Wavelet-Based Representations for the $1/f$ Family of Fractal Processes

GREGORY W. WORNELL, MEMBER, IEEE

The $1/f$ family of fractal random processes model a truly extraordinary range of natural and man-made phenomena, many of which arise in a variety of signal processing scenarios. Yet despite their apparent importance, the lack of convenient representations for $1/f$ processes has, at least until recently, strongly limited their popularity. In this paper, we demonstrate that $1/f$ processes are, in a broad sense, optimally represented in terms of orthonormal wavelet bases. Specifically, via a useful frequency-domain characterization for $1/f$ processes, we develop the wavelet expansion's role as a Karhunen–Loève-type expansion for $1/f$ processes. As an illustration of potential, we show that wavelet-based representations naturally lead to highly efficient solutions to some fundamental detection and estimation problems involving $1/f$ processes.

I. Introduction

There is a wide range of engineering contexts in which there is a need to be able to synthesize, analyze, and process fractal signals. Indeed, fractal geometry abounds in nature. Fractal waveforms arise, for example, in natural landscapes, in the distribution of earthquakes, in ocean waves, in turbulent flow, in the pattern of errors on communication channels, and even in fluctuations of the stock market.

Some of the most prevalent forms of fractal geometry arise out of statistical scaling behavior in physical phenomena. An important class of fractal signals with this character are the $1/f$ processes [1]. These statistically self-similar random processes exhibit rich behavior. They are generally nonstationary, yet they possess stationary attributes: when viewed through certain linear time-invariant filters they appear stationary. Furthermore, their sample functions exhibit strong long-term correlation structure with polynomial-type decay that cannot be captured with traditional—e.g., autoregressive-moving average—time-series models. In fact, the development of models for $1/f$-type behavior remains a research topic of much interest in engineering, physics, and mathematics communities [2].

Manuscript received January 16, 1992; revised October 30, 1992. This work was supported in part by the Advanced Research Projects Agency monitored by ONR under Contract N00014-89-J-1489, and the Air Force Office of Scientific Research under Grant AFOSR-91-0034.

The author is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9212427.

While universally acceptable characterizations and representations for $1/f$ processes have, for the most part, remained elusive, a number of useful models for $1/f$ processes can be found in the literature. One class of models is based upon a fractional integral formulation. A popular example of this class is the fractional Brownian motion framework which, although developed more recently by Mandelbrot and Van Ness [3], also appeared independently in the work of Kolmogorov. This model can be viewed as a refinement of an earlier construction due to Barnes and Allan [4].

Other models for $1/f$ processes fall into the category of extended or infinite-order ARMA models. An example is the well-known "superposition of Lorenzian spectra" model developed by van der Ziel [5]. The origins of this model can be traced back at least to the work of Bernamont in the 1930's [6]. Another example is the infinite, continuous transmission line model of Keshner [1] which corresponds to an ARMA model with alternating poles and zeros exponentially spaced along the negative real axis of the s-plane.

While these various models have provided much insight into $1/f$ behavior, their use in signal processing contexts has been limited. In the case of fractional Brownian motion, issues of mathematical tractability inhibit the solution of many basic detection and estimation problems, although significant progress in this direction has been reported in Barton and Poor [7]. In addition, the development of practical data-driven algorithms for processing $1/f$ data using fractional Brownian motion models has traditionally proved challenging. However, work in this direction is described in Lundahl *et al.* [8] and, more recently, both in Deriche and Tewfik [9] and in Kaplan and Kuo [10]. Finally, in the signal processing context, the use of extended ARMA models has generally proven useful only in the synthesis of $1/f$ processes; see, e.g., [11].

In this paper, we develop a powerful, signal-processing-oriented representation for $1/f$ processes in terms of orthonormal wavelet bases, i.e., bases in which the basis functions are all dilations and translations of a single prototype function. Because $1/f$ processes simultaneously

0018–9219/93$03.00 © 1993 IEEE

exhibit both statistical scale invariance and a particular notion of time invariance, wavelet-based representations are, in many respects, ideally suited for these processes. In fact, the wavelet transform constitutes as natural a tool for the manipulation of these processes as the Fourier transform does for stationary processes. Furthermore, just as the discovery of fast Fourier transform (FFT) algorithms dramatically increased the viability the Fourier-based processing of stationary signals in real systems, the existence of fast discrete wavelet transform (DWT) algorithms for implementing wavelet transformations means that wavelet-based representations of $1/f$ signals are also of great practical significance. Indeed, they allow the development of highly efficient signal processing algorithms for $1/f$ data.

As the core of the paper, we develop the wavelet expansion's role as a Karhunen–Loève-type expansion for $1/f$ processes. Specifically, via a useful frequency-domain characterization for $1/f$ processes, we obtain both synthesis and analysis results. In effect, the synthesis result establishes that under suitable but mild conditions on the wavelet basis, $1/f$-like behavior can be synthesized from wavelet expansions whose coefficients are mutually uncorrelated and have a particular variance structure. Analogously, the analysis result establishes that under similar conditions, the coefficients of a wavelet-based expansion of a $1/f$ process are effectively uncorrelated and have the expected variance structure. Not only are we able to validate these results through simulations, but, in addition, we are able to identify potential examples of physiological and economic data consistent with the model.

The final aspect of the paper demonstrates the potential convenience and efficiency of wavelet-based representations in signal processing applications. In particular, we show that these representations lead to both convenient descriptions of the whitening filters for $1/f$ processes, and, in turn, to efficient and robust solutions to a number of fundamental signal processing problems involving $1/f$ processes. The examples considered include minimum probability of error signal detection and discrimination, maximum-likelihood parameter estimation, and minimum-mean-square error smoothing. While not explored in this paper, there are some interesting parallels between the these algorithms and the multiscale signal processing algorithms discussed in [12].

The format of the paper is as follows. Section II briefly reviews the theory of orthonormal wavelet bases, emphasizing a signal processing perspective. The first half of Section III reviews $1/f$ processes and their properties, and the fractional Brownian motion framework. The latter half of the section develops a useful frequency-domain characterization for $1/f$ processes, and explores its properties and relationships to fractional Brownian motion. Section IV then develops the wavelet-based representation for $1/f$ processes, and its properties. Finally, Section V applies the representation to the solution of some example problems of optimal detection and estimation, and Section VI contains some concluding remarks.

II. Orthonormal Wavelet Bases

In this section, we review those aspects of wavelet theory that are important in the context of this paper and introduce some notational conventions. In light of our application, we explicitly adopt a signal processing perspective. From this point of view, it is apparent that while wavelet theory is in some sense new, many of the ideas underlying wavelets—including constant-Q filter banks and time–frequency analysis [13], and quadrature mirror and conjugate quadrature filters [14]—are rather familiar. For the interested reader, a number of extensive introductions to wavelet theory are available; see, e.g., [15]–[18].

An orthonormal wavelet transformation of a signal $x(t)$ is generally described in terms of the synthesis/analysis equations

$$x(t) = \sum_m \sum_n x_n^m \psi_n^m(t) \tag{1a}$$

$$x_n^m = \int_{-\infty}^{\infty} x(t)\, \psi_n^m(t)\, dt \tag{1b}$$

and has the special property that the orthogonal basis functions are all dilations and translations of a single function referred to as the *basic wavelet* $\psi(t)$. In particular

$$\psi_n^m(t) = 2^{m/2}\psi(2^m t - n) \tag{2}$$

where m and n are the dilation and translation indices, respectively.

An important example of a wavelet basis, and one to which we will frequently refer, is that derived from the ideal bandpass wavelet. This wavelet, which we specifically denote by $\tilde{\psi}(t)$, is the impulse response of an ideal bandpass filter with frequency response

$$\tilde{\psi}(\omega) = \begin{cases} 1, & \pi < |\omega| \le 2\pi \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

More generally, however, the constraints on the Fourier transform $\Psi(\omega)$ of the basic (or "mother") wavelet $\psi(t)$ are less stringent. These include

$$|\Psi(\omega)| \le 1 \tag{4a}$$

$$\Psi(0) = 0 \tag{4b}$$

$$\Psi(\omega) \sim \mathcal{O}(|\omega|^{-1}), \qquad |\omega| \to \infty \tag{4c}$$

where (4c) is a consequence of restricting our attention to *regular* bases.[1] For wavelet bases with higher order regu-

[1] A function $f(t)$ is Rth-order regular if its Fourier transform $F(\omega)$ decays according to

$$F(\omega) \sim \mathcal{O}(|\omega|^{-R}), \qquad |\omega| \to \infty$$

where the notation $\mathcal{O}(\cdot)$ is to be understood to mean that

$$\lim_{\omega \to \infty} \frac{F(\omega)}{|\omega|^{-R}} < \infty.$$

Such functions have $R-1$ regular derivatives, where the term "regular" denotes a function that is at least first-order regular.

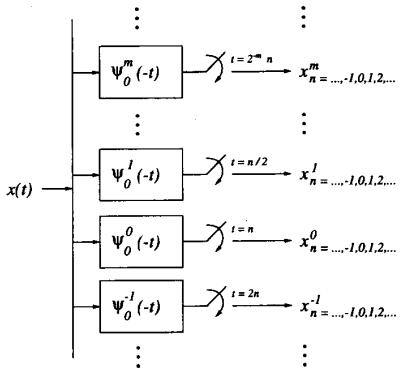

Fig. 1. Critically sampled filter bank interpretation of an orthonormal wavelet decomposition.

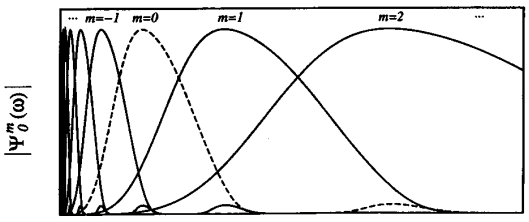

Fig. 2. The octave band filters corresponding to an orthonormal wavelet decomposition. The wavelet basis in this example is one due to Daubechies [19].

larity, as are required in many contexts, the corresponding spectral decay is even stronger. For a wavelet basis to have Rth-order regularity, it is sufficient (though not necessary) that $\psi(t)$ possess R vanishing moments:

$$\int_{-\infty}^{\infty} t^r \psi(t)\, dt = (j)^r \Psi^{(r)}(0) = 0, \qquad r = 0, 1, \cdots, R-1.$$

In any event, $\psi(t)$ is often the impulse response of an at least roughly bandpass filter. In these cases, the wavelet transformation can typically be interpreted either in terms of a generalized constant-Q (specifically, octave-band) filter bank, or in terms of a multiresolution signal analysis.

A. An Octave-Band Filter Bank Interpretation

The filter bank interpretation of the wavelet transform, depicted in Fig. 1, is obtained by viewing the analysis equation (1b), for each m, as a *filter-and-sample operation,* viz.,

$$x_n^m = \{x(t) * \psi_0^m(-t)\}|_{t=2^{-m}n}.$$

Although the interpretation applies to a broader class of bases, it is often convenient to visualize the basis associated with the ideal-bandpass wavelet (3). In this case, the output of each filter in the bank corresponding to $m = \cdots, -1, 0, 1, 2, \cdots$ is sampled at the corresponding Nyquist rate. More generally, the filter bank is *critically sampled* [13] in the sense that reconstruction is not possible if any of the sampling rates are reduced regardless of the choice of wavelet. When the wavelet is reasonably localized in both time and frequency, the decomposition's role as a time–frequency analysis is also apparent: each x_n^m constitutes a measure of the information in $x(t)$ within a frequency band roughly of width 2^m near $\omega \approx 2^m\pi$ and within a time interval roughly of duration 2^{-m} near $t \approx 2^{-m}n$.

For a particular choice of wavelet basis, there can be significant spectral overlap in the frequency response magnitudes of the filters in the bank, as depicted in the example of Fig. 2. In such cases, however, orthogonality is still preserved through an appropriate choice of phase

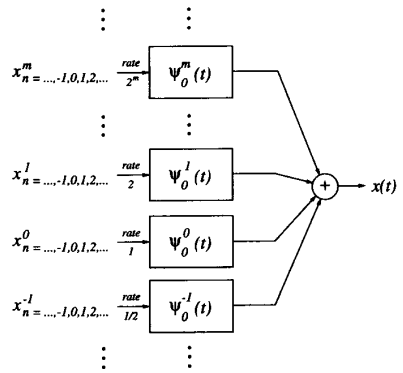

Fig. 3. Interpretation of an orthonormal wavelet expansion as a multirate modulation scheme.

in the prototype filter. Nevertheless, it is possible to construct practical wavelet bases such that the spectral overlap between channels is small in applications where this is important.

Finally, with the wavelet transform analysis equation (1b) interpreted as a filter bank decomposition, the corresponding *synthesis* equation (1a) may be interpreted as multirate modulation. In particular, as depicted in Fig. 3, for a given m each sequence of coefficients x_n^m is modulated onto the corresponding wavelet dilate $\psi_0^m(t)$ at rate 2^m. For the case of the ideal-bandpass wavelet, this corresponds to modulating each such sequence x_n^m into the distinct octave frequency band $2^m\pi < \omega \leq 2^{m+1}\pi$.

B. Multiresolution Signal Analysis Interpretation

The multiresolution signal analysis interpretation allows us to view wavelet analysis as a technique for isolating variations in the signal that occur on different temporal scales. The wavelet decomposition specifically corresponds to a class of linear multiresolution signal analyses in which a signal space $\boldsymbol{V}$ is decomposed into a nested sequence of approximation spaces

$$\cdots \subset \boldsymbol{V}_{-1} \subset \boldsymbol{V}_0 \subset \boldsymbol{V}_1 \subset \boldsymbol{V}_2 \subset \cdots$$

each of which contains those signals of resolution 2^m, with the notion of resolution dependent upon the particular wavelet basis involved. Associated with each $\boldsymbol{V}_m$ is a linear

approximation operator A_m defining projections from $\boldsymbol{V}$ onto $\boldsymbol{V}_m$, so that $A_m x(t)$ defines the "closest" signal of resolution 2^m to $x(t)$. Conveniently, the nesting structure of the $\boldsymbol{V}_m$ guarantees that coarser scale approximations can be derived successively from finer scale approximations.

Wavelet-based multiresolution analyses have the special property that the nature of the resolution-limited approximations is similar in all time intervals and at all scales. As a consequence, each such analysis can be completely described in terms of a single *scaling function* (or "father" wavelet) $\phi(t)$. Specifically, for each m,

$$\cdots, \phi_{-1}^m(t), \phi_0^m(t), \phi_1^m(t), \phi_2^m(t), \cdots$$

generates an orthonormal basis for $\boldsymbol{V}_m$, where the basis functions are all dilations and translations of one another, i.e.,

$$\phi_n^m(t) = 2^{m/2}\phi(2^m t - n). \tag{5}$$

The corresponding resolution-2^m approximation of a signal $x(t)$ is then expressed as

$$A_m x(t) = \sum_n a_n^m \phi_n^m(t) \tag{6}$$

with the coefficients a_n^m computed according to the projections

$$a_n^m = \int_{-\infty}^{\infty} x(t)\, \phi_n^m(t)\, dt. \tag{7}$$

In general, $\phi(t)$ has a Fourier transform $\Phi(\omega)$ that is at least roughly lowpass. At the very least, however, $\Phi(\omega)$ satisfies the properties

$$|\Phi(\omega)| \leq 1 \tag{8a}$$

$$|\Phi(0)| = 1 \tag{8b}$$

$$\Phi(\omega) \sim \mathcal{O}(|\omega|^{-1}), \qquad \omega \to \infty. \tag{8c}$$

Consistent with our remarks concerning (4c), the last of these is a consequence of a regularity requirement on the bases.

