example. Moreover, we note that although the averaged auto-covariance function is indistinguishable from the true covariance function, in any one realization the covariance can vary significantly from the true auto-covariance. This deviation is more pronounced in LRD data with H near one, and is exacerbated by skewed distributions with heavy tails, such as the Pareto.[10] In

[10]Furthermore, the averaged auto-covariances plotted in Figure 5 were obtained assuming a known mean. With LRD data, for any one realization the empirical mean can vary substantially, and covariance estimates can behave even more erratically.

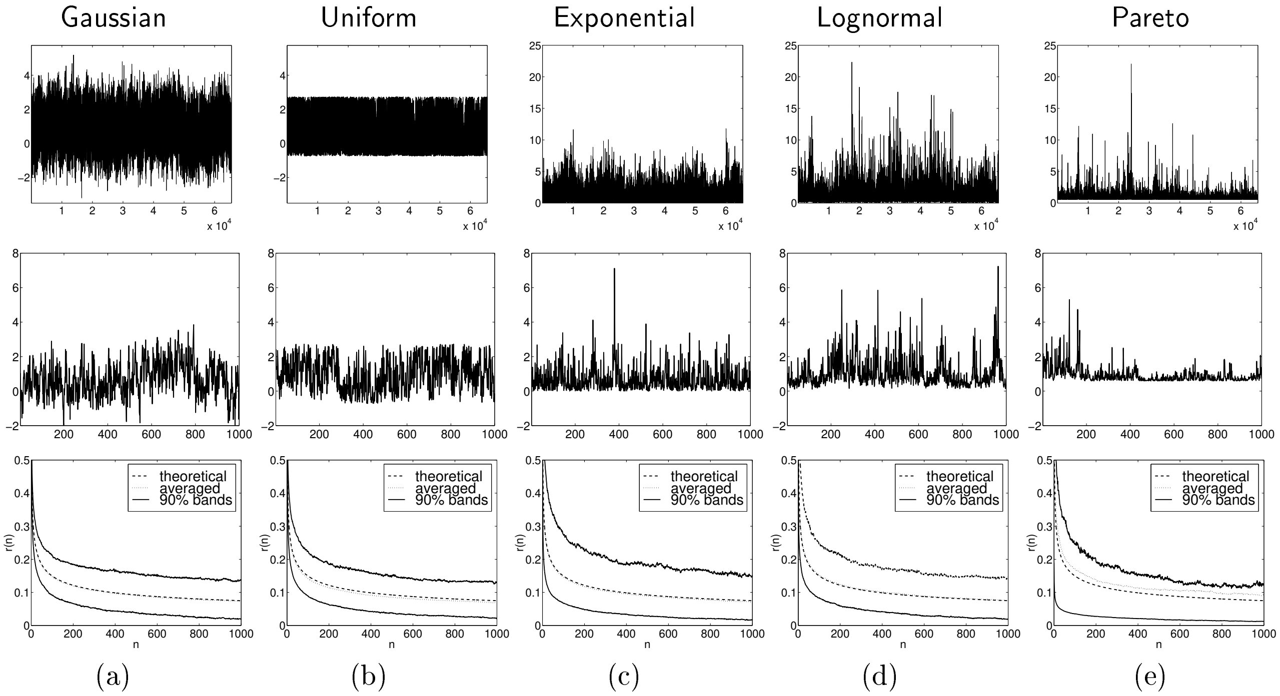
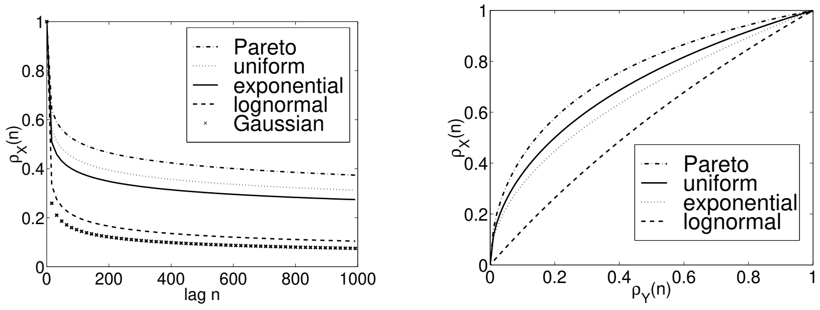


Figure 5: *Realizations of the exactly second-order self-similar (ESS) process $Y[n]$ for (a) Gaussian, (b) uniform, (c) exponential, (d) lognormal, and (e) Pareto data synthesized via the FFT method of Section 5.3. The parameters of each process were chosen so that $\mu_Y = 1$, $\sigma_Y^2=1$, and $H = 0.85$. At top are the full length-65537 traces. At middle are the first 1000 samples. The processes have the same theoretical second-order statistics, but appear strikingly different. At bottom are the averaged auto-covariance functions with 90% confidence intervals computed over 500 trials. Except for the Pareto process, the averaged auto-covariance function is nearly indistinguishable from the true covariance function. In any one realization, however, the covariance can vary significantly from the true auto-covariance, particularly for H near one and for heavier-tailed densities such as the Pareto.*

these instances, great care must be taken in interpreting and estimating parameters from a single LRD trace.