Generally, we may interpret the projection (7) as a lowpass-like filter-and-sample operation, viz.,

$$a_n^m = \{x(t) * \phi_0^m(-t)\}|_{t=2^{-m}n} \tag{9}$$

and (6) as a modulation of these samples onto a lowpass-like waveform. In fact, an important example of a multiresolution analysis, especially in the context of this paper, is generated from the *ideal* lowpass scaling function. This scaling function, which we denote by $\tilde{\phi}(t)$, has a Fourier transform that is the frequency response of an ideal lowpass filter, i.e.,

$$\tilde{\Phi}(\omega) = \begin{cases} 1, & |\omega| \leq \pi \\ 0, & |\omega| > \pi. \end{cases} \tag{10}$$

In this case, the corresponding multiresolution analysis is based upon perfectly bandlimited signal approximations, whereby $A_m x(t)$ represents $x(t)$ bandlimited to $\omega = 2^m\pi$. Furthermore, we may interpret (9) and (6) in the context of classical sampling theory [20]. In particular, $\phi(t)$ in (9) plays the role of an anti-aliasing filter and (6) is the interpolation formula associated with the sampling theorem. While the bandlimited multiresolution analysis is technically unrealizable, one can construct practical analyses which constitute a close approximation to this analysis.

The relationship between a wavelet basis and its associated multiresolution analysis is apparent when we consider the information lost between successive resolution-limited approximations. Each such *detail* signal

$$D_m x(t) = A_{m+1}x(t) - A_m x(t) \tag{11}$$

lies in the orthogonal complement subspace $\boldsymbol{O}_m$ of $\boldsymbol{V}_m$ in $\boldsymbol{V}_{m+1}$. In turn, it is the basic wavelet that directly leads to the orthonormal basis

$$\cdots, \psi_{-1}^m(t), \psi_0^m(t), \psi_1^m(t), \psi_2^m(t), \cdots$$

for $\boldsymbol{O}_m$, where $\psi_n^m(t)$ is as defined in terms of dilations and translations of $\psi(t)$ as per (2). Hence, we have

$$D_m x(t) = \sum_n x_n^m \psi_n^m(t) \tag{12}$$

with (cf. (1b))

$$x_n^m = \int_{-\infty}^{\infty} x(t)\, \psi_n^m(t)\, dt.$$

Furthermore, it is from this point of view that one can recognize that it is the ideal bandpass wavelet (3) that is associated with the bandlimited multiresolution analysis defined via (10).

Combining (11) with (12) we obtain the representation

$$A_M x(t) = \sum_{m<M} D_m x(t) = \sum_{m<M} \sum_n x_n^m \psi_x^m(t) \tag{13}$$

consistent with the notion that $A_M x(t)$ is an approximation to $x(t)$ in which details on scales smaller than 2^M are discarded. Finally, accumulating the detail signals over all m by taking $M \to \infty$ in (13) yields the efficient orthonormal representation

$$x(t) = \sum_m \sum_n x_n^m \psi_n^m(t)$$

that is the synthesis formula (1a).

The multiresolution signal analysis viewpoint not only provides a highly useful conceptual interpretation of the wavelet transform, but perhaps even more importantly, leads to a particularly efficient algorithm for the computation of the wavelet transform of a signal. Furthermore, it is the existence of this algorithm that ultimately makes the representations discussed in this paper of such practical significance. However, a detailed understanding of this algorithm, which is based on an intimate relationship between orthonormal wavelet bases and a class of discrete-time conjugate quadrature filter banks, is not essential to an appreciation of the main results of the paper. A summary for the interested reader is provided in Appendix I.

C. Examples

The ideal bandpass wavelet (3) leads to a basis with excellent frequency localization properties but very poor temporal localization since $\tilde{\psi}(t)$ decays only like $1/t$ for large t. Fortunately, in applications one can choose from among many possible wavelet bases to make a variety of different tradeoffs. In this section, we briefly summarize some other families of wavelet bases.

For example, at another extreme from the bandlimited multiresolution analysis lies the Haar-based multiresolution analysis in which approximations at resolution 2^m are piecewise-constant on intervals of length 2^{-m}. In this case, the corresponding scaling function is given by

$$\phi(t) = \begin{cases} 1, & 0 < t < 1 \\ 0, & \text{otherwise} \end{cases}$$

and the associated wavelet is

$$\psi(t) = \begin{cases} 1, & 0 < t < 1/2 \\ -1, & 1/2 \le t < 1 \\ 0, & \text{otherwise.} \end{cases}$$

This analysis is realizable and exhibits excellent temporal localization but very poor frequency localization— $\Psi(\omega)$ falls off only like $1/\omega$ for $\omega \to \infty$.

The family of Battle–Lemarie wavelet bases [21], [19] is obtained from multiresolution analyses based upon orthogonalized Pth-order spline functions. In fact, in this class the Haar-based wavelet basis corresponds to the case $P = 0$, while the bandpass wavelet basis corresponds to $P \to \infty$. The Battle–Lemarie bases have very reasonable localization properties: they are characterized by exponential decay in the time domain and decay like $1/|\omega|^{P+1}$ in the frequency domain. Hence while they are, strictly speaking, unrealizable, the exponential decay property ensures that good approximations may be realized via truncation.

Finally, an important class of practical wavelet bases due to Daubechies [19] has not only reasonable localization in both time and frequency, but finite-extent basis functions as well. In particular, Daubechies bases of order R are derived from conjugate quadrature filter pairs of length $2R$ for $R = 1, 2, \cdots$, and have the property that the basis functions are *maximally regular* in the sense of having the maximum number of vanishing moments (R) for a given order. More generally, the development of other families of wavelets based on multirate filter bank theory continues to receive considerable attention in the literature, as described in, e.g., [22]–[24].

III. $1/f$ Process Characterizations

The term "$1/f$ process" has traditionally been used to describe a broad class of physical signals having measured power spectra obeying a power law relationship of the form

$$S_x(\omega) \sim \frac{\sigma_x^2}{|\omega|^\gamma} \tag{14}$$

for some spectral parameter γ. In any spectral analysis, of course, data length generally limits access to information at lower frequencies and data resolution limits access to detail at higher frequencies. Nevertheless, there are many examples of phenomena for which arbitrarily large data records justify a $1/f$ spectrum of the form (14) over all accessible frequencies. However, (14) is not integrable and, hence, strictly speaking, does not constitute a valid power spectrum in the theory of stationary random processes. As a consequence, a number of subtleties arise in the development of mathematical models for $1/f$ behavior.

In spite of the paradox of the $1/f$ spectrum, a truly enormous and tremendously varied collection of natural and man-made phenomena exhibit $1/f$-type spectral behavior over many decades of frequency. A partial list includes (see, e.g., [1], [25]–[29] and the references therein):

- geophysical time series such as variation in temperature and rainfall records, measurements of oceanic flows, flood level variation in the Nile river, wobble in the Earth's axis, frequency variation in the Earth's rotation, and sunspot variations;
- economic time series such as the Dow Jones Industrial Average;
- physiological time series such as instantaneous heart rate records for healthy patients, EEG variations under pleasing stimuli, and insulin uptake rate data for diabetics;
- biological time series such as voltages across nerve and synthetic membranes;
- electromagnetic fluctuations such as in galactic radiation noise, the intensity of light sources, and flux flow in superconductors;
- electronic device noises in field effect and bipolar transistors, vacuum tubes, and Schottky, Zener, and tunnel diodes;
- resistance fluctuations in metal film, semiconductor films and contacts, germanium filaments in carbon and aqueous solution, thermocells, and concentrations cells;
- frequency variation in hourglasses, quartz crystal oscillators, atomic clocks, and superconducting cavity resonators;
- man-induced phenomena including variations in traffic flow and amplitude and frequency variation in Western, African, Asian, and Indian music, both modern and traditional;
- generation of perceptually pleasing physiological stimuli, such as artificial music and breezes;
- burst errors on communication channels;
- texture variation in natural terrain, landscapes, and cloud formations.

While $\gamma \approx 1$ in many of these examples, more generally $0 \le \gamma \le 2$. However, there are many examples of phenomena in which γ lies well outside this range. For $\gamma \ge 1$, the lack of integrability of (14) in a neighborhood of the spectral origin reflects the preponderance of low-frequency energy in the corresponding processes. This phenomenon is termed the infrared (IR) catastrophe. For many physical phenomena, measurements corresponding

to very small frequencies show no low-frequency roll off, which is usually understood to reveal an inherent nonstationarity in the underlying process. Such is the case for the Wiener process, with which it is sometimes convenient to associate a $1/f$ spectrum with $\gamma = 2$. For $\gamma \leq 1$, the lack of integrability in the tails of the spectrum reflects a preponderance of high-frequency energy and is termed the ultraviolet (UV) catastrophe. Such behavior is familiar for generalized Gaussian processes such as stationary white Gaussian noise, which may be considered a $1/f$ process with $\gamma = 0$ and its usual derivatives. In general, it is agreed that the transition from stationary to nonstationary behavior occurs at some parameter value in the range $0 < \gamma < 2$, although opinions differ as to the exact transition point [27].

Physical $1/f$ processes possess a number of special characteristics. For example, as is reflected by the fact that (14) obeys, for every a, the scaling equation

$$S_x(\omega) = |a|^\gamma S_x(a\omega).$$

$1/f$ processes possess *statistical self-similarity*, i.e., their statistics are invariant to dilations and contractions of the time axis, to within an amplitude factor. As such, $1/f$ processes have no "characteristic scale." In turn, the self-similarity inherent in $1/f$ processes typically gives rise to fractal structure in the associated time waveforms.

Another related property of $1/f$ processes is their so-called persistent statistical dependence. Indeed, the generalized Fourier pair [30]

$$\frac{|\tau|^{\gamma-1}}{2\Gamma(\gamma)\cos(\gamma\pi/2)} \leftrightarrow \frac{1}{|\omega|^\gamma} \tag{15}$$

valid for $\gamma > 0$ but $\gamma \neq 1, 2, 3, \cdots$ suggests that the autocorrelation $R_x(\tau)$ associated with the spectrum (14) for $0 < \gamma < 1$ is characterized by slow decay of the form

$$R_x(\tau) \sim |\tau|^{\gamma-1}.$$

In fact, this power-law decay in correlation structure distinguishes $1/f$ processes from many traditional models used for time series analysis. For example, the well-studied family of autoregressive moving-average (ARMA) models have a correlation structure invariably characterized by *exponential* decay. As a consequence, ARMA models are generally inadequate for capturing long-term dependence in data.

Perhaps the most popular mathematical characterization of $1/f$ processes is obtained through the concept of fractional Brownian motion. In the next section, we briefly review this model, and, in the process, reveal many important properties of $1/f$ processes. Unavoidably, several mathematical subtleties arise in the development of fractional Brownian motion, making this section somewhat less accessible to the nonspecialist than the remainder of the paper. Furthermore, while insightful, a detailed understanding of these subtleties is not essential to appreciating the main results of the paper. For these reasons, the reader may find it easier to skip over this section on a first reading, proceeding directly to Section III-B. There we develop a powerful but much simpler mathematical characterization for $1/f$ processes.

A. Fractional Brownian Motion and Fractional Gaussian Noise

The $1/f$ processes can be viewed as a particular class of statistically self-similar random processes. Using a rather well-established definition, a random process $x(t)$ is said to be statistically self-similar with parameter H if for any real $a > 0$ it obeys the scaling relation

$$x(t) \stackrel{\mathcal{P}}{=} a^{-H} x(at) \tag{16}$$

where $\stackrel{\mathcal{P}}{=}$ denotes equality in a statistical sense. For *strict-sense* self-similar processes, this equality is in the sense of all finite-dimensional joint probability distributions. For *wide-sense* self-similar processes, the equality may be interpreted in the sense of second-order statistics, i.e., mean and covariance functions. In this latter case, the self-similarity relation (16) may be alternately expressed as

$$M_x(t) \triangleq E[x(t)] = a^{-H} M_x(at) \tag{17a}$$

$$R_x(t,s) \triangleq E[x(t)\,x(s)] = a^{-2H} R_x(at, as). \tag{17b}$$

For Gaussian processes, upon which we will focus our attention, the two definitions are, of course, equivalent. Let us further restrict our attention to zero-mean processes.

A seemingly natural definition of a Gaussian $1/f$ process $x(t)$ with spectral exponent γ would be the result of driving stationary white Gaussian noise $w(t)$ through a linear time-invariant system with impulse response, for $H > -1/2$

$$v(t) = \frac{1}{\Gamma(H+1/2)} t^{H-1/2} u(t) \tag{18}$$

for which the system function is, for $\gamma = 2H + 1$, [31]

$$\Upsilon(s) = \frac{1}{s^{\gamma/2}}$$

where $\Gamma(\cdot)$ is the gamma function. Unfortunately, however, because the system defined via (18) is unstable except for the degenerate case $H = -1/2$, the convolution

$$x(t) = v(t) * w(t) = \frac{1}{\Gamma(H+1/2)} \cdot \int_{-\infty}^{t} (t-\tau)^{H-1/2} w(\tau)\, d\tau \tag{19}$$

is not well-defined.

The remedy of Barnes and Allan [4] to the dilemma posed by this construction was to key the integration in (19) to the time domain origin in their model, thereby defining their model for $1/f$ behavior through the fractional (Riemann–Liouville) integral [32]

$$x(t) = \frac{1}{\Gamma(H+1/2)} \int_0^t |t-\tau|^{H-1/2} w(\tau)\, d\tau. \tag{20}$$

Unfortunately, while statistically self-similar with parameter H, the Barnes–Allan process fails to exhibit the spectral

behavior associated with $1/f$ processes, even when the definition is extended for $t < 0$ through the convention

$$\int_0^t \triangleq -\int_t^0. \tag{21}$$

Fractional Brownian motion (fBm) represents a useful refinement of the Barnes–Allan process. In the fractional Brownian motion framework, processes corresponding to $1 < \gamma < 3$, for which there is infinite low-frequency power, are developed as nonstationary self-similar random processes having finite power in any finite time interval. These processes are the fractional Brownian motions, and classical Brownian motion is a special case corresponding to $\gamma = 2$. By contrast, processes corresponding to $-1 < \gamma < 1$, for which there is infinite high-frequency power, are developed as generalized stationary Gaussian processes corresponding to the derivative of a fractional Brownian motion. It has become popular to refer to these processes as fractional Gaussian noises, with stationary white Gaussian noise as a special case corresponding to $\gamma = 0$. The theory does not accommodate the cases $\gamma > 3$ and $\gamma < -1$. Furthermore, the models are degenerate for the cases $\gamma = -1, \gamma = 1$, and $\gamma = 3$.

Loosely speaking, fractional Brownian motions are the class of Gaussian statistically self-similar random processes having, in some sense, stationary derivatives. More precisely, because such processes are technically not differentiable, a fractional Brownian motion is a nonstationary Gaussian self-similar process $x(t)$ satisfying $x(0) = 0$ whose increment process

$$\Delta x(t;\epsilon) \triangleq \frac{x(t+\epsilon) - x(t)}{\epsilon} \tag{22}$$

is stationary[2] (and self-similar) for every $\epsilon > 0$. The stationary increment condition is important, in the sense that this is the stationary attribute that ultimately leads to a meaningful notion of spectra for these nonstationary processes.