In Figure 6, we plot the prewarped Gaussian covariance functions required to synthesize the traces in Figure 5 and the prewarping functions themselves. These plots demonstrate that the underlying Gaussian process exhibits stronger LRD than the target nonGaussian process.

Although the focus of this work is not on the measurement and estimation of LRD in Gaussian and nonGaussian data, this example illustrates how the tools developed here might be useful for such a study. Our experiments indicate that for highly-skewed nonGaussian data with H close to one, on any given realization the sample covariance can deviate significantly from the theoretical covariance. In this case, accurate parameter estimation from a single trace is difficult. Our FFT algorithm could be used to test the robustness of different LRD parameter estimation algorithms,



(a) (b)

Figure 6: *(a) Gaussian covariance functions (normalized) required to synthesize the traces in Figure 5. In all cases, the underlying Gaussian process X is more highly correlated than the target nonGaussian process Y. (b) Normalized Gaussian covariance values $\rho_X[n] = r_X[n]/\sigma_X^2$ as a function of the normalized nonGaussian covariance values $\rho_Y[n] = r_Y[n]/\sigma_Y^2$.*

particularly those derived under a Gaussian assumption [27].

6 Conclusions

Although continuous-time $1/f$ processes such as fBm and fGn have a number of fascinating properties, at the end of the day a digital computer-based synthesis of these signals must be performed in discrete time. We have demonstrated that an exact synthesis should be based on the *time-domain* properties of these processes. Moreover, we have proved that an FFT approach based on embedding a Toeplitz matrix into a circulant matrix is exact and, because of its computational superiority ($O(N \log N)$ versus $O(N^2)$), clearly the best method for exactly synthesizing sampled fBm and dfGn.

The simplicity of the time-domain approach combined with the fast FFT algorithm has allowed us to develop a general framework for modeling and exactly synthesizing Gaussian and nonGaussian LRD processes. For instance, we have introduced a new Gaussian process called kinked fractional Brownian motion, with increments process kinked discrete fractional Gaussian noise. KfBm exhibits different fBm-like scaling behavior over different resolutions of analysis and for several applications may be a more accurate model than pure fBm. In the nonGaussian arena, we have demonstrated efficient synthesis techniques for exactly second-order self-similar processes — processes with a covariance structure identical to dfGn, but with a nonGaussian first-order density. These types of processes occurs in areas as diverse as data networking and finance and can be synthesized exactly for many nonGaussian pdfs of interest.

Several issues beg to be addressed. Synthesis of general nonGaussian LRD processes, such as nonGaussian equivalents to kdfGn and adfGn of Section 4, via the FFT method would be straightforward from the work developed here. In general it could be difficult to guarantee that this synthesis procedure is exact, but we have noticed that if the Gaussian covariance function satisfies Theorem 2, then the covariance will often be valid for FFT-based nonGaussian synthesis as well.

Synthesis of multi-dimensional LRD processes is another application of clear interest, particularly for image processing. Unfortunately, as shown in theory [14] and practice [4], many of the useful 1-d conditions for guaranteeing exactness do not hold in higher dimensions. Moreover, for higher-dimensional spaces, there does not exist a single fGn-like increments process for fBm [4]. Generalizations of this work to multi-dimensional processes could be quite difficult.

Lastly, a general polynomial warping [18, 19], in particular a Hermite polynomial expansion, would allow us to synthesize LRD processes with a wider class of nonGaussian pdfs than the examples listed here. However, in general such a method requires the use of numerical integration to calculate the Hermite polynomial coefficients.

With algorithms to exactly synthesize both Gaussian and nonGaussian LRD processes, we provide network researchers with tools that will help them understand and differentiate the effects of LRD and nonGaussianity in their data. Since the FTT synthesis is exact, we can be confident that the properties of the synthesized traces are indeed authentic and not due to some artifact of the synthesis procedure. In addition to its potential applicability to networking, this research could also be useful for applications such as simulation of financial markets, where volatility is known to exhibit LRD and nonGaussian behavior [8, 45].

We are not limited to these few applications, however. The FFT algorithm could also be quite useful for benchmark testing of LRD parameter estimation algorithms. For instance, LRD processes with varying short-term covariances and nonGaussian marginals could be synthesized to test the accuracy and robustness of LRD estimation algorithms. More generally, our framework provides synthesis and modeling tools that could lead to new insights into the general behavior of Gaussian and nonGausssian $1/f$ and LRD processes, which are still poorly understood relative to classical

Markov, Poisson, and AR processes.