Fractional Brownian motion defined in this manner can be expressed, for $0 < H < 1$, in the form

$$x(t) \triangleq \frac{1}{\Gamma(H+1/2)} \cdot \left[\int_{-\infty}^{0} (|t-\tau|^{H-1/2} - |\tau|^{H-1/2}) w(\tau)\, d\tau + \int_0^t |t-\tau|^{H-1/2} w(\tau)\, d\tau\right] \tag{23}$$

where $w(t)$ is stationary white Gaussian noise, and where for $t < 0, x(t)$ is defined through the convention (21). From (23), we see directly that for $H = 1/2$, fractional Brownian motion specializes to the Wiener process, i.e., classical Brownian motion.

[2] In fact, fractional Brownian motions for $0 < H < 1$ constitute the *only* statistically self-similar, zero-mean, mean-square continuous, finite-variance, Gaussian random processes satisfying $x(0) = 0$ and having stationary increments.

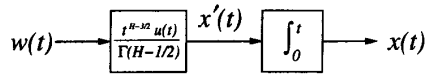

Fig. 4. Synthesis of fractional Brownian motion $x(t)$ in terms of fractional Gaussian noise $x'(t)$ and stationary white Gaussian noise $w(t)$.

Likewise, the correlation function for fractional Brownian motion can be readily derived as

$$R_x(t,s) = E[x(t)\,x(s)] = \frac{\sigma_H^2}{2}(|s|^{2H} + |t|^{2H} - |t-s|^{2H}) \tag{24}$$

where

$$\sigma_H^2 = \operatorname{Var} x(1) = \Gamma(1-2H)\frac{\cos(\pi H)}{\pi H}. \tag{25}$$

From (24), it is trivial to verify that the process is statistically self-similar with parameter H.

While fractional Brownian motion is not differentiable, like regular Brownian motion, it does have a generalized derivative. From (23) we can express this derivative, which is referred to as the associated "fractional Gaussian noise," in the form

$$x'(t) = \frac{d}{dt}x(t) = \lim_{\epsilon\to 0}\Delta x(t;\epsilon) = \frac{1}{\Gamma(H'+1/2)} \cdot \int_{-\infty}^{t} |t-\tau|^{H'-1/2} w(\tau)\, d\tau \tag{26}$$

where we have defined

$$H' = H - 1. \tag{27}$$

Equation (26) now provides an interpretation for convolutions of the form (19), and allows us to view fractional Gaussian noise as a $1/f$ process with spectral exponent $\gamma' = 2H' + 1$. The resulting conceptually useful synthesis for fractional Brownian motion depicted in Fig. 4.

The character of fractional Gaussian noise $x'(t)$ strongly depends on the value of H. Indeed, as shown in [7], for large lags the correlation of the increment process has the same algebraic sign as $H - 1/2$, i.e., for $|\tau| \gg \epsilon$

$$R_{\Delta x}(\tau;\epsilon) = E[\Delta x(t;\epsilon)\,\Delta x(t-\tau;\epsilon)] \approx \sigma_H^2(2H-1)|\tau|^{2H-2}.$$

Consequently, for $1/2 < H < 1$, fractional Gaussian noise exhibits long-term dependence, i.e., persistent correlation, while for $0 < H < 1/2$, it exhibits persistent anticorrelation. For $H = 1/2$, fractional Gaussian noise specializes to ordinary stationary white Gaussian noise, for which there is, of course, no correlation.

From the fractional Brownian motion framework one can derive a number of other important properties of $1/f$ processes as well, among which are its fractal characteristics. Specifically, sample functions of fractional Brownian motions whose self-similarity parameters lie in the range

$0 < H < 1$ (i.e., $1 < \gamma < 3$) have a fractal (Hausdorff–Besicovitch) dimension given by [27]

$$D = 2 - H$$

that gives a quantitative measure of their roughness.

Perhaps the principal limitation of the fractional Brownian motion framework is that it does not provide useful models for $1/f$ processes corresponding to $\gamma \leq -1, \gamma \geq 3$, and perhaps the most important and ubiquitous case, $\gamma = 1$. Indeed, for $\gamma = 3\,(H = 1)$, fractional Brownian motion as defined by (23) degenerates to a process whose sample paths are all lines through the origin, viz.,

$$x(t) = |t|x(1)$$

while for $\gamma = 1\,(H = 0)$, fractional Brownian motion degenerates to the trivial process

$$x(t) \equiv 0.$$

More generally, choosing $H < 0$ in (23) leads to processes that are not mean-square-continuous, while choosing $H > 1$ in (23) leads to processes whose increments are not stationary [3], [33].

B. A Frequency-Based Characterization of $1/f$ Processes

In this section, we introduce a potentially more general mathematical characterization of $1/f$ processes than is provided by the fractional Brownian motion framework, and one which is perhaps closer in spirit to the empirical characterization through which they were introduced. In fact, we shall use this characterization to *define* $1/f$ processes in a sufficiently precise sense for the purposes of this paper.

The basic notion is that $1/f$ processes are those statistically self-similar random processes that appear stationary when viewed through ideal bandpass filters. Given that measurements of spectra for physical processes can only be obtained over a range of frequencies governed by data length and resolution limitations suggests that this is a rather natural means for distinguishing $1/f$ processes from other statistically self-similar processes. More precisely, we choose the following definition.

Definition 1 A wide-sense statistically self-similar zero-mean random process $x(t)$ shall be said to be a $1/f$ process if there exist ω_0 and ω_1 satisfying $0 < \omega_0 < \omega_1 < \infty$ such that when $x(t)$ is filtered by an ideal bandpass filter with frequency response

$$B_1(\omega) = \begin{cases} 1, & \omega_0 < |\omega| \leq \omega_1 \\ 0, & \textit{otherwise} \end{cases} \tag{28}$$

the resulting process $y_1(t)$ is wide-sense stationary and has finite variance.

While choosing an ideal bandpass filter in this definition may not be critical—that is, it might suffice to choose any filter whose frequency response $B(\omega)$ has sufficient decay as $\omega \to 0$ and $\omega \to 0$—the use of ideal filters is rather convenient. Indeed, the fundamental appeal of Definition 1 as a characterization for $1/f$ processes is its basis in the frequency domain. This allows the well-established tools of Fourier analysis to be extended to this important class of generally nonstationary processes, and, in turn, the derivation of a number of properties of $1/f$ processes in a highly straightforward manner.

In accordance with this definition it suffices to find just one bandpass filter through which the process is stationary. In fact, the following theorem establishes that if a statistically self-similar process is stationary when filtered by one bandpass filter, then it is stationary when filtered by *any* bandpass filter. The theorem also justifies the terminology "$1/f$ process" in Definition 1, and provides a natural interpretation of the spectrum (14). A detailed but straightforward proof of this result is provided in Appendix II-A.

Theorem 1 A $1/f$ process $x(t)$, when filtered by an ideal bandpass filter with frequency response

$$B(\omega) = \begin{cases} 1, & \omega_L < |\omega| \leq \omega_U \\ 0, & \textit{otherwise} \end{cases} \tag{29}$$

for any $0 < \omega_L < \omega_U < \infty$, yields a wide-sense stationary random process $y(t)$ with finite variance and having power spectrum, for some $\sigma_x^2 > 0$

$$S_y(\omega) = \begin{cases} \sigma_x^2/|\omega|^\gamma, & \omega_L < |\omega| \leq \omega_U \\ 0, & \textit{otherwise} \end{cases} \tag{30}$$

where the spectral exponent γ is related to the self-similarity parameter H according to $\gamma = 2H + 1$.

This characterization not only captures the behavior of many physical $1/f$ processes, but also encompasses some important existing mathematical models for $1/f$ behavior. As an example, we present the following theorem, whose proof is provided in Appendix II-B.

Theorem 2 Fractional Brownian motions corresponding to $0 < H < 1$ and the associated fractional Gaussian noises are $1/f$ processes in the sense of Definition 1.

In essence, this theorem states that self-similar random processes that appear stationary when viewed through a differentiator, i.e., fractional Brownian motions, also appear stationary when viewed through any ideal bandpass filter. Furthermore, we have a trivial corollary that the special case corresponding to the Wiener process and its derivative, stationary white Gaussian noise, are also $1/f$ processes. However, the Barnes–Allan process with its pronounced time origin is not a $1/f$ process in the sense of Definition 1.

A number of interesting technical questions are raised by the characterization of $1/f$ processes presented here, many of which are as yet unexplored. For instance, how broad a class of processes this definition admits remains an open question. Indeed, it is not known whether it is actually possible to construct nontrivial Gaussian processes that satisfy Definition 1 for values of H outside $0 < H < 1$. Similarly, it is not known whether there other Gaussian processes besides fractional Brownian motion that satisfy the definition for $0 < H < 1$. Nevertheless, results relevant to these and other related questions has been recently reported by Ramanathan and Zeitouni [34].

IV. Wavelet-Based Models for $1/f$ Processes

In this section, we present the role of orthonormal wavelet basis expansions as Karhunen–Loève-like expansions for $1/f$ processes. That is, we show that such wavelet expansions in terms of uncorrelated random variables constitute good models for $1/f$ behavior. Karhunen–Loève expansions have, in general, proven enormously useful in development and interpretation of classical detection and estimation theory [35]. Consequently, the results we discuss in this section have some important implications which we explore later. In the meantime, we identify two categories of results: synthesis results and analysis results.

A. Synthesis

First, we demonstrate that nearly-$1/f$ behavior may be generated from orthonormal wavelet basis expansions in terms of collections of uncorrelated wavelet coefficients. In particular we present the following theorem, an earlier version of which appears in [36], and whose proof is provided in Appendix II-C.

Theorem 3 Consider any orthonormal wavelet basis with Rth-order regularity for some $R \geq 1$. Then the random process constructed via the expansion

$$x(t) = \sum_m \sum_n x_n^m \psi_n^m(t) \tag{31}$$

where the x_n^m are a collection of mutually uncorrelated, zero-mean random variables with variances

$$\operatorname{Var} x_n^m = \sigma^2 2^{-\gamma m}$$

for some parameter $0 < \gamma < 2R$, has a time-averaged spectrum

$$S_x(\omega) = \sigma^2 \sum_m 2^{-\gamma m} |\Psi(2^{-m}\omega)|^2 \tag{32}$$

that is nearly $1/f$, i.e.,

$$\frac{\sigma_L^2}{|\omega|^\gamma} \leq S_x(\omega) \leq \frac{\sigma_U^2}{|\omega|^\gamma} \tag{33}$$

for some $0 < \sigma_L^2 \leq \sigma_U^2 < \infty$, and has octave-spaced ripple, i.e., for any integer k

$$|\omega|^\gamma S_x(\omega) = |2^k\omega|^\gamma S_x(2^k\omega). \tag{34}$$

Several remarks concerning this theorem are appropriate. First, we emphasize that the nearly $1/f$ spectrum (32), an example of which is depicted in Fig. 5, is to be interpreted in the same manner that (14) is for true $1/f$ processes. That is, if $x(t)$ is filtered by an ideal bandpass filter with frequency response of the form (29), the output of the filter will have finite power and correspond to a spectrum of the form (32) over the passband $\omega_L < |\omega| \leq \omega_U$. Furthermore, we also emphasize that this spectrum is a *time-averaged* one. Indeed, the output of such a bandpass filter will *not*, in general, be stationary in any sense as a consequence of the discrete nature of the synthesis.

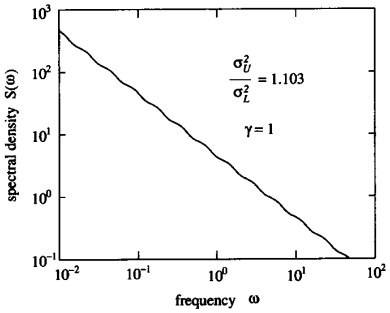

Fig. 5. The time-averaged spectrum of a $1/f$-type process synthesized from the first-order Battle–Lemarie orthonormal wavelet basis.

In general, when the orthonormal wavelet decomposition is viewed in terms of a generalized octave-band filter bank, the result of this theorem is intuitively reasonable. In fact, for the case of the ideal bandpass wavelet basis, it can be inferred from geometric arguments that the tightest bounding constants are

$$\sigma_L^2 = \sigma^2 \pi^\gamma$$
$$\sigma_U^2 = \sigma^2 (2\pi)^\gamma.$$

Interestingly, Theorem 3 also leads to a special interpretation of the case $\gamma = 1$, arguably the most prevalent of the $1/f$ -type processes. In particular, the choice of the variance progression

$$\operatorname{Var} x_n^m = \sigma^2 2^{-m}$$

corresponds to distributing power equally among the detail signals at all resolution scales, since we have for each m

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} P_m(\omega) |\Psi(2^{-m}\omega)|^2 \, d\omega = 1. \tag{35}$$

Not surprisingly, a much stronger theorem holds for the case $\gamma = 0$. Here we have a particular instance of an orthonormal basis expansion in terms of uncorrelated random variables with identical variances, which always generates a stationary white-noise process whose spectral density is the variance of the coefficients. Hence, for $\gamma = 0$ we have

$$S_x(\omega) = \sigma^2 = \sigma^2 \sum_m |\Psi(2^{-m})|^2$$

where the last equality is a consequence of the wavelet identity [37]

$$\sum_m |\Psi(2^{-m}\omega)|^2 = 1. \tag{36}$$

The wavelet basis constraints in Theorem 3 have a number of practical consequences. To generate $1/f$-like behavior for $0 < \gamma < 2$, it suffices to use a wavelet basis for which the corresponding multiresolution analysis is at least *regular*. Virtually any practical wavelet basis satisfies

this condition, even the Haar basis. However, to generate $1/f$-like behavior for $\gamma > 2$, higher regularity ($R > 1$) is required, as can be verified experimentally. For instance, when a Haar-based synthesis ($R = 1$) of a $1/f$ process with $\gamma = 5$ is attempted, the resulting sample functions exhibit abrupt discontinuities. More generally, without sufficient regularity, the characteristics of the basis functions manifest themselves in the sample functions generated by the expansion. However, using more regularity than required by the theorem would appear to be of no advantage.

As a final remark, we mention that Theorem 3 may, in principle, be extended to $\gamma < 0$ provided the wavelet basis used in the synthesis has a sufficient number of vanishing moments. This can be deduced from the proof in Appendix II-C. Typically, however, this case is of relatively little interest since few, if any, physical $1/f$ processes are associated with negative γ.

B. Analysis

This section presents a collection of complementary results to suggest that wavelet bases are equally useful in the analysis of $1/f$ processes. In particular, we provide both theoretical and empirical evidence suggesting that when $1/f$ processes are expanded in terms of orthonormal wavelet bases, the resulting wavelet coefficients are typically rather weakly correlated, particularly in contrast to the rather strong correlation present in the original process.

The following theorem provides a convenient expression for the correlation between arbitrary wavelet coefficients of a $1/f$ process. A proof is outlined in Appendix II-D.