A Exactness of FFT-Based Synthesis of DfGn for $1/2 < H < 1$

To verify that the FFT synthesizes dfGn exactly, we must demonstrate the fGn covariance function $r_{X_H}[n]$ (15) satisfies the conditions of Theorem 2. That is, we must show that it is convex, decreasing, and non-negative. Hence, we need to show for any integer $n > 0$: $r_{X_H}[n] \geq 0$, $\Delta r_{X_H}[n] \leq 0$, and $\Delta^2 r_{X_H}[n] \geq 0$. We will split this into two cases.

A.1 Case I: $n \geq 1$

It will be convenient to replace the discrete-time function $r_{X_H}[n]$ with its continuous-time equivalent $r_{X_H}(\tau)$, since convexity, monotonicity, and non-negativity for the continuous function $r_{X_H}(\tau), \tau \in [n_1, n_2]$, $n_1, n_2 \in \mathbb{Z}$ immediately lead to the identical properties in the sequence $r_{X_H}[n]$, $n = n_1, n_1+1, \ldots, n_2$.

For example, if $\frac{d^2 r_{X_H}(\tau)}{d\tau^2}$ is non-negative on (n_1, n_2), then for $\tau_1 > \tau_2$, we observe that $\frac{dr_{X_H}(\tau_1)}{d\tau} \geq \frac{dr_{X_H}(\tau_2)}{d\tau}$, implying that

$$
\begin{aligned}
\Delta^2 r_{X_H}[n] &= (r_{X_H}[n] - r_{X_H}[n+1]) - (r_{X_H}[n+1] - r_{X_H}[n+2]) \\
&= -\int_n^{n+1} \frac{dr_{X_H}(\tau)}{d\tau} d\tau + \int_{n+1}^{n+2} \frac{dr_{X_H}(\tau)}{d\tau} d\tau \\
&= \int_n^{n+1} \left[\frac{dr_{X_H}(\tau+1)}{d\tau} - \frac{dr_{X_H}(\tau)}{d\tau} \right] d\tau \geq 0. \qquad (92)
\end{aligned}
$$

Hence, it becomes clear that

$$
\frac{d^2 r_{X_H}(\tau)}{d\tau^2} \geq 0, \ \forall \tau \in (n_1, n_2) \ \Rightarrow \ \Delta^2 r_{X_H}[n] \geq 0, \ n_1 \leq n \leq n_2 - 2, \qquad (93)
$$

which from (31) implies that the discrete-time sequence is convex over the interval $[n_1, n_2]$.

To establish the convexity, monotonicity, and non-negativity properties of the covariance function, we will find the following expansion useful for qualitative analysis of $r_{X_H}(\tau)$ for $\tau > 1$:

$$
\begin{aligned}
(\tau \pm 1)^c &= \tau^c + \sum_{k=1}^{\infty} \frac{(\pm 1)^k \tau^{c-k}}{k!} \prod_{j=0}^{k-1} (c-j) \qquad (94) \\
&= \tau^c \pm c\tau^{c-1} + c(c-1)\frac{\tau^{c-2}}{2!} \pm \cdots \qquad (95)
\end{aligned}
$$

By substituting (94) into (15) and combining terms, we obtain the series expansions for r_{X_H} and its first derivatives for $\tau > 1$:

$$r_{X_H}(\tau) = \sigma_X^2 \sum_{k=1}^{\infty} \frac{\tau^{2H-2k}}{(2k)!} \prod_{j=0}^{2k-1} (2H-j), \tag{96}$$

$$r'_{X_H}(\tau) = \sigma_X^2 \sum_{k=1}^{\infty} \frac{\tau^{2H-2k-1}}{(2k)!} \prod_{j=0}^{2k} (2H-j), \tag{97}$$

$$r''_{X_H}(\tau) = \sigma_X^2 \sum_{k=1}^{\infty} \frac{\tau^{2H-2k-2}}{(2k)!} \prod_{j=0}^{2k+1} (2H-j). \tag{98}$$

For $H > \frac{1}{2}$, each term in (96) and (98) is positive (the product of an even number of negative numbers), while each term in (97) is negative (the product of an odd number of negative numbers). For $0 < H < 1/2$ the opposite statements are true. Thus, it is easily verified that

$$r_{X_H}(\tau) > \sigma_X^2 H(2H-1)\tau^{2H-2} > 0, \quad 1/2 < H < 1, \quad \tau > 1, \tag{99}$$

$$r'_{X_H}(\tau) < \sigma_X^2 H(2H-1)(2H-2)\tau^{2H-3} < 0, \quad 1/2 < H < 1, \quad \tau > 1, \tag{100}$$

$$r''_{X_H}(\tau) > \sigma_X^2 H(2H-1)(2H-2)(2H-3)\tau^{2H-4} > 0, \quad 1/2 < H < 1, \quad \tau > 1. \tag{101}$$

We have thus established that the continuous-time function $r_{X_H}(\tau)$ is convex, positive, and decreasing for $\tau > 1$. This fact and the continuity of $r_{X_H}(\tau)$ at $\tau = 1$ imply that the discrete-time function $r_{X_H}[n]$ is convex, non-negative, and decreasing for $n \geq 1$.

A.2 Case II: $n = 0$

A continuous-time analysis in this case is more difficult, since the underlying continuous function $r_{X_H}(\tau)$ is not convex, with a second-derivative undefined for $\tau = 0, 1$. Instead we can look at the first and second differences directly to verify that the function is decreasing with positive second difference. At $n = 0$, we have

$$\Delta r_{X_H}[0] = r[0] - r[1] = 2 - 2^{2H-1} \geq 0, \tag{102}$$

$$\Delta^2 r_{X_H}[0] = \frac{1}{2}\left(7 - 4\ 2^{2H} + 3^{2H}\right) \geq 0. \tag{103}$$

To see the inequality in (103), note that

$$\frac{d(\Delta^2 r_{X_H}[0])}{dH} = \frac{2^{2H}}{2}\left(-4\log 2 + \left(\frac{3}{2}\right)^{2H}\log 3\right) \leq \frac{2^{2H}}{2}\left(-4\log 2 + \frac{9}{4}\log 3\right) \leq 0, \tag{104}$$

with $\Delta^2 r_{X_H}[0] = 0$ for $H = 1$. Hence, $\Delta^2 r_{X_H}[0]$ must be positive for $H \in (1/2, 1)$. Thus, we conclude that FFT-based synthesis of dfGn is exact for $H \in (1/2, 1)$. $\square$

B Exactness of FFT-Based Synthesis of NonGaussian ESS Processes for $1/2 < H < 1$

We will verify that the prewarped Gaussian covariance used for synthesizing a nonGaussian ESS processes is convex, monotonically decreasing, and positive for several types of nonGaussian pdfs. Just as in Appendix A, we split the test for convexity conditions into two cases: $n \geq 2$ and $n = 0, 1$. Note that we investigate the case $n \geq 2$ instead of $n \geq 1$, since the continuous-time bounds we will derive are not meaningful for $n = 1$.

B.1 Case I: $n \geq 2$

B.1.1 General bound

Consider data with a covariance function $r_X = W^{-1}(r_Y)$, where W^{-1} is a twice-differentiable, monotonically increasing map such that $W^{-1}[1] = \sigma_X^2$. As in Appendix A, we will treat r_X as a function of a continuous-time lag $\tau \in \mathbb{R}$, using the fact that if $r_X(\tau)$ is convex for $\tau \geq 2$, then $\Delta^2 r_X[n] > 0$ for $n \geq 2$.

By the chain and product rules of the calculus we have

$$\frac{d^2 r_X(\tau)}{d\tau^2} = \frac{d^2 W^{-1}(r_Y)}{dr_Y^2}(\tau)\left(\frac{dr_Y(\tau)}{d\tau}\right)^2 + \frac{dW^{-1}(r_Y)}{dr_Y}(\tau)\frac{d^2 r_Y(\tau)}{d\tau^2}. \tag{105}$$

We will now examine the properties of the inverse warping function W^{-1} that are necessary to satisfy that the convexity constraint $\frac{d^2 r_X(\tau)}{d\tau^2} \geq 0$. From (105) we obtain

$$\frac{\frac{d^2 W^{-1}(r_Y)}{dr_Y^2}(\tau)}{\frac{dW^{-1}(r_Y)}{dr_Y}(\tau)} \geq \frac{-\frac{d^2 r_Y(\tau)}{d\tau^2}}{\left(\frac{dr_Y(\tau)}{d\tau}\right)^2}. \tag{106}$$

This bound holds for general LRD processes, not just for ESS processes.

B.1.2 Application to ESS processes

In the case of an ESS process Y, from (101) we observe that for $\tau > 1$

$$\frac{d^2 r_Y(\tau)}{d\tau^2} \geq \sigma_Y^2 H(2H-1)(2H-2)(2H-3)\tau^{2H-4}. \tag{107}$$

Next, we find from (97) that for $\tau > 1$

$$\frac{dr_Y(\tau)}{d\tau} = \sigma_Y^2 H(2H-1)(2H-2)\tau^{2H-3}\left(1 + \sum_{k=2}^{\infty} \tau^{2k-2} \frac{2\prod_{j=2}^{2k-1}(2H-1-j)}{(2k!)}\right). \tag{108}$$

It is easily verified that $\frac{2\prod_{j=2}^{2k-1}|(2H-1-j)|}{(2k!)} < \frac{1}{2}$ for $k \geq 2$. Hence, for $\tau > 1$ we obtain

$$\left|1+\sum_{k=2}^{\infty} t^{2k-2}\frac{2\prod_{j=2}^{2k-1}(2H-1-j)}{(2k!)}\right| \leq 1+\frac{1}{2}\sum_{k=2}^{\infty}\tau^{2k-2} \leq \left(1+\frac{1}{2}\frac{\tau^{-2}}{1-\tau^{-2}}\right). \tag{109}$$