Theorem 4 Let $x(t)$ be a $1/f$ process whose spectral parameters, in the sense of Theorem 1, are σ_x^2 and γ. Furthermore, let the x_n^m be projections of $x(t)$ onto an orthonormal wavelet basis corresponding to a wavelet $\psi(t)$ with R vanishing moments. Then provided $0 < \gamma < 2R$, the correlation between an arbitrary pair of such coefficients x_n^m and $x_{n'}^{m'}$ is given by

$$E[x_n^m x_{n'}^{m'}] = \frac{2^{-(m+m')/2}}{2\pi} \cdot \int_{-\infty}^{\infty} \frac{\sigma_x^2}{|\omega|^\gamma} \Psi(2^{-m}\omega)\, \Psi^*(2^{-m'}\omega) \cdot e^{-j(n2^{-m} - n'2^{-m'})\omega}\, d\omega. \quad (37)$$

Several properties of the second-order statistics of wavelet coefficients of $1/f$ processes may be readily derived from this theorem. For instance, an immediate consequence is that the variance of each x_n^m is of the form

$$\operatorname{Var} x_n^m = \sigma^2 2^{-\gamma m}$$

where

$$\sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sigma_x^2}{|\omega|^\gamma} |\Psi(\omega)|^2\, d\omega$$

reminiscent of—and consistent with—our synthesis result. Similarly, defining

$$\rho_{n,n'}^{m,m'} \triangleq \frac{E[x_n^m x_{n'}^{m'}]}{\sqrt{(\operatorname{Var} x_n^m)(\operatorname{Var} x_{n'}^{m'})}} \quad (38)$$

as the *normalized* wavelet correlation, a second consequence is that the wavelet coefficients are wide-sense-stationary at each scale, i.e., for a fixed scale m, $\rho_{n,n'}^{m,m}$ is a function only of $n - n'$. Indeed, specializing (37) to the case $m' = m$ yields

$$\rho_{n,n'}^{m,m} = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} \frac{\sigma_x^2}{|\omega|^\gamma} |\Psi(\omega)|^2 e^{-j(n-n')\omega}\, d\omega. \quad (39)$$

These results can also be established specifically in the context of the fractional Brownian motion framework as shown by Flandrin [38].

Normalized wavelet coefficients which correspond to synchronous time instants possess a kind of stationarity *across* scales as well. From the filter bank interpretation of wavelet analysis in which the output of the mth filter is sampled at rate $t = 2^{-m}n$ for $n = \cdots, -1, 0, 1, 2, \cdots$, we observe that a pair of wavelet coefficients x_n^m and $x_{n'}^{m'}$ at distinct scales m and m' correspond to synchronous time instants precisely when

$$2^{-m}n = 2^{-m'}n'. \quad (40)$$

Our stationarity result in this case is then that the normalized correlation among time-synchronous wavelet coefficients corresponding to scales m and m' is a function only of $m - m'$. More precisely, specializing (37) yields

$$\rho_{n,n'}^{m,m'} = \frac{1}{2\pi\sigma^2} 2^{-(m-m')/2} \cdot \int_{-\infty}^{\infty} \frac{\sigma_x^2}{|\omega|^\gamma} \Psi(2^{-(m-m')}\omega)\, \Psi(\omega)\, d\omega \quad (41)$$

whenever (40) holds. This result, too, can be shown in the context of fractional Brownian motion, as shown by Flandrin [39].

While the stationarity results provide some insight into the correlation structure among wavelet coefficients, additional insight is obtained by examining the magnitude of the correlation among wavelet coefficients both along a scale and across scales. The following theorem [37] identifies some asymptotic properties of the intercoefficient correlation. A proof is provided in Appendix II-E. A version of this theorem has also been established in the case of fractional Brownian motion by Tewfik and Kim [40].

Theorem 5 Consider an orthonormal wavelet basis such that $\psi(t)$ has R vanishing moments, i.e.,

$$\Psi^{(r)}(\omega) = 0, \qquad r = 0, 1, \cdots, R-1 \quad (42)$$

for some integer $R \geq 1$. Then provided $0 < \gamma < 2R$, the wavelet coefficients obtained by projecting a $1/f$ process onto this basis have a correlation whose magnitude decays

according to[3]

$$|\rho_{n,n'}^{m,m'}| \sim \mathcal{O}(|2^{-m}n - 2^{-m'}n'|^{-\lceil 2R-\gamma \rceil}) \qquad (43)$$

as

$$|2^{-m}n - 2^{-m'}n'| \to \infty.$$

While this theorem makes an interesting statement about the *relative* correlation among some wavelet coefficients well-separated in (m, n)-space, we must avoid inferring some stronger statements. First, it says nothing about the correlation among time-synchronous wavelet coefficients (i.e., those satisfying (40)), regardless of how well separated they are. Furthermore, while plausible, the theorem itself does not assert that choosing an analysis wavelet with a larger number of vanishing moments can reduce the correlation among wavelet coefficients in the analysis of $1/f$ processes. Likewise, the theorem does not actually validate the reasonable hypothesis that choosing a wavelet with an insufficient number of vanishing moments will lead to strong correlation among the wavelet coefficients of $1/f$ processes. In fact, the theorem identifies neither a range of m, m', n, n' over which (43) holds, nor a leading multiplicative constant in (43). Consequently, this precludes us from inferring anything about the *absolute* correlation between *any* particular pair of coefficients.

For the case of the ideal bandpass wavelet basis, however, we may obtain some more useful bounds on the correlation among wavelet coefficients. In this case, the basis functions corresponding to distinct scales have nonoverlapping frequency support. Hence, carefully exploiting the stationarity properties of $1/f$ processes developed in Theorem 1, we may conclude that the wavelet coefficients corresponding to distinct scales are uncorrelated. However, at a given scale the correlation at integral lag $l \geq 0$ is nonzero and may be expressed in the form

$$\rho_{n,n-l}^{m,m} = \frac{\sigma_x^2}{\pi\sigma^2} \int_{\pi}^{2\pi} \omega^{-\gamma} \cos(\omega l)\, d\omega \qquad (44)$$

with

$$\frac{\sigma^2}{\sigma_x^2} = \begin{cases} (2^{1-\gamma} - 1)/(\pi^{\gamma}(1-\gamma)), & \gamma \neq 1 \\ (\ln 2)/(\pi), & \gamma = 1. \end{cases} \qquad (45)$$

While (44) cannot be evaluated in closed form, integrating by parts twice and using the triangle inequality gives the useful closed-form bound

$$|\rho_{n,n-l}^{m,m}| \leq \frac{\sigma_x^2}{\sigma^2} \frac{\gamma}{l^2\pi^{2+\gamma}} \left\{ 1 + \frac{1}{2^{1+\gamma}} + \frac{1+\gamma}{l\pi} \left[1 - \frac{1}{2^{2+\gamma}} \right] \right\} \qquad (46)$$

valid for $\gamma \geq 0$ and integer-valued $l \geq 1$.

In Fig. 6, we plot the exact magnitude of the normalized correlation (44) obtained by numerical integration as a function of lag l together with the bound (46). Note that correlation among wavelet coefficients is extremely small: adjacent coefficients have a correlation coefficient of less than 15%, and more widely separated coefficients have a correlation coefficient less than 3%.

[3]The ceiling function $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

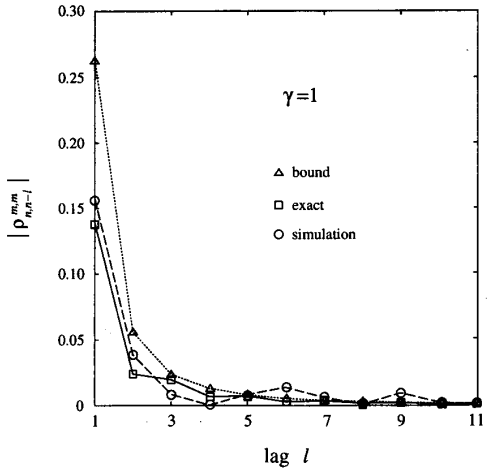

Fig. 6. Along-scale correlation between wavelet coefficients for a $1/f$ process for which $\gamma = 1$. The squares □ indicate a numerical estimate of the exact magnitude of the normalized correlation between wavelet coefficients as a function of the lag l between them. The ideal bandpass wavelet was assumed in the analysis. The triangles △ indicate the corresponding values of the closed-form bound obtained in the text. The circles ○ show the average sample-correlation as computed from a projections of synthesized $1/f$ processes onto the fifth-order Daubechies wavelet basis.

On the same plot we superimpose the average along-scale sample correlation between wavelet coefficients obtained from a $1/f$ process synthesized using the method of Corsini and Saletti [11] based on the extended-ARMA model of Keshner [1]. In this simulation, a 65 536-sample segment of a $1/f$ process was generated for $\gamma = 1$ and analyzed using Daubechies fifth-order wavelet basis. Here, the sample-correlation function of the coefficients at each scale was computed, and averaged appropriately with the sample-correlation functions at the other scales. That the experimental result so closely matches the exact result for the bandlimited basis suggests that our analysis result for the bandlimited basis may, in fact, be more broadly applicable. However, as yet no theoretical statement to this effect has been established.

As a final remark, the results of this section can be interpreted in the context of spectral analysis. In particular, in light of the octave-band filter bank interpretation of wavelet bases we may view wavelet-based analysis, in some sense, as spectral analysis on a logarithmic frequency scale. The results of this section, then, together with a consideration of the spectral characteristics of $1/f$ processes viewed through bandpass filters, suggest that it is this kind of spectral analysis that is, broadly speaking, best suited to $1/f$-type behavior.

C. Examples

In this section, we demonstrate, using wavelet-based analysis, two instances of time series that would appear to be well-modeled as $1/f$ processes. The first is an example

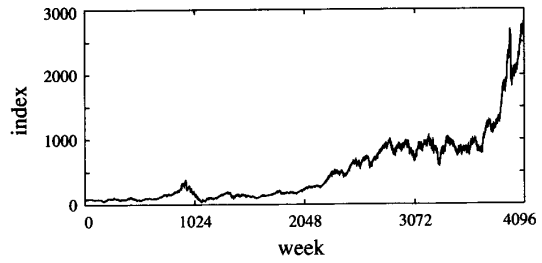

Fig. 7. Weekly Dow Jones Industrial Average data, to present.

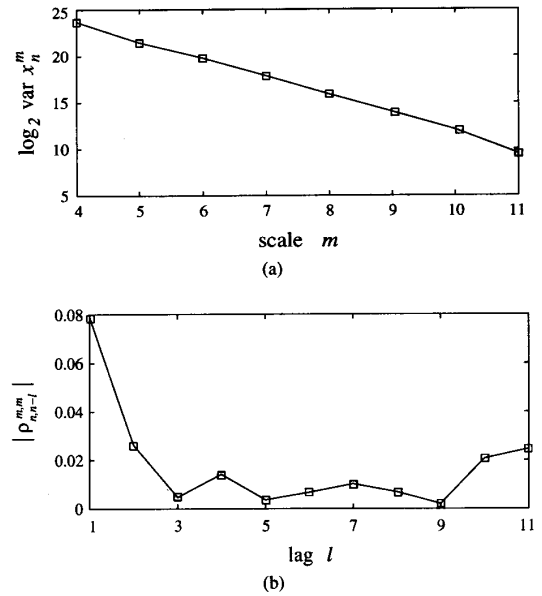

Fig. 8. Wavelet-based analysis of weekly Dow Jones Industrial Average data. The time series is analyzed using a fifth-order Daubechies wavelet basis. (a) Scale-to-scale wavelet coefficient sample-variance progression. (b) Average magnitude of the normalized along-scale sample correlation between wavelet coefficients.

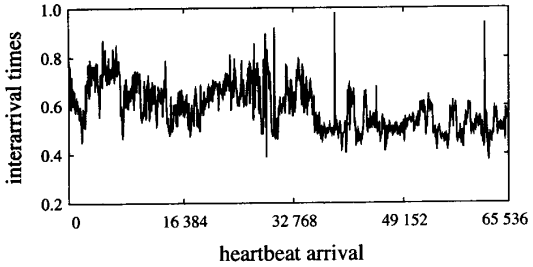

Fig. 9. Heartbeat interarrival times for a healthy human patient.

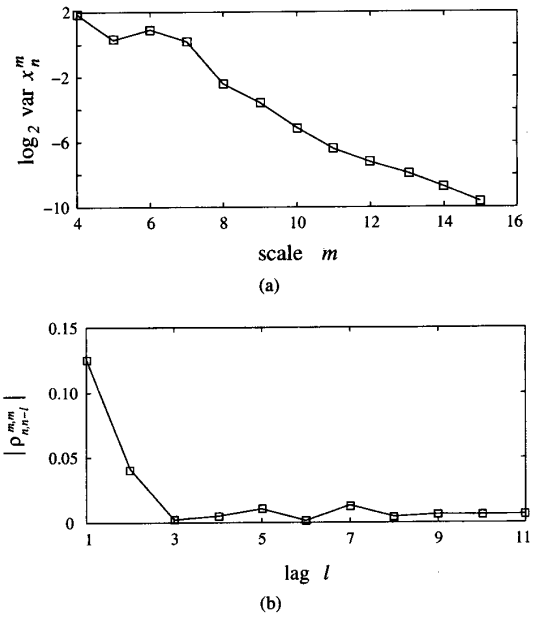

Fig. 10. Wavelet-based analysis of the heartbeat interarrival times for a health patient. The time series is analyzed using a fifth-order Daubechies wavelet basis. (a) Scale-to-scale wavelet coefficient sample-variance progression. (b) Average magnitude of the normalized along-scale sample correlation between wavelet coefficients.

involving economic data, specifically the 80 years of raw weekly Dow Jones Industrial Average data depicted in Fig. 7.

As shown in Fig. 8(a), the sample variances of wavelet coefficients extracted from this data obey a geometric scale-to-scale progression consistent with a $1/f$ process of $\gamma \approx 2$. In Fig. 8(b), we see that the average along-scale sample correlation among wavelet coefficients is rather weak. Since adjacent coefficients have a correlation of less than 8%, and more widely separated coefficients have a correlation of less than 3%, it would appear reasonable to neglect the intercoefficient correlation in the analysis of such data.

A second example involves physiological data, specifically the record of healthy human heart beat interarrival times depicted in Fig. 9. For these data, which correspond to approximately 11 h of continuously acquired data, the quantization levels of the interarrival times are spaced 4 ms apart. As shown in Fig. 10(a), the sample variances of wavelet coefficients extracted from these data also generally follow a geometric scale-to-scale progression consistent with a $1/f$ process of $\gamma \approx 1$. When viewing this plot, keep in mind that the sample-variance measurements are progressively less reliable estimates of the true variances at coarser scales (smaller m). In Fig. 10(b), we show the weak average along-scale sample correlation between wavelet coefficients. In this case, coefficients separated by lags of two or more have less than 2% correlation, again suggesting that such intercoefficient correlation may be neglected in any wavelet-based analysis.

V. Signal Processing with $1/f$ Processes

Having presented both theoretical and empirical results to suggest that the orthonormal wavelet transform is an apparently useful and convenient tool in the synthesis and analysis of $1/f$-type processes, we now explore the equally important role of the wavelet transform in *processing* such signals.

There are a great many signal processing problems that arise which involve $1/f$ processes either as the signals of interest, such as in the case of economic data, or as sources of noise or other interference, such as is the case in optical systems. From a wealth of applications there arises need for algorithms for signal detection, identification, classification, restoration, and enhancement. In this section, we demonstrate the wavelet expansion's role in solving some of the fundamental problems of optimal detection and estimation with $1/f$ processes that form the basis of these algorithms.