Thus, for $\tau \geq 2$ from (108) and (109) we obtain the bound

$$\left|\frac{dr_Y(\tau)}{d\tau}\right|^2 \leq \frac{49}{36}\sigma^4(H(2H-1)(2H-2))^2\tau^{4H-6}, \tag{110}$$

yielding

$$\frac{\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau)}{\frac{dW^{-1}(r_Y)}{dr_Y}(\tau)} \geq -\frac{36}{49}\frac{(2H-3)}{\sigma_Y^2 H(2H-1)(2H-2)\tau^{2H-2}}. \tag{111}$$

Since on the interval $\frac{2H-3}{2H-2} \geq 2$ for $H \in [1/2, 1]$, we can write

$$\frac{\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau)}{\frac{dW^{-1}(r_Y)}{dr_Y}(\tau)} \geq -\frac{72}{49}\frac{1}{\sigma_Y^2 H(2H-1)\tau^{2H-2}} \geq -\frac{72}{49}\frac{1}{r_Y(\tau)}. \tag{112}$$

Equation (112) is a simple, but somewhat overly restrictive, condition to check for the convexity of the underlying Gaussian covariance function for $n \geq 2$. For instance, as τ becomes large, the covariance function $r_Y(\tau)$ is well-approximated by the second-derivative of the structure function $\sigma_Y^2|\tau|^{2H}$, and it is easy to show that we must have the asymptotic behavior

$$\frac{\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau)}{\frac{dW^{-1}(r_Y)}{dr_Y}(\tau)} \geq -\frac{2}{r_Y(\tau)},\ \tau \to \infty \tag{113}$$

We now focus on the examples from the text.

B.1.3 Lognormal density

For the lognormal warping (73), we have

$$\frac{dW^{-1}(r_Y(\tau))}{dr_Y} = \left(\mu_Y^2 + r_Y(\tau)\right)^{-1}, \tag{114}$$

$$\frac{d^2W^{-1}(r_Y(\tau))}{dr_Y^2} = -\left(\mu_Y^2 + r_Y(\tau)\right)^{-2}. \tag{115}$$

This leads us to

$$\frac{\frac{d^2W^{-1}(r_Y(\tau))}{dr_Y^2}}{\frac{dW^{-1}(r_Y(\tau))}{dr_Y}} = -\left(\mu_Y^2 + r_Y(\tau)\right)^{-1} \geq \frac{-1}{r_Y(\tau)}, \tag{116}$$

which satisfies (112).

B.1.4 Uniform density

For the uniform warping (64), we see that

$$\begin{aligned}
\frac{dW^{-1}(r_Y)}{dr_Y}(\tau) &= 2(r_Y(\tau))^{-1/2}(1+4r_Y(\tau))^{-3/2}, && (117)\\
\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau) &= -(16r_Y(\tau)+1)(r_Y(\tau))^{-3/2}(1+4r_Y(\tau))^{-5/2}, && (118)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\frac{d^2W^{-1}(r_Y(\tau))}{dr_Y^2}}{\frac{dW^{-1}(r_Y(\tau))}{dr_Y}} &= \frac{-1}{2}\frac{1+16r_Y(\tau)}{r_Y(\tau)(1+4r_Y(\tau))}\\
&= \frac{-8}{1+4r_Y(\tau)} - \frac{-1}{2r_Y(\tau)(1+4r_Y(\tau))}\\
&\geq -8 - \frac{-1}{2r_Y(\tau)}\\
&\geq -\frac{2}{3r_Y(\tau)} - \frac{-1}{2r_Y(\tau)}\\
&\geq -\frac{7}{6r_Y(\tau)}, && (119)
\end{aligned}$$

which satisfies (112).

B.1.5 Exponential density

For the exponential warping (82), we observe that

$$\begin{aligned}
\frac{dW^{-1}(r_Y)}{dr_Y}(\tau) &= \frac{1}{2\mu_Y}(r_Y(\tau))^{-1/2}, && (120)\\
\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau) &= -\frac{1}{4\mu_Y}(r_Y(\tau))^{-3/2}, && (121)
\end{aligned}$$

and

$$\frac{\frac{dW^{-1}(r_Y)}{dr_Y}(\tau)}{\frac{d^2W^{-1}(r_Y)}{dr_Y^2}(\tau)} = \frac{-1}{2r_Y(\tau)}, \qquad (122)$$

which satisfies (112).

B.1.6 Pareto density: $a \geq 3$

In this case, we work with the normalized covariance function $\rho_Y(\tau)$. Since σ_Y^2 just provides a scaling, the bound of (112) applies with $r_Y(\tau)$ replaced by $\rho_Y(\tau)$. For the Pareto warping (91), we obtain

$$\frac{dW^{-1}(\rho_Y)}{d\rho_Y}(\tau) = \frac{a(a-1)(a-2)}{2(\rho_Y(\tau))^{1/2}(a^2-2a+\rho_Y(\tau))^{3/2}}, \quad (123)$$

$$\frac{d^2W^{-1}(\rho_Y)}{d\rho_Y^2}(\tau) = -\frac{a(a-1)(a-2)(a^2-2a+4\rho_Y(\tau))}{4(\rho_Y(\tau))^{3/2}(a^2-2a+\rho_Y(\tau))^{5/2}}, \quad (124)$$

and

$$\frac{\frac{dW^{-1}(\rho_Y)}{d\rho_Y}(\tau)}{\frac{d^2W^{-1}(\rho_Y)}{d\rho_Y^2}(\tau)} = -\frac{a^2-2a+4\rho_Y(\tau)}{2\rho_Y(\tau)(a^2-2a+\rho_Y(\tau))}. \quad (125)$$

We provide bounds for the case $a > 4$, when the fourth moment exists. It can be shown that for $a > 2$ the Pareto satisfies the asymptotic bound of (113), but not necessarily that of (112).

We bound the numerator from above and the denominator from below using the fact that $0 \leq \rho_Y(\tau) \leq 1$ and $a > 4$ to obtain

$$\frac{\frac{dW^{-1}(\rho_Y)}{d\rho_Y}(\tau)}{\frac{d^2W^{-1}(\rho_Y)}{d\rho_Y^2}(\tau)} > -\frac{7}{6\rho_Y(\tau)}, \quad (126)$$

which satisfies (112).

B.2 Case II: $n = 0, 1$

For each of the aforementioned pdfs, to guarantee convexity we must also verify that

$$\Delta^2 r_X[0] = W^{-1}(r_Y[0]) - 2W^{-1}(r_Y[1]) + W^{-1}(r_Y[2]) \geq 0, \quad (127)$$

$$\Delta^2 r_X[1] = W^{-1}(r_Y[1]) - 2W^{-1}(r_Y[2]) + W^{-1}(r_Y[3]) \geq 0. \quad (128)$$

For these pdfs, the second differences at $n = 0$ and $n = 1$ are complicated expressions that are difficult to bound analytically. Therefore, we provides plots that, in conjunction with the smoothness of the underlying functions as a function of H, indicate that the second differences $\Delta^2 r_X[0]$ and $\Delta^2 r_X[1]$ are non-negative for $1/2 \leq H \leq 1$, with the zero value achieved only at the boundaries $H = 1/2$ and $H = 1$. The first step to applying these plots and keeping them 1-d is to show that the values of the parameters of the density functions do not effect the convexity.

For instance, we can set the parameter $\mu_Y = 1$ for the exponential function and assume the uniform lies in the region $[0, 1]$ without loss of generality. Examination of (78) combined with the fact that $r_Y[n] = \frac{(a-b)^2}{12}\rho_Y[n]$ shows that the choice of region has no effect on the covariance $r_X[n]$.

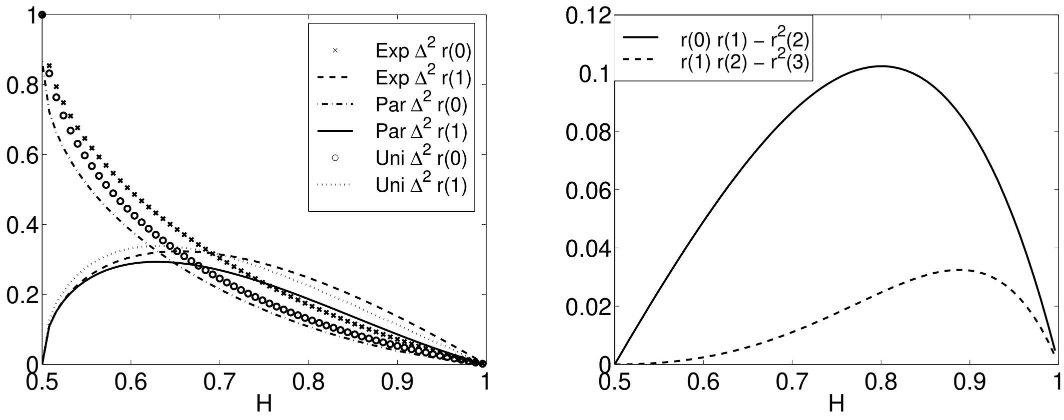
For the Pareto density, we now show that we can set the parameter $a = 3$ without loss of generality. The second difference of the underlying Gaussian correlation function is given by

$$\Delta^2 r_X[n] = (a-1)\left(\frac{\sqrt{\rho_Y[n]}}{\sqrt{a^2-2a+\rho_Y[n]}} - 2\frac{\sqrt{\rho_Y[n+1]}}{\sqrt{a^2-2a+\rho_Y[n+1]}} + \frac{\sqrt{\rho_Y[n+2]}}{\sqrt{a^2-2a+\rho_Y[n+2]}}\right). \quad (129)$$