In the development of signal processing algorithms for $1/f$ data, robustness is of tremendous importance. This is because even when the $1/f$ signal is of primary interest, the observed data are invariably subject to various forms of distortion. Typically, they will be time-limited, resolution-limited, and corrupted by broadband noise.[4] By the nature of the $1/f$ spectrum, there is a large amount of high-frequency information in the data, much of which lies below the broadband noise floor. Consequently, algorithms for processing $1/f$ data that do not take into account such noise have a tendency to rely too strongly on high-frequency information. The result is often intolerable noise sensitivity. For this reason, we pay particular attention to robustness issues in the problems we consider in this section.

As in earlier sections of the paper, the Corsini–Saletti algorithm [11] based on Keshner's extended-ARMA model [1] is used to synthesis $1/f$ data in simulations. Because this synthesis is fundamentally different from a wavelet-based synthesis, such simulations play an important role in verifying the robustness of the wavelet-based algorithms with respect to modeling error. Additionally, in these simulations we use the Daubechies' finite-extent wavelet basis with fifth-order regularity for which the corresponding conjugate quadrature filters have ten nonzero coefficients. In accordance with the theorems of Section IV, this practical basis has more than enough vanishing moments to accommodate spectral parameters in our principal range of interest, $0 < \gamma < 2$.

All the algorithms of this section exploit the property that the coefficients of suitable wavelet expansions of Gaussian $1/f$ processes can be modeled as collections of mutually independent random variables obeying a particular geometric scale-to-scale variance progression. This not only makes the derivation and analysis of the algorithm mathematically highly tractable, but also leads to some computationally highly efficient structures for implementing the resulting algorithms.

As an illustration, consider the linear reversible whitening filters for $1/f$ processes. For a $1/f$ process $x(t)$ with parameter γ, the wavelet coefficients x_n^m derived out of such a process may be expressed in the form

$$x_n^m = [\sigma 2^{-\gamma m/2}] v_n^m$$

where the v_n^m are then zero-mean, unit-variance, uncorrelated random variables. Because the wavelet coefficients v_n^m thereby correspond to a process $v(t)$ that is stationary and white, the associated whitening filter for a $1/f$ process $x(t)$ involves computing the wavelet coefficients of the input, scaling those coefficients by $[\sigma 2^{-\gamma/2}]^{-1}$, and using the resulting coefficients in a wavelet expansion to generate $v(t)$. This linear filter has kernel[5]

$$\kappa_\omega^0(t,\tau) = \sum_m \sum_n \psi_n^m(t) \frac{1}{\sigma \beta^{-m/2}} \psi_n^m(\tau).$$

Furthermore, the inverse of this whitening filter, i.e., the corresponding $1/f$ synthesis filter, has kernel

$$\kappa_s^0(t,\tau) = \sum_m \sum_n \phi_n^m(t)\, \sigma \beta^{-m/2} \psi_n^m(\tau).$$

Because $1/f$ processes are typically accompanied by white noise, it is, in fact, the whitening filters for processes that are the superposition of $1/f$ and white components that arise most often. In this case, the kernel associated with whitening filter for these $1/f$-plus-white processes takes the form

$$\kappa_w(t,\tau) = \sum_m \sum_n \psi_n^m(t) \frac{1}{\sigma_m} \psi_n^m(\tau) \tag{47}$$

where $\sigma_m > 0$ is defined by

$$\sigma_m^2 = \sigma^2 2^{-\gamma m} + \sigma_w^2 \tag{48}$$

with σ_w^2 the spectral density of the white noise component. The wavelet-based realization of this whitening filter is depicted in Fig. 11.

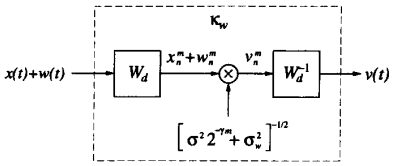

Fig. 11. Canonical form realization of the whitening filter for $1/f$-plus-white processes.

A. Discriminating Between $1/f$ Signals

As our first example, consider the ability of an optimal Bayesian detector to discriminate between Gaussian $1/f$ processes of distinct parameters in a background of sta-

[4] Actually, the coexistence of $1/f$ and white noises in electronic and optical systems is well-documented. In electronic systems, for instance, the predominant noise is $1/f$ noise at frequencies below about 1 kHz, while at higher frequencies, it is white noise in the form of thermal (i.e., Johnson) and shot noise [41].

[5] In our notation, the kernel $k(t,\tau)$ of a linear system defines the response of the system at time t to a unit impulse at time τ. Consequently the response of the system to a suitable input $x(t)$ is expressed as

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, k(t,\tau)\, d\tau.$$

tionary white Gaussian noise. It is convenient to formulate this problem in terms of a binary hypothesis test involving noisy observations $r(t)$. Specifically, we have as our two hypotheses[6]

$$H_0 : r(t) = \dot{x}(t) + w(t) \tag{49a}$$
$$H_1 : r(r) = \ddot{x}(t) + w(t) \tag{49b}$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ are Gaussian $1/f$ processes with distinct parameters, and $w(t)$ is independent white measurement noise.

Rewriting the hypothesis test in terms of the corresponding wavelet coefficients

$$H_0 : r_n^m = \dot{x}_n^m + w_n^m$$
$$H_1 : r_n^m = \ddot{x}_n^m + w_n^m$$

leads to a substantial simplification. Under each hypothesis, the r_n^m are collections of zero-mean statistically independent Gaussian random variables where

$$\text{Var}\{r_n^m|H_0\} = \dot{o}_m^2 = \dot{o}^2 2^{-\dot{\gamma}m} + \sigma_w^2 \tag{50a}$$
$$\text{Var}\{r_n^m|H_1\} = \ddot{o}_m^2 = \ddot{o}^2 2^{-\ddot{\gamma}m} + \sigma_w^2. \tag{50b}$$

Most generally, the collection of available observation coefficients take the form

$$\boldsymbol{r} = \{r_n^m \in \mathcal{R}\} = \{r_n^m, m \in \mathcal{M}, n \in \mathcal{N}(m)\}$$

where

$$\mathcal{M} = \{m_1, m_2, \cdots, m_M\} \tag{51a}$$
$$\mathcal{N}(m) = \{n_1(m), n_2(m), \cdots, n_{N(m)}(m)\} \tag{51b}$$

are arbitrary. However, the special case

$$\mathcal{M} = \{1, 2, \cdots, M\} \tag{52a}$$
$$\mathcal{N}(m) = \{1, 2, \cdots, N_0 2^{m-1}\} \tag{52b}$$

is consistent with the collection of coefficients that would be available in practice from an DWT-based implementation of the wavelet decomposition involving $N = N_0 2^M$ samples of observed data, where N_0 is a constant that depends on the length of the filter $h[n]$.

For a minimum probability of error ($\Pr(\epsilon)$) decision rule under the assumption of equally likely hypotheses, the likelihood ratio test for the problem simplifies to a test of the form

$$l = \frac{1}{2}\sum_{m\in\mathcal{M}} N(m)\left\{\left[\frac{1}{\dot{o}_m^2} - \frac{1}{\ddot{o}_m^2}\right](r_n^m)^2 - \ln\frac{\ddot{o}_m^2}{\dot{o}_m^2}\right\} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{53}$$

where the σ_m^2 are sample variances defined via

$$\hat{\sigma}_m^2 = \frac{1}{N(m)}\sum_{n\in\mathcal{N}(m)} (r_n^m)^2 \tag{54}$$

[6]We use the notation $\dot{}$ and $\ddot{}$ to distinguish the $1/f$ processes and their respective parameters under the two hypotheses. These symbols should not be confused with differentiation operators, for which we have generally reserved the notation $'$ and $''$.

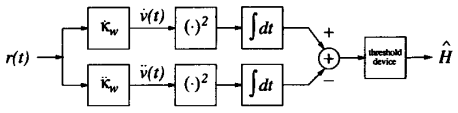

Fig. 12. A canonical form implementation of the optimal receiver for discriminating between $1/f$ models with distinct parameters based on noisy observations $r(t)$.

summarize the aspects of the data required in the discrimination. This rest can be implemented in the canonical form depicted in Fig. 12. In this form, the observations $r(t)$ are processed by $1/f$-plus-white whitening filters corresponding to each hypothesis, for which the respective kernels are

$$\dot{\kappa}_w(t,\tau) = \sum_m\sum_n \psi_n^m(t)\frac{1}{\dot{\sigma}_m}\psi_n^m(\tau)$$
$$\ddot{\kappa}_w(t,\tau) = \sum_m\sum_n \psi_n^m(t)\frac{1}{\ddot{\sigma}_m}\psi_n^m(\tau).$$

Consequently, only one of the residual processes $\dot{v}(t)$ and $\ddot{v}(t)$ is white, depending on which hypothesis is true. To decide between the two hypotheses, the receiver computes the difference in energy in the two residuals and compares it to the appropriate threshold.

Although evaluating the performance of the optimal receiver is rather difficult in general, a useful gauge of performance is obtained by exploiting a popular approach based on the Chernoff bound [35]. Specifically, we can bound the performance by

$$\Pr(\epsilon) \le \tfrac{1}{2}e^{\mu(s_*)} \tag{55}$$

where, for any real parameter s

$$\mu(s) \triangleq \ln E[e^{sl}|H_0]$$
$$= \frac{1}{2}\sum_{m\in\mathcal{M}} N(m)\left\{s\ln\frac{\dot{\sigma}_m^2}{\ddot{\sigma}_m^2} - \ln\left[s\frac{\dot{\sigma}_m^2}{\ddot{\sigma}_m^2} + (1-s)\right]\right\}$$

and where s_* is the parameter value yielding the best possible bound, i.e.,

$$s_* = \arg\min_s \mu(s).$$

Because

$$\mu(0) = \mu(1) = 0$$

and

$$\mu(s) \ge 0$$

the parameter s_* is located by a simple numerical search in the range $0 \le s \le 1$.

Using (55) as estimates of the probability of error performance, we obtain the performance plots of Figs. 13–15. Because the intent of these experiments is to provide a measure of the degree to which $1/f$ processes of different spectral exponents γ are distinguishable, the variance

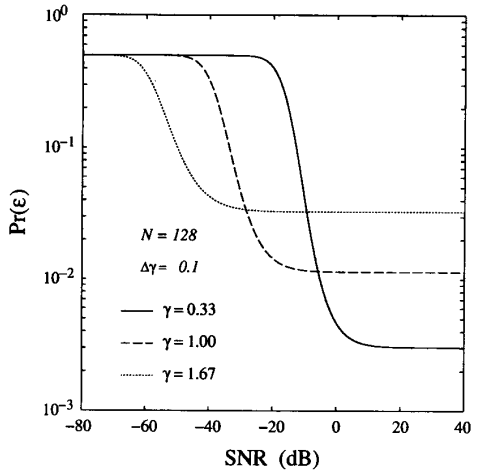

Fig. 13. Optimal discriminator performance as a function of SNR, as estimated via the Chernoff bound.

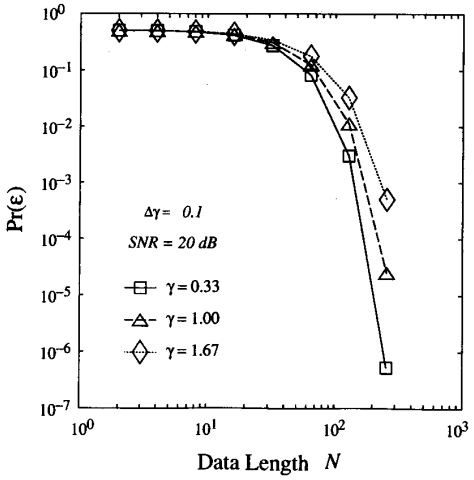

Fig. 14. Optimal discriminator performance as a function of the number of samples N of noisy observations, as estimated by the Chernoff bound. The symbols □, △, and ◇ correspond to actual estimates; the lines are provided as visual aides only in this case.

parameters $\dot{\sigma}^2$ and $\ddot{\sigma}^2$ have been chosen so as to correspond to observations of the same variance under each hypothesis. The results therefore indicate the probability of error performance of an optimal detector in discriminating between two equal-variance $1/f$ processes whose spectral exponents differ by $\Delta\gamma$ based on noisy observations of length N corresponding to a prescribed SNR. In the tests, three different spectral exponent regimes are considered, corresponding to $\gamma = 0.33, \gamma = 1.00$, and $\gamma = 1.67$. While these tests are obviously not substitutes for more comprehensive performance studies involving Monte Carlo simulations with synthetic and real data, the figures provide at least some indication of the behavior and tradeoffs inherent in optimal discrimination of this type.

There is, of course, a broad class of related detection and discrimination problems involving $1/f$ processes that

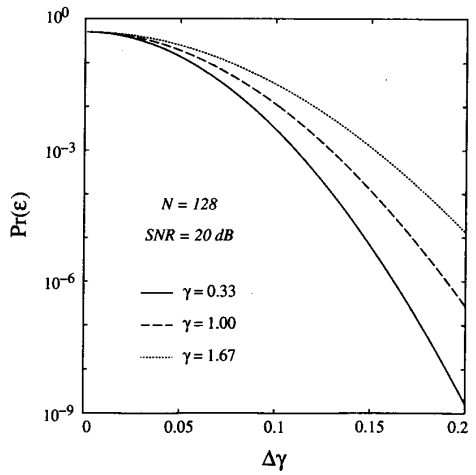

Fig. 15. Optimal discriminator performance as a function of the parameter separation $\Delta\gamma$ between the two hypotheses, as estimated via the Chernoff bound.

may addressed using a related approach. These would include multiple hypothesis testing problems corresponding to discrimination among $1/f$ processes corresponding to several distinct sets of parameters. When one passes from a discrete collection of hypotheses to the continuum, it is natural to re-interpret the discrimination problem as a parameter estimation problem. The wavelet-based framework has proved equally useful in solving problems of robust parameter estimation with noisy $1/f$ signals.

B. Parameter Estimation for $1/f$ Signals

In this section, we describe a class of wavelet-based algorithms developed in [42] for generating Maximum Likelihood (ML) estimates of the parameters of a Gaussian $1/f$ signal from observations corrupted by stationary white Gaussian noise of unknown spectral density. Such parameter estimates, in addition to solving a $1/f$ spectrum estimation problem, play an important role in classifying and detecting signals on the basis of their fractal dimension. In fact, in image processing, where two-dimensional extensions of $1/f$ processes are used to model natural terrain and other patterns and textures [8], [26], fractal dimension can be exploited in distinguishing among various man-made and natural objects.

From observations of the form

$$r(t) = x(t) + w(t) \tag{56}$$

where $x(t)$ is a zero-mean Gaussian $1/f$ process and $w(t)$ is independent zero-mean stationary white Gaussian noise, a collection of wavelet coefficients

$$\boldsymbol{r} = \{r_n^m, m \in \mathcal{M}, n \in \mathcal{N}(m)\}$$

may be extracted. Because we can again model such coefficients as mutually independent zero-mean Gaussian random variables with variance

$$\operatorname{Var} r_n^m = \sigma_m^2 = \sigma^2 2^{-\gamma m} + \sigma_w^2$$

the corresponding log-likelihood function takes the particularly simple form

$$L(\gamma, \sigma^2, \sigma_w^2) = \ln p_{\boldsymbol{r}}(\boldsymbol{r}|\gamma, \sigma^2, \sigma_w^2) = -\frac{1}{2} \sum_{m \in \mathcal{M}} N(m) \left\{ \frac{\hat{\sigma}_m^2}{\sigma_m^2} + \ln (2\pi\sigma_m^2) \right\} \tag{57}$$

where the M sample variances $\hat{\sigma}_m^2$ defined in (54) summarize the aspects of the data required in the estimation.