Since $\rho_Y[n+1] \geq \rho_Y[n+2]$, (129) is bounded below by

$$\begin{aligned}
\Delta^2 r_X[n] &\geq (a-1)\left(\frac{\frac{\sqrt{a^2-2a+\rho_Y[n+1]}}{\sqrt{a^2-2a+\rho_Y[n]}}\sqrt{\rho_Y[n]} - 2\sqrt{\rho_Y[n+1]} + \sqrt{\rho_Y[n+2]}}{\sqrt{a^2-2a+\rho_Y[n+1]}}\right) \\
&\geq \frac{\sqrt{a^2-2a+\rho_Y[n+1]}}{\sqrt{a^2-2a+\rho_Y[n]}}\sqrt{\rho_Y[n]} - 2\sqrt{\rho_Y[n+1]} + \sqrt{\rho_Y[n+2]}. \quad (130)
\end{aligned}$$

Since $\rho_Y[n] \geq \rho_Y[n+1]$, (130) is monotonically increasing in a. Hence, in Figure 7(a), we plot (130) for $a = 3$, $H \in [1/2, 1]$, $n = 0$, and $n = 1$ in order to show that $\Delta^2 r_X[0]$ and $\Delta^2 r_X[1]$ are non-negative for all $a \geq 3$.



(a) (b)

Figure 7: *Demonstration of the convexity of the underlying Gaussian covariance function $r_X[n]$ at $n = 0$ and $n = 1$ for the problem of nonGaussian ESS process synthesis. (a) The second differences $\Delta^2 r_X[0]$ and $\Delta^2 r_X[1]$ (or lower bounds) are non-negative for $H \in [1/2, 1]$ for the exponential, Pareto, and uniform densities. (b) The terms $r_Y[n]r_Y[n+2] - r_Y^2[n+1]$ are non-negative for the lognormal density with $H \in [1/2, 1]$ and $n = 0, 1$. This implies that $\Delta^2 r_X[0] > 0$ and $\Delta^2 r_X[1] > 0$.*

Finally, for the lognormal, in order to demonstrate non-negativity in a single one-dimensional plot, we must remove the extra parameter μ_Y^2. Note that the second difference for the lognormal function is given by

$$\Delta^2 r_X[n] = \log\left(\frac{(\mu_Y^2 + r_Y[n-1])(\mu_Y^2 + r_Y[n+1])}{(\mu_Y^2 + r_Y[n])^2}\right). \quad (131)$$

Since $r_Y[n-1]+r_Y[n+1]>2r_Y[n]$ by the convexity of $r_Y[n]$, we see that (131) is non-negative if $r_Y[n-1]r_Y[n+1]\geq r_Y^2[n]$. Hence, in Figure 7(b) we plot $r_Y[n]r_Y[n+2]-r_Y^2[n+1]$ as a function of H for $n=0$ and $n=1$ to demonstrate that $\Delta^2 r_X[0]$ and $\Delta^2 r_X[1]$ are non-negative.

References

[1] J. Feder, *Fractals.* New York: Plenum Press, 1988.

[2] M. Keshner, "$1/f$ noise," *Proc. of the IEEE*, vol. 70, pp. 212–218, March 1982.

[3] B. West and M. Shlesinger, "On the ubiquity of $1/f$ noise," *International Journal of Modern Physics, B*, vol. 3, no. 6, pp. 795–819, 1989.

[4] L. M. Kaplan and C.-C. J. Kuo, "An improved method for 2-d self-similar image synthesis," *IEEE Trans. Image Proc.*, vol. 5, pp. 754–761, May 1996.

[5] T. Lundahl, W. Ohley, S. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. on Medical Imaging*, vol. 5, pp. 152–161, Sep. 1986.

[6] A. E. Jacquin, "Fractal image coding: A review," *Proc. of IEEE*, vol. 81, pp. 1451–1465, Oct. 1993.

[7] T. Bollerslev and H. O. Mikkelsen, "Modeling and pricing long memory in stock market volatility," *Journal of Econometrics*, vol. 73, pp. 151–184, 1996.

[8] B. Mandelbrot, "A multifractal walk down wall street," *Scientific American*, pp. 70–73, Feb. 1999.

[9] A. Arneodo, Y. D'Aubenton-Carafa, B. Audit, E. Bacry, J. Muzy, and C. Thermes, "What can we learn with wavelets about DNA sequences," *Physica A*, vol. 249, pp. 439–448, 1998.

[10] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.

[11] A. Erramilli, O. Narayan, and W. Willinger, "Experimental queueing analysis with long-range dependent traffic," *IEEE/ACM Trans. Networking*, pp. 209–223, April 1996.

[12] B. Mandelbrot and J. W. V. Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, pp. 422–437, Oct. 1968.

[13] D. P. Heyman and T. V. Lakshman, "What are the implications of long-range dependence for VBR-video traffic engineering?," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 301–317, June 1996.

[14] C. Dietrich and G. Newsam, "Fast and exact simulation of stationary Gaussian processes processes through circulant embedding of the covariance matrix," *Siam J. Sci. Comput.*, vol. 18, pp. 1088–1107, July 1997.

[15] A. Feldmann, A. C. Gilbert, W. Willnger, and T. Kurtz, "Looking behind and beyond self-similarity: Scaling phenomena in measured WAN traffic," in *Proc. of 35th Annual Allerton Conf. on Comm., Control, and Computing*, pp. 269–280, June 1997.

[16] A. L. Neidhardt and J. L. Wang, "The concept of relevant time scales and its application to queuing analysis of self-similar traffic," in *Proc. SIGMETRICS '98/PERFORMANCE '98*, pp. 222–232, 1998.

[17] D. Avnir, O. Biham, D. Lidar, and O. Malcai, "Is the geometry of nature fractal?," *Science*, vol. 279, pp. 39–40, Jan 1998.

[18] M. Grigoriu, *Applied Non-Gaussian Processes.* Englewood Cliffs, NJ: Prentice Hall, 1995.

[19] B. Liu and D. Munson, "Generation of a random sequence having a jointly specified marginal distribution and autocovariance," *IEEE Trans. on Acoustics, Speech, and Signal Proc.*, vol. 30, pp. 973–983, Dec. 1982.

[20] G. W. Wornell, *Signal Processing with Fractals: A Wavelet-based Approach.* New Jersey: Prentice Hall, 1995.

[21] P. Flandrin, "On the spectrum of fractional Brownian motion," *IEEE Trans. Inform. Theory*, vol. 35, pp. 197–199, Jan. 1989.

[22] D. Cox, "Long-range dependence: A review," *Statistics: An Appraisal*, pp. 55–74, 1984.

[23] S. M. Kay, *Modern Spectral Estimation: Theory and Application.* EngleWood Cliffs, New Jersey: Prentice-Hall, 1988.

[24] B. Tsybakov and N. Georganas, "Self-similar processes in communications networks," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1713–1725, Sep. 1998.

[25] L. Kaplan and C.-C. Kuo, "Extending self-similarity for fractional Brownian motion," *IEEE Trans. Signal Proc.*, vol. 42, pp. 3526–3530, Dec. 1994.

[26] A. M. Yaglom, *Correlation Theory of Stationary and Related Random Functions: Basic Results.* New York: Springer-Verlag, 1987.

[27] M. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for long-range dependence: An empirical study," *Fractals.*, vol. 3, pp. 785–798, 1995.

[28] V. Paxson, "Fast, approximate synthesis of fractional Gaussian noise for generating self-similar network traffic," *Computer Communication Review*, vol. 27, pp. 5–18, Oct. 1997.

[29] W.-C. Lau, A. Erramilli, J. Wang, and W. Willinger, "Self-similar traffic generation: The random midpoint displacement algorithm and its properties," in *IEEE Int. Conf. on Comm.*, pp. 466–472, June 1995.

[30] G. Corsini and R. Saletti, "A $1/f^{\gamma}$ power spectrum noise sequence generator," *IEEE Trans. on Instrumentation and Measurement*, vol. 37, pp. 615–619, Dec. 1988.

[31] J. Hosking, "Modeling persistence in hydrological time series using fractional differencing," *Water Resources Research*, vol. 20, pp. 1898–1908, Dec. 1984.

[32] N. Kasdin, "Discrete simulation of colored noise and stochastic process and $1/f^{\alpha}$ power law noise generation," *Proc. of the IEEE*, vol. 83, pp. 802–827, May 1995.

[33] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, pp. 910–916, Mar. 1992.

[34] L. Kaplan and C.-C. Kuo, "Fractal estimation from noisy data via discrete fractional Gaussian noise (DFGN) and the Haar basis," *IEEE Trans. Signal Proc.*, vol. 41, pp. 3554–3562, Dec. 1993.

[35] R. H. Riedi, M. S. Crouse, V. Ribiero, and R. G. Baraniuk, "A multifractal wavelet model with application to TCP network traffic," *IEEE Trans. Inform. Theory*, vol. 45, pp. 992–1018, April 1999.

[36] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications.* New York: John Wiley & Sons, 1993.

[37] A. Dembo, C. Mallows, and L. Shepp, "Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1206–1212, Nov. 1989.

[38] G. Newsam and C. Dietrich, "Bounds on the size of nonnegative definite circulant embeddings of positive definite Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1218–1220, July 1994.

[39] S. Barnett, *Matrices, Methods, and Applications.* Oxford: Clarendon Press, 1990.

[40] C. Burrus and T. Parks, *DFT/FFT and Convolution Algorithms.* New York: John Wiley & Sons, 1985.

[41] L. Devroye, *Non-uniform random variate generation.* New York: Springer-Verlag, 1986.

[42] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1-2. New York: John Wiley & Sons, 1994.

[43] W. J. Szajnowski, "The generation of correlated Weibull clutter for signal detection problems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-13, pp. 536–540, Sep. 1977.

[44] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1991.

[45] F. Breidt, N. Crato, and P. de Lima, "Modeling the persistent volatility of asset returns," in *IEEE/IAFE Conf. on Computational Intelligence for Financial Engineering*, (New York, NY), pp. 266–272, Mar. 1997.