In the absence of noise ($\sigma_w^2 = 0$), the parameters which maximize the likelihood function are obtained by solving for the roots of a polynomial in $\beta = 2^\gamma$. In particular

$$\hat{\beta}_{\mathrm{ML}} \leftarrow \sum_{m \in \mathcal{M}} \left[\frac{mN(m)}{\sum_{m \in \mathcal{M}} mN(m)} - \frac{N(m)}{\sum_{m \in \mathcal{M}} N(m)} \right] \hat{\sigma}_m^2 \beta^m = 0 \tag{58}$$

from which we get

$$\hat{\gamma}_{\mathrm{ML}} = \log_2 \hat{\beta}_{\mathrm{ML}} \tag{59a}$$

$$\hat{H}_{\mathrm{ML}} = (\hat{\gamma}_{\mathrm{ML}} - 1)/2 \tag{59b}$$

$$\hat{D}_{\mathrm{ML}} = 2 - \hat{H}_{\mathrm{ML}}. \tag{59c}$$

More generally, when $\sigma_w^2 \neq 0$ is unknown, a simple iterative extension of this solution based on an Estimate–Maximize algorithm [43] can be used to obtain the corresponding ML estimates [42].

While the ML parameter estimates for this problem are biased, they are asymptotically efficient and consistent. Cramér–Rao bounds on estimator performance can, in fact, be derived in a equally straightforward manner [42]. For the special case corresponding to $\sigma_w^2 = 0$ and $N = N_0 2^M$ samples of observed data, evaluating the bounds asymptotically in the scenario described by (52) yields

$$\mathrm{var}\, \hat{\gamma}_{\mathrm{ML}} \sim 2/[(\ln 2)^2 N] \tag{60a}$$

$$\mathrm{var}\, (\hat{\sigma}_{\mathrm{ML}}^2/\sigma^2) \sim 2(\log_2 N)^2)/N. \tag{60b}$$

Figure 16 demonstrates the performance, as measured via 64-trial Monte Carlo simulations, of the estimator for γ both as a function of SNR and as a function of data length N. These plots correspond to the case in which all signal and noise parameters are unknown. The percentage rms error in $\hat{\sigma}_{\mathrm{ML}}^2$ exhibits qualitatively similar behavior both as a function of SNR and as a function of data length, though convergence can be particularly slow for this parameter [42].

The parameter estimation algorithm can also be used in parameter tracking applications. As an example, Fig. 17 demonstrates the performance of the algorithms in tracking both the location and size of a step change in the spectral exponent γ of a noise-free $1/f$ signal. The signal in Fig. 17(a) has left and right halves corresponding to $\gamma = 0.90$ and $\gamma = 1.10$, respectively, but identical variances. Local estimates of γ are computed by applying the parameter estimation algorithm to the signal under a sliding window of length 16 384 centered about the point of interest. Naturally, with such a tracking algorithm, a wider estimation window

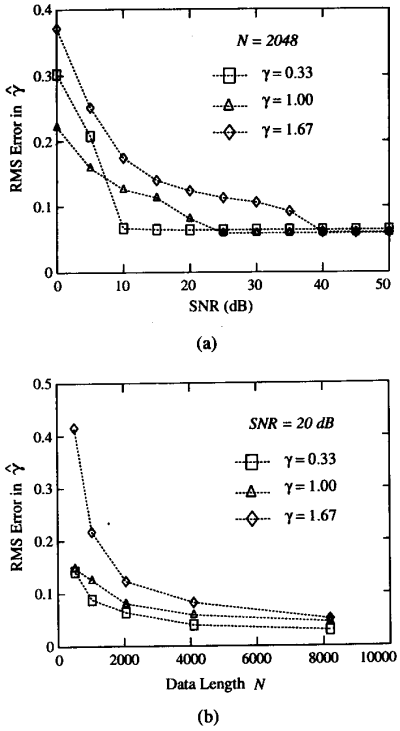

Fig. 16. RMS error in the ML estimate of γ as a function of (a) SNR and (b) data length N. The symbols associated with each γ mark the actual empirical measurements; dashed lines are provided as visual aides only.

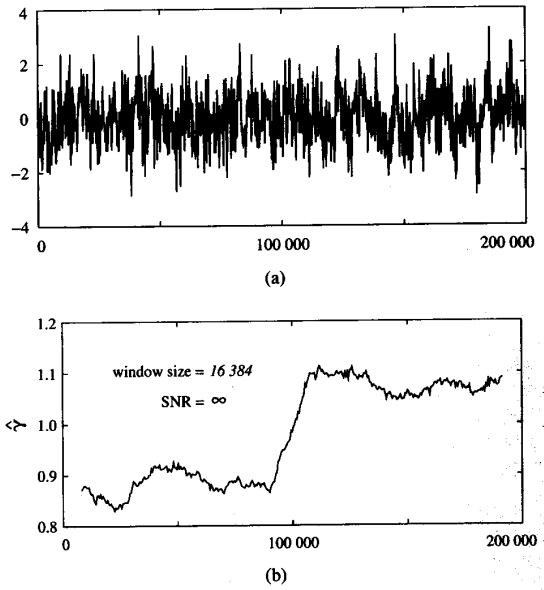

Fig. 17. Tracking the time-varying spectral exponent γ of a noise-free $1/f$-type signal. (a) $1/f$ signal with step change in γ. For the left half of the signal, $\gamma = 0.90$, while for right half, $\gamma = 1.10$. (b) Estimate of γ under a time-limited window.

could be used to reduce the variance in the parameter estimates within each half of the waveform, but at the expense of an increase in the width of the transition zone.

Smoothing is inherent in any parameter estimation algorithm involving signals embedded in noise. In particular, the iterative parameter estimation algorithm mentioned in this section involves repeated smoothing. At each step of the algorithm, the current parameter estimates are first used in a extract the signal from the noise, after which new parameter estimates are obtained from the signal and noise independently. As a final demonstration, we show how the problem of smoothing noisy $1/f$ signals can be addressed in its own right using the wavelet-based framework.

C. Smoothing of $1/f$ Signals

In this section, we consider the problem of extracting a $1/f$ signal from a background of additive stationary white noise. For purposes of illustration, we present Bayesian estimation algorithms developed in [42] that are optimal with respect to a mean-square error criterion for Gaussian processes, keeping in mind that for non-Gaussian processes we still obtain the best possible linear signal estimates.

Again, given the collection of mutually independent wavelet coefficients

$$\boldsymbol{r} = \{r_n^m \in \mathcal{R}\} = \{r_n^m, m \in \mathcal{M}, n \in \mathcal{N}(m)\}$$

extracted from observations $r(t)$ of the form (56), in which the signal and noise parameters $\gamma, \sigma^2, \sigma_w^2$ are all known, the optimal estimate $\hat{x}(t)$ is readily constructed. Specifically, we may express $\hat{x}(t)$ in the form

$$\begin{aligned}\hat{x}(t) &= \sum_{m,n\in\mathcal{R}} E[x_n^m | r_n^m]\, \psi_n^m(t) \\ &= \sum_{m,n\in\mathcal{R}} \left[\frac{\sigma^2\beta^{-m}}{\sigma^2\beta^{-m} + \sigma_w^2}\right] r_n^m \psi_n^m(t). \end{aligned} \quad (61)$$

From (61), we see that the estimator effectively retains coarse scale information where there is good signal-to-noise ratio and discards fine-scale information where the noise predominates. In fact the Wiener filtering that takes place is consistent with the characteristics of the generalized spectra involved. Indeed, because at high frequencies the white-noise spectrum dominates, while at low frequencies the $1/f$ signal spectrum dominates, one can interpret the optimal filter as effectively implementing a form of low-pass filtering.

It is also possible to describe the optimal estimator (61) in canonic form, which consists of two stages as depicted in Fig. 18. In the first stage, the noisy observations $r(t)$ are processed by a whitening filter with kernel $\kappa_w(t,\tau)$ given by (47) to generate an intermediate white "innovations" process $v(t)$ whose wavelet coefficients are

$$v_n^m = \frac{r_n^m}{\sigma_m}.$$

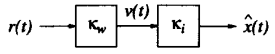

Fig. 18. The canonic whitening-innovations implementation of the optimal filter for extracting a $1/f$ signal $x(t)$ from noisy observations $r(t)$.

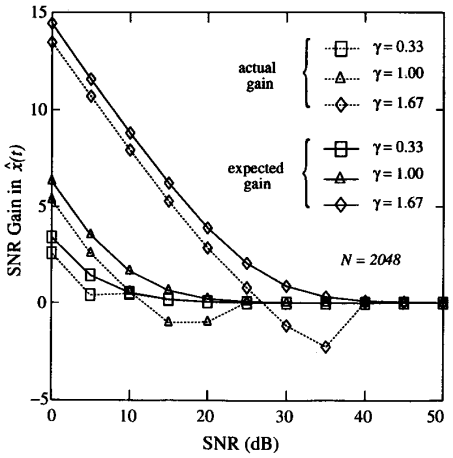

Fig. 19. SNR gain (dB) of the signal estimate as a function of the SNR of the observations. Both the gains predicted by (63) and gains actually obtained are indicated.

In the second stage, $v(t)$ is processed by an innovations filter with kernel

$$\kappa_i(t,\tau) = \sum_m \sum_n \psi_n^m(t) \left[\frac{\sigma^2\beta^{-m}}{\sigma_m}\right] \psi_n^m(\tau) \quad (62)$$

to generate the optimal estimate $\hat{x}(t)$.

Because of the preponderance of low-frequency energy in $1/f$ processes, good performance is achieved by these estimators even at low SNR. The total mean-square estimation error, relative to noise-free reconstruction of $x(t)$ from $x_n^m \in \mathcal{R}$, can be expressed as

$$\begin{aligned}\epsilon &= \sum_{m,n\in\mathcal{R}} E'(\hat{x}_n^m - x_n^m)^2] \\ &= \sum_{m\in\mathcal{M}} N(m) \left[\frac{\sigma^2\beta^{-m}\cdot\sigma_w^2}{\sigma^2\beta^{-m} + \sigma_w^2}\right]. \end{aligned} \quad (63)$$

As is apparent from the results of 64-trial Monte Carlo simulations depicted in Fig. 19, (63) gives a realistic estimate of the smoothing performance to be expected in practice.

As a final demonstration, Fig. 20 shows a segment of a 65 536-sample $1/f$ signal, the same signal embedded in noise, and the optimal signal estimate. In this example, the spectral exponent is $\gamma = 1.67$, and the SNR in the observations of 0 dB. The estimated spectral exponent is $\hat{\gamma}_{\mathrm{ML}} = 1.66$, and the SNR gain of the signal estimate is 13.9 dB.

VI. Concluding Remarks

In this paper, we have presented a number of results on the characterization, representation, and application of

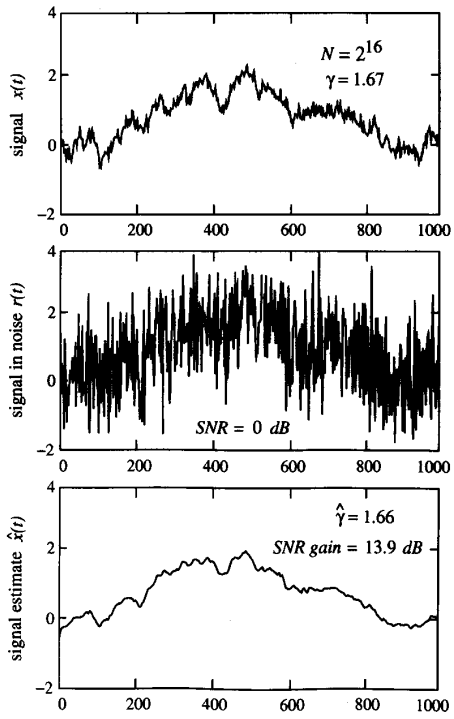

Fig. 20. Optimal smoothing of a noisy $1/f$ signal.

$1/f$ processes for signal processing. Because the wavelet-based approach to the topic is relatively new, there remain, quite naturally, many unanswered questions. A number of these have been identified in the appropriate sections of the paper. As an example, better bounds on the correlation among wavelet coefficients of $1/f$ processes are needed. Furthermore, extensions of representation to nondyadic wavelet bases, i.e., bases in which the dilation factors may be powers of rational constants other than two, are comparatively unexplored.

In another direction, there is interest in more general modeling structures. In many applications, the two-parameter $1/f$ model is overly constrained. More flexible generalizations of these fractal models based on a broader notion of statistical scaling, and the characteristics of their wavelet-based representations, are subjects of ongoing research. Preliminary results in this area are reported in [44].

Section V provided some representative examples of how wavelet-based models can be exploited to develop signal processing algorithms for $1/f$ processes. Many other detection and estimation problems can be similarly addressed, and some of these are described in [37], [45]. However, still others cannot be addressed so directly, an example of which is optimal prediction of $1/f$ processes. The degree to which wavelet-based representations are useful in addressing these problems remains to be explored.

Finally, while this paper has focussed on the use of family of fractal signals in the analysis of natural and man-made phenomena, from an engineering perspective

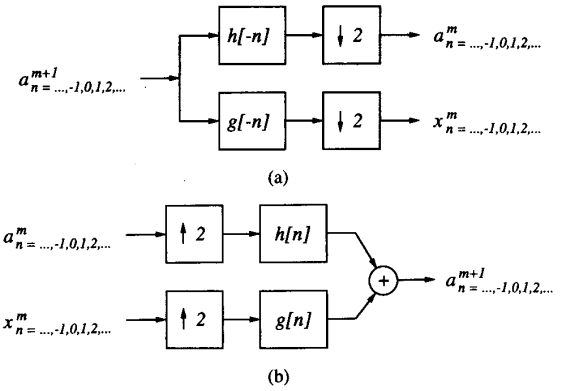

Fig. 21. A single stage of the wavelet transform computation. (a) The analysis step: filter–downsample. (b) The synthesis step: upsample–filter–merge.

the synthesis of fractal signals for applications ranging from communications to remote sensing may ultimately prove to be equally valuable. A particular application of a class of fractal signals to problems of communication over uncertain channels was developed in [46]. Interestingly, this class of fractal signals, which obey a deterministic scale-invariance characterization, have equally important representations in terms of orthonormal wavelet bases. Exploring further applications of fractal signals in these and other areas remains a rich direction for further research.

Appendix I
Computation of the Wavelet Transform

In this Appendix, we summarize a practical, efficient, discrete-time algorithm due to Mallat [16] for computing the orthonormal wavelet transform of a signal. This algorithm exploits a natural correspondence between orthonormal wavelet bases and a class of multirate filter banks.

We first describe the analysis algorithm for recursively computing the coefficients of an orthonormal wavelet expansion of a signal. In this algorithm; the sequence of continuous-time filter-and-sample operations described in Section II-A is reformulated into a single continuous-to-discrete conversion procedure followed by iterative discrete-time processing. Specifically, to extract x_n^m for $m \leq M$ for a signal $x(t)$, the approximation coefficients a_n^{M+1} are obtained via the filter-and-sample procedure of (9), after which the following filter–downsample algorithm:

$$a_n^m = \sum_l h[l-2n]a_l^{m+1} \tag{64a}$$

$$x_n^m = \sum_l g[l-2n]a_l^{m+1} \tag{64b}$$

depicted in Fig. 21(a) is applied recursively to extract coefficients x_n^m at successively coarser scales m.

The synthesis algorithm provides a recursive reconstruction of resolution-limited signal approximations from wavelet coefficients. For this algorithm, the series of conventional modulations (1a) is reformulated into an

iterative discrete-time procedure followed by a single discrete-to-continuous conversion. In particular, given x_n^m for $m \leq M$, the upsample–filter–merge algorithm

$$a_n^{m+1} = \sum_l \{h[n-2l]a_l^m + g[n-2l]x_l^m\} \tag{64c}$$

depicted in Fig. 21(b) is applied recursively until the approximation coefficients a_n^M for some desired scale M are obtained, from which $A_{M+1}x(t)$ is constructed via the modulation (6).

Central to both synthesis and analysis algorithms are the conjugate quadrature filter (CQF) pair $h[n], g[n]$ defined in terms of the wavelet basis by

$$h[n] = \int_{-\infty}^{\infty} \phi_n^1(t)\,\phi_0^0(t)\,dt \tag{65a}$$

$$g[n] = \int_{-\infty}^{\infty} \phi_n^1(t)\,\psi_0^0(t)\,dt. \tag{65b}$$

Typically, $h[n]$ and $g[n]$ have Fourier transforms $H(\omega)$ and $G(\omega)$ that have roughly half-band low-pass and high-pass characteristics, respectively. In fact, for the case of the bandlimited multiresolution signal analysis, $h[n]$ and $g[n]$ are *ideal* low-pass and high-pass filters, specifically

$$H(\omega) = \begin{cases} 1, & 0 < |\omega| \leq \pi/2 \\ 0, & \pi/2 < |\omega| \leq \pi \end{cases}$$

$$G(\omega) = \begin{cases} 0, & 0 < |\omega| \leq \pi/2 \\ 1, & \pi/2 < |\omega| \leq \pi. \end{cases}$$

For filters $h[n]$ and $g[n]$ of length L, a highly efficient implementation of the DWT via an FFT-based algorithm using polyphase forms generally has an asymptotic computational complexity of $\mathcal{O}(\log L)$ per input sample [47]. However, we remark that there are many subtle issues associated with measuring complexity of the algorithm, and this complexity measure can often be misleading [22].

APPENDIX II
PROOFS

A. Proof of Theorem 1

Let ω_0 and ω_1 be constants from Definition 1, and let $\lambda = \omega_1/\omega_0$. We first establish the following useful lemma.

Lemma 1 When a $1/f$ process $x(t)$ is passed through a filter with frequency response

$$B_a(\omega) = \begin{cases} 1, & a\omega_0 < |\omega| \leq a\omega_1 \\ 0, & \textit{otherwise} \end{cases} \tag{66}$$

for any $a > 0$, the output $y_a(t)$ is wide-sense stationary, has finite variance and has an autocorrelation satisfying

$$R_{y_a}(\tau) = E[y_a(t)\,y_a(t-\tau)] = a^{-2H} R_{y_1}(a\tau) \tag{67}$$

for all $a > 0$. Furthermore, for any distinct integers m and k, the processes $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ are jointly wide-sense stationary.

Proof: First, from Definition 1 we have immediately that $y_1(t)$ is wide-sense stationary. More generally, consider the case $a > 0$. Let $b_a(t)$ be the impulse response of the filter with frequency response (66). To establish (67), it suffices to note that $y_a(t)$ has correlation function

$$\begin{aligned} R_{y_a}(t,s) &= E[y_a(t)\,y_a(s)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b_a(t-\alpha)\,b_a(s-\beta) \\ &\quad \cdot R_x(\alpha,\beta)\,d\alpha\,d\beta \\ &= a^{-2H}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b_1(at-\alpha)\,b_1(as-\beta) \\ &\quad \cdot R_x(\alpha,\beta)\,d\alpha\,d\beta \\ &= a^{-2H} R_{y_1}(at,as) \end{aligned} \tag{68}$$

where we have exploited the identities (17b) and

$$b_a(t) = a b_1(at).$$

However, since $y_1(t)$ is wide-sense stationary, the right side of (68) of a function only of $t-s$. Hence, $y_a(t)$ is wide-sense stationary and (67) follows. Furthermore, $y_a(t)$ has variance

$$R_{y_a}(0,0) = a^{-2H} R_{y_1}(0,0) < \infty$$

where the inequality is a consequence of Definition 1. To establish our final result, since $B_{\lambda^m}(\omega)$ and $B_{\lambda^k}(\omega)$ occupy disjoint frequency intervals for $m \neq k$, the spectra of $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ likewise occupy disjoint frequency intervals. Thus $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ are uncorrelated, and, hence, jointly wide-sense sense stationary as well. ■

Proceeding, now to a proof of our main theorem, let us establish that $y(t)$ is wide-sense stationary. Let M_L and M_U be any pair of integers such that

$$\lambda^{M_L}\omega_0 < \omega_L < \omega_U < \lambda^{M_U}\omega_1$$

and consider preceding the filter (29) with a filter whose frequency response is

$$\tilde{B}(\omega) = \begin{cases} 1, & \lambda^{M_L}\omega_0 < |\omega| \leq \lambda^{M_U}\omega_1 \\ 0, & \text{otherwise} \end{cases} \tag{69}$$

since this will not affect the output $y(t)$.

Let $\tilde{y}(t)$ be the output of the filter (69) when driven by $x(t)$. Then since

$$\tilde{B}(\omega) = \sum_{m=M_L}^{M_U} B_{\lambda^m}(\omega)$$

where $B_{\lambda^m}(\omega)$ is as defined in (66) of Lemma 1, we can decompose $\tilde{y}(t)$ according to

$$\tilde{y}(t) = \sum_{m=M_L}^{M_U} y_{\lambda^m}(t) \tag{70}$$

where $y_{\lambda^m}(t)$ is the response of the filter with frequency response $B_{\lambda^m}(\omega)$ to $x(t)$. Since, by Lemma 1, all the terms comprising the summation (70) are jointly wide-sense

stationary, $\tilde{y}(t)$ is wide-sense stationary. Then since $y(t)$ is obtained from $\tilde{y}(t)$ through the filter (29), the stationarity of $y(t)$ is an immediate consequence of the stationarity of $\tilde{y}(t)$ [48].

Let us now derive the form of the spectrum of $y(t)$, i.e., (30). We begin by rewriting (67) of Lemma 1 in the frequency domain as

$$S_{y_1}(a\omega) = a^{-(2H+1)} S_{y_a}(\omega) \tag{71}$$

where $S_{y_a}(\omega)$ is the power spectrum associated with $y_a(t)$. For $1 < a < \lambda$, we observe that $S_{y_1}(\omega)$ and $S_{y_a}(\omega)$ have spectral overlap in the frequency range $a\omega_0 < |\omega| < \omega_1$, and can therefore conclude that the two spectra must be identical in this range. The reasoning is as follows. If we pass either $y_a(t)$ or $y_1(t)$ through the bandpass filter with frequency response

$$B^\dagger(\omega) = \begin{cases} 1, & a\omega_0 < |\omega| \le \omega_1 \\ 0, & \text{otherwise} \end{cases}$$

whose impulse response is $b^\dagger(t)$, the outputs must identical, i.e.,

$$b^\dagger(t) * y_a(t) = b^\dagger(t) * y_1(t) = b^\dagger(t) * x(t).$$

Since $y_a(t)$ and $y_1(t)$ are jointly wide-sense stationary, we then conclude

$$S_{y_a}(\omega)|B^\dagger(\omega)|^2 = S_{y_1}(\omega)|B^\dagger(\omega)|^2$$

whence

$$S_{y_a}(\omega) = S_{y_1}(\omega), \qquad a\omega_0 < |\omega| < \omega_1. \tag{72}$$

Combining (72) with (71) we get

$$S_{y_1}(a\omega) = a^{-(2H+1)} S_{y_1}(\omega), \qquad a\omega_0 < |\omega| < \omega_1 \tag{73}$$

for any $1 < a < \lambda$. Differentiating (73) with respect to a and letting $a \to 1+$, we find that

$$\omega S'_{y_1}(\omega) = -(2H+1) S_{y_1}(\omega), \qquad \omega_0 < \omega < \omega_1$$

and note that all positive, even, regular solutions to this equation are of the form

$$S_{y_1}(\omega) = \sigma_x^2 / |\omega|^\gamma, \qquad \omega_0 < |\omega| \le \omega_1 \tag{74}$$

for some $\sigma_x^2 > 0$ and $\gamma = 2H + 1$. Using (74) with (71) we find, further, that

$$S_{y_{\lambda^m}}(\omega) = \begin{cases} \sigma_x^2 / |\omega|^\gamma, & \lambda^m \omega_0 < |\omega| \le \lambda^m \omega_1 \\ 0, & \text{otherwise.} \end{cases}$$

Via Lemma 1, the $y_{\lambda^m}(t)$ are uncorrelated, so we deduce that $\tilde{y}(t)$ has spectrum

$$\begin{aligned} S_{\tilde{y}}(\omega) &= \sum_{m=M_L}^{M_U} S_{y_{\lambda^m}}(\omega) \\ &= \begin{cases} \sigma_x^2 / |\omega|^\gamma, & \lambda^{M_L} \omega_0 < |\omega| \le M^{M_U} \omega_1 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Finally, since

$$S_y(\omega) = |B(\omega)|^2 S_{\tilde{y}}(\omega)$$

our desired result (30) follows.

■

B. Proof of Theorem 2

To show that a fractional Brownian motion $x(t)$, for $0 < H < 1$, is a $1/f$ process according to Definition 1, it suffices to consider the effect on $x(t)$ of *any* LTI filter with a regular finite-energy impulse response $b(t)$ and frequency response $B(\omega)$ satisfying $B(\omega) = 0$. In particular, since $x(t)$ has correlation given by (24), the output of the filter

$$y(t) = \int_{-\infty}^{\infty} b(t-\tau)\, x(\tau)\, d\tau \tag{75}$$

has autocorrelation

$$\begin{aligned} R_y(t,s) &= E[y(t)\, y(s)] \\ &= \frac{\sigma_H^2}{2} \int_{-\infty}^{\infty} b(v)\, dv \\ &\quad \cdot \int_{-\infty}^{\infty} |t - s + u - v|^{2H} b(u)\, du \end{aligned}$$

as first shown by Flandrin [38]. Since $R_y(t,s)$ is a function only of $t - s$, the process is stationary, and has spectrum

$$S_y(\omega) = |B(\omega)|^2 \cdot \frac{1}{|\omega|^{2H+1}}.$$

When we restrict our attention to the case in which $B(\omega)$ is the ideal bandpass filter (28), we see that $y(t)$ is not only stationary, but has finite variance. This establishes that any fractional Brownian motion $x(t)$ satisfies the definition of a $1/f$ process.

That the generalized derivative, fractional Gaussian noise $x'(t)$, is also a $1/f$ process follows almost immediately. Indeed, when $x'(t)$ is processed by the LTI filter with impulse response $b(t)$ described above, the output is $y'(t)$, the derivative of (75). Since $y(t)$ is stationary, so is $y'(t)$. Moreover, $y'(t)$ has spectrum

$$S_{y'}(\omega) = |B(\omega)|^2 \cdot \frac{1}{|\omega|^{2H'+1}}$$

where H' is as given by (27). Again, when $B(\omega)$ is given by (28), $y'(t)$ is not only stationary, but has finite variance, which is our desired result.

■

C. Proof of Theorem 3

Without loss of generality, let us assume $\sigma^2 = 1$. Next, we define

$$x_M(t) = \sum_{m=-M}^{M} \sum_n x_n^m \psi_n^m(t) \tag{76}$$

as a resolution-limited approximation to $x(t)$ in which information at resolutions coarser than 2^{-M} and finer than 2^M is discarded, so

$$x(t) = \lim_{M \to \infty} x_M(t) = \sum_m \sum_n x_n^m \psi_n^m(t).$$

Since for each m the wavelet coefficient sequence x_n^m is wide-sense-stationary with spectrum $2^{-\gamma m}$, the approximation $x_M(t)$ is cyclostationary [48] with period 2^M, has finite

variance, and has the associated time-averaged spectrum

$$S_M(\omega) = \sum_{m=-M}^{M} 2^{-\gamma m} |\Psi(2^{-m}\omega)|^2. \tag{77}$$

The limiting time-averaged spectrum

$$S_x(\omega) = \lim_{M\to\infty} S_M(\omega)$$

gives the desired spectrum expression (32), and corresponds to the time-averaged spectrum of $x(t)$ as measured at the output of a bandpass filter for each frequency ω in the passband. The desired octave-spaced ripple relation (34) for arbitrary integer k follows immediately from (32).

To establish (33), we begin by noting that, given ω, we can choose m_0 and ω_0 such that $\omega = 2^{m_0}\omega_0$ and $1 \le |\omega_0| < 2$. Hence, using (34) we see

$$S_x(\omega) = 2^{-m_0\gamma} S_x(\omega_0)$$

from which it follows that

$$\left[\inf_{1\le|\omega_0|<2} S_x(\omega_0)\right] \frac{1}{|\omega|^\gamma} \le S_x(\omega) \le \left[\sup_{1\le|\omega_0|<2} S_x(\omega_0)\right] \frac{2^\gamma}{|\omega|^\gamma}.$$

It suffices, therefore to find upper and lower bounds for $S_x(\omega_0)$ on $1 \le |\omega_0| < 2$.

Since $\psi(t)$ is Rth-order regular, $\Psi(\omega)$ decays at least as fast as $1/\omega^R$ as $\omega \to \infty$. This, together with the fact that $\Psi(\omega)$ is bounded according to (4a), implies that

$$|\Psi(\omega)| \le \frac{C}{1+|\omega|^R}$$

for some $C \ge 1$. Using this with (8a) in (32) leads to the upper bound

$$S_x(\omega_0) \le \sum_{m=0}^{\infty} 2^{-\gamma m} + \sum_{m=1}^{\infty} 2^{\gamma m} C^2 2^{-2Rm} < \infty.$$

To establish the lower bound it suffices to show $S_x(\omega) > 0$ for every $1 \le \omega \le 2$, which we establish by contradiction.

Suppose for some $1 \le \omega_0 \le 2$

$$S_x(\omega_0) = \sum_m 2^{-\gamma m} |\Psi(2^{-m}\omega_0)|^2 = 0.$$

Then since all the terms in the sum are nonnegative, this would imply that each term is zero, from which we could conclude

$$\sum_m |\Psi(2^{-m}\omega_0)|^2 = 0.$$

However, this contradicts the wavelet basis identity (36). Hence, we must have that $S(\omega) > 0$ for every $\pi \le \omega_0 \le 2\pi$. The complete theorem follows. ■

D. Proof of Theorem 4

We begin by defining the process $x_K(t)$ as the result of filtering $x(t)$ with the ideal bandpass filter whose frequency response is given by

$$B_K(\omega) = \begin{cases} 1, & 2^{-K} < |\omega| \le 2^K \\ 0, & \text{otherwise} \end{cases}$$

so that

$$\lim_{K\to\infty} x_K(t) = x(t).$$

Then by Theorem 1, $x_K(t)$ is wide-sense stationary and has power

$$S_K(\omega) = \begin{cases} \sigma_x^2/|\omega|^\gamma, & 2^{-K} < |\omega| \le 2K \\ 0, & \text{otherwise.} \end{cases} \tag{78}$$

If we denote its corresponding autocorrelation by

$$R_K(\tau) = E[x_K(t)\, x_K(t-\tau)]$$

and its wavelet coefficients by

$$x_n^m(K) = \int_{-\infty}^{\infty} x_K(t)\, \psi_n^m(t)\, dt$$

the correlation between wavelet coefficients may be expressed

$$\begin{aligned} E[x_n^m(K)\, x_{n'}^{m'}(K)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_n^m(t)\, R_K(t-\tau)\, \psi_{n'}^{m'}(\tau)\, dt\, d\tau \\ &= \int_{-\infty}^{\infty} \psi_n^m(t)[R_K(t) * \psi_{n'}^{m'}(t)]\, dt. \end{aligned} \tag{79}$$

Applying Parseval's theorem and exploiting (78), we may rewrite (79) in the frequency domain as

$$\begin{aligned} E[x_n^m(K)\, x_{n'}^{m'}(K)] &= \frac{2^{-(m+m')/2}}{2\pi} \cdot \left\{ \int_{-2^K}^{-2^{-K}} \frac{\sigma^2}{|\omega|^\gamma} \Psi(2^{-m}\omega)\, \Psi^*(2^{-m'}\omega)\, d\omega \right. \\ &\quad \left. + \int_{2^{-K}}^{2^K} \frac{\sigma_x^2}{|\omega|^\gamma} \Psi(2^{-m}\omega)\, \Psi^*(2^{-m'}\omega)\, d\omega \right\}. \end{aligned} \tag{80}$$

Interchanging limits, we get

$$x_n^m = \lim_{K\to\infty} x_n^m(K)$$

and, in turn,

$$E[x_n^m x_{n'}^{m'}] = \lim_{K\to\infty} E[x_n^m(K)\, x_{n'}^{m'}(K)]. \tag{81}$$

Substituting (80) into (81) yields (37). Since

$$|E[x_n^m x_{n'}^{m'}]|^2 \le \operatorname{var} x_n^m \cdot \operatorname{var} x_{n'}^{m'}$$

and since

$$\operatorname{var} x_n^m = \frac{\sigma^2 \sigma_x^2 2^{-\gamma m}}{\pi} J$$

where

$$J = \int_0^\infty \omega^{-\gamma} |\Psi(\omega)|^2 \, d\omega \tag{82}$$

it suffices to show that (82) converges. Because $\psi(t)$ has R vanishing moments, there exist constants C_0 and C_1 such that

$$|\Psi(\omega)| < C_0 |\omega|^R \tag{83a}$$

$$|\Psi(\omega)| < C_1 |\omega|^{-R}. \tag{83b}$$

Using (83) in (82), we obtain, for $0 < \gamma < 2R$ and $R \geq 1$

$$J = \int_0^1 C_0^2 \omega^{2R-\gamma} \, d\omega + \int_1^\infty C_1^2 \omega^{-2R-\gamma} \, d\omega < \infty$$

as required. ■

E. Proof of Theorem 5

Let us define

$$\Delta = 2^{-m} n - 2^{-m'} n'$$

and

$$\Xi(\omega) = \omega^{-\gamma} \Psi(2^{-m}\omega) \, \Psi^*(2^{-m'}\omega)$$

for $\omega > 0$, so that (38) may be expressed, via (37), as

$$\rho_{n,n'}^{m,m'} = \frac{\sigma_x^2}{\pi \sigma^2} \mathrm{Re}\, I(\Delta) \tag{84}$$

where

$$I(\Delta) = \int_0^\infty \Xi(\omega) e^{-j\Delta\omega} \, d\omega. \tag{85}$$

Thus to establish the desired result, it suffices to show that $I(\Delta)$ has the appropriate decay.

We first note that if $\gamma \geq 2R+1$, then we cannot guarantee that $I(\Delta)$ converges for any Δ. Indeed, since

$$\Xi(\omega) \sim \mathcal{O}(\omega^{2R-\gamma}), \qquad \omega \to 0$$

we see that $I(\Delta)$ is not absolutely integrable. However, provided $\gamma \leq 2R$, $I(\Delta)$ is absolutely integrable, i.e.,

$$\int_0^\infty |\Xi(\omega)| \, d\omega < \infty.$$

In this case, we have, by the Riemann–Lebesgue lemma [30], that

$$I(\Delta) \to 0, \qquad \Delta \to \infty.$$

When $0 < \gamma < 2R$, we may integrate (85) by parts Q times, for some positive integer Q, to obtain

$$I(\Delta) = \frac{1}{(j\Delta)^Q} \int_0^\infty \Xi^{(Q)}(\omega) e^{-j\Delta\omega} \, d\omega + \sum_{q=0}^{Q-1} \frac{1}{(j\Delta)^q} \Big\{ \lim_{\omega \to 0} [\Xi^{(q)}(\omega) e^{-j\Delta\omega}] - \lim_{\omega \to \infty} [\Xi^{(q)}(\omega) e^{-j\Delta\omega}] \Big\}. \tag{86}$$

Due to the vanishing moments of the wavelet we have

$$\Xi^{(q)}(\omega) \sim \mathcal{O}(\omega^{2R-\gamma-q}), \qquad \omega \to 0 \tag{87}$$

while due to the regularity of the wavelet, $\Psi(\omega)$ decays at least as fast as $1/\omega^R$ as $\omega \to \infty$, whence

$$\Xi^{(q)}(\omega) \sim \mathcal{O}(\omega^{-2R-\gamma-q}), \qquad \omega \to \infty. \tag{88}$$

Hence, the limit terms in (86) for which $-2R - \gamma < q < 2R - \gamma$ all vanish.

Moreover, when we substitute $q = Q$, (87) and (88) imply that $\Xi^{(Q)}(\omega)$ is absolutely integrable, i.e.,

$$\int_0^\infty |\Xi^{(Q)}(\omega)| \, d\omega < \infty \tag{89}$$

whenever $-2R - \gamma + 1 < Q < 2R - \gamma + 1$, which implies, again via the Riemann–Lebesgue lemma, that the integral in (86) vanishes asymptotically, i.e.,

$$\int_0^\infty \Xi^{(Q)}(\omega) e^{-j\Delta\omega} \, d\omega \to 0, \qquad \Delta \to \infty. \tag{90}$$

Hence, choosing $Q = \lceil 2R - \gamma \rceil$ in (86) (so $2R - \gamma \leq Q < 2R - \gamma + 1$) allows us to conclude

$$I \sim \mathcal{O}(\Delta^{-\lceil 2R-\gamma \rceil}), \qquad \Delta \to \infty. \tag{91}$$

Substituting (91) into (84) then yields the desired result. ■

ACKNOWLEDGMENT

The author wishes to thank Prof. A. V. Oppenheim, Prof. A. S. Willsky, and Prof. W. M. Siebert, all at MIT, for their valuable input to this work. The author is also grateful to Prof. L. Gould of MIT for providing the Dow Jones data, and to Dr. D. Rigney and his colleagues at Beth Israel Hospital in Boston for providing the heart rate data.

REFERENCES

[1] M. S. Keshner, "$1/f$ noise," *Proc. IEEE*, vol. 70, pp. 212–218, Mar. 1982.

[2] C. M. Van Vliet, Ed., *Noise in Physical Systems and* $1/f$ *Noise*. Singapore: World Scientific, 1987.

[3] B. B. Mandelbrot and H. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, pp. 422–436, Oct. 1968.

[4] J. A. Barnes and D. W. Allan, "A statistical model of flicker noise," *Proc. IEEE*, vol. 54, pp. 176–178, Feb. 1966.

[5] A. van der Ziel, "On the noise spectra of semi-conductor noise and of flicker effect," *Physica*, vol. 16, no. 4, pp. 359–372, 1950.

[6] J. Bernamont, "Fluctuations in the resistance of thin films," *Proc. Phys. Soc.*, vol. 49, pp. 138–139, 1937.

[7] R. J. Barton and V. H. Poor, "Signal detection in fractional Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 34, pp. 943–959, Sept. 1988.

[8] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Medical Imag.*, vol. MI-5, pp. 152–161, Sept. 1986.

[9] M. Deriche and A. H. Tewfik, "Maximum likelihood estimation of the parameters of discrete fractionally differenced Gaussian noise processes," *IEEE Trans. Signal Process.*, vol. 41, no. 10, pp. 2977–2989, Oct. 1993.

[10] L. M. Kaplan and C.-C. J. Kuo, "Fractal estimation from noisy measurements via discrete fractional Gaussian noise (DFGN) and the haar basis," *IEEE Trans. Signal Processing,* submitted for publication, 1992.
[11] G. Corsini and R. Saletti, "Design of a digital $1/f^{\nu}$ noise simulator," in *Noise in Physical Systems and* $1/f$ *Noise,* C. M. Van Vliet, Ed. Singapore: World Scientific, 1987, pp. 82–86.
[12] M. Basseville, A. Benveniste, K. C. Chou, S. A. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory,* vol. 38, pp. 766–784, Mar. 1992.
[13] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1983.
[14] P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proc. IEEE,* vol. 78, pp. 56–93, Jan. 1990.
[15] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Proc. Mag.,* vol. 8, pp. 14–38, Oct. 1991.
[16] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 11, pp. 674–693, July 1989.
[17] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory,* vol. 36, pp. 961–1005, Sept. 1990.
[18] G. Strang, "Wavelets and dilation equations: A brief introduction," *SIAM Rev.,* vol. 31, pp. 614–627, Dec. 1989.
[19] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.,* vol. 41, pp. 909–996, Nov. 1988.
[20] A. J. Jerri, "The Shannon sampling theorem—its various extensions and applications: A tutorial review," *Proc. IEEE,* vol. 65, pp. 1565–1596, Nov. 1977.
[21] S. G. Mallat, "Multiresolution representations and wavelets," Ph.D. dissertation, Univ. of Pennsylvania, Philadelphia, PA, Aug. 1988.
[22] M. Vetterli, "Wavelets and filter banks for discrete-time signal processing," Tech. Rep. CU/CTR/TR 218-91-48, Columbia Univ., Jan. 1991.
[23] I. Daubechies, "Orthonormal bases of compactly supported wavelets II: Variations on a theme," *SIAM J. Math. Anal.,* , vol. 24, pp. 499–519, Mar. 1993.
[24] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing,* vol. 40, pp. 2207–2232, Sept. 1992.
[25] M. S. Keshner, "Renewal process and diffusion models of $1/f$ noise," Ph.D. dissertation, MIT, Cambridge, MA, May 1979.
[26] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-6, pp. 661–674, Nov. 1984.
[27] B. B. Mandelbrot, *The Fractal Geometry of Nature.* San Francisco, CA: Freeman, 1982.
[28] R. F. Voss, "$1/f$ (flicker) noise: A brief review," in *Proc. Ann. Symp. Freq. Contr.,* 1979, pp. 40–46.
[29] A. van der Ziel, "Unified presentation of $1/f$ noise in electronic devices: Fundamental $1/f$ noise sources," *Proc. IEEE,* pp. 233–258, Mar. 1988.
[30] D. C. Champeney, *A Handbook of Fourier Theorems.* Cambridge, England: Cambridge Univ. Press, 1987.
[31] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions.* New York: Dover, 1965.
[32] K. B. Oldham and J. Spanier, *The Fractional Calculus.* New York: Academic Press, 1974.
[33] M. S. Taqqu, "Self-similar processes and related ultraviolet and infrared catastrophes," in *Random Fields: Rigorous Results in Statistical Mechanics and Quantum Field Theory,* J. Fritz, J. L. Lebowitz, and D. Szasz, Eds. Amsterdam, The Netherlands: North-Holland, 1981, pp. 1057–1096.
[34] J. Ramanathan and O. Zeitouni, "On the wavelet transform of fractional Brownian motion," *IEEE Trans. Inform. Theory,* vol. 37, no. 4, pp. 1156–1158, 1991.
[35] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.
[36] G. W. Wornell, "A Karhunen-Loève-like expansion for $1/f$ processes via wavelets," *IEEE Trans. Inform. Theory,* vol. 36, pp. 859–861, July 1990.
[37] ——, "Synthesis, analysis, and processing of fractal signals," RLE Tech. Rep. 566, MIT, Cambridge, MA, Oct. 1991.
[38] P. Flandrin, "On the spectrum of fractional Brownian motions," *IEEE Trans. Inform. Theory,* vol. 35, pp. 197–199, Jan. 1989.
[39] ——, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Inform. Theory,* vol. 38, pp. 910–917, Mar. 1992.
[40] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional Brownian motions," *IEEE Trans. Inform. Theory,* vol. 38, Mar. 1992.
[41] S. D. Senturia and B. D. Wedlock, *Electronic Circuits and Applications.* New York: Wiley, 1975.
[42] G. W. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. Signal Processing,* vol. 40, pp. 611–623, Mar. 1992.
[43] N. M. Laird, A. P. Dempster, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. Roy. Stat. Soc.,* pp. 1–38, Dec. 1977.
[44] W. M. Lam, "Modeling algorithms for a class of fractal signals," S.M. thesis, MIT, Cambridge, MA, Aug. 1992.
[45] G. W. Wornell, "Communication over fractal channels," in *Proc. Int. Conf. Acoust. Speech, Signal Processing,* May 1991.
[46] G. W. Wornell and A. V. Oppenheim, "Wavelet-based representations for a class of self-similar signals with application to fractal modulation," *IEEE Trans. Inform. Theory,* vol. 38, pp. 785–800, Mar. 1992.
[47] M. Vetterli and C. Herley, "Wavelets and filter banks: Relationships and new results," in *Proc. Int. Conf. Acoust. Speech, Signal Processing,* 1990.
[48] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* 2nd ed. New York: McGraw-Hill, 1984.

Gregory W. Wornell (Member, IEEE) was born in Montréal, Que., Canada, in 1962. He received the B.A.Sc. degree (with honors) from the University of British Columbia, Vancouver, Canada, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, all in electrical engineering, in 1985, 1987, and 1991, respectively.

Since 1991 he has been on the faculty at MIT, where he is ITT Career Development Assistant Professor in the Department of Electrical Engineering and Computer Science. During the 1992–1993 academic year, he was on leave at AT&T Bell Laboratories, Murray Hill, NJ. During 1990 he was a Visiting Investigator at the Woods Hole Oceanographic Institution, Woods Hole, MA, and between 1985 and 1989 he held several research staff appointments with MacDonald Dettwiler and Associates, Canada. His current research interests include signal processing, broadband and wireless communications, and applications of fractal geometry and nonlinear dynamical system theory in these areas. He is co-author on a pending patent. From 1985 to 1989 he held a 1967 Science and Engineering Scholarship from the Natural Sciences and Engineering Research Council of Canada, and, in 1991, he received the MIT Goodwin Medal for "conspicuously effective teaching."

Dr. Wornell is a member of Tau Beta Pi and Sigma Xi.