On the usefulness of wavelet-based simulation of fractional Brownian motion*†

Vladas Pipiras
University of North Carolina at Chapel Hill

September 16, 2004

Abstract

We clarify some ways in which wavelet-based synthesis of fractional Brownian motion is used and can be useful. In particular, we examine the choice of an initial scale in the wavelet-based synthesis method, compare it to other methods for simulation of fractional Brownian motion, and discuss connections to strong invariance principles encountered in Probability and Statistics.

1 Introduction

Wavelet-based synthesis is a method to simulate fractional Brownian motion (fBm, in short) by using wavelets. Based on a special wavelet decomposition of fBm established by Sellan (1995) and Meyer, Sellan and Taqqu (1999), it was first implemented to simulate fBm in practice by Abry and Sellan (1996). The wavelet-based synthesis was subsequently clarified by Pipiras (2003*a*).

Despite a theoretical result of Sellan (1995) and Meyer et al. (1999), and a practical implementation of Abry and Sellan (1996) and Pipiras (2003*a*), there remain several open questions related to the wavelet-based synthesis of fBm:

How useful is the wavelet-based synthesis of fBm? How does it compare to other simulation methods? When should it be used?

Our goal is to provide answers to these questions. We should also point out that the focus here is specifically on the wavelet-based synthesis of fBm and not on the wavelet decomposition of fBm or its usefulness. The wavelet decomposition of fBm is useful by its own right, for example, to study path properties of fBm through its wavelet detail coefficients or as an example of an optimal decomposition of fBm (see Ayache and Taqqu (2002)).

The paper is organized as follows. In Section 2, we recall some basic facts behind the theoretical wavelet decomposition of fBm and its implementation in practice, and in Section 3, we introduce the notation used throughout the paper. The wavelet-based synthesis of fBm involves generating

---

* *AMS Subject classification.* Primary 60G18; secondary 41A58, 60F15.

† *Keywords and phrases*: fractional Brownian motion, wavelets, low and high-pass filters, fast wavelet transform, FARIMA time series, simulation.

approximation coefficients at a fine scale (final approximation) from those at a coarser scale (initial approximation) by using the fast wavelet transform. In Section 4, we study what coarsest (initial) scale should be used in the synthesis. We argue that, in the sense to be specified below, it doesn't really matter at which coarsest scale the initial approximation is to be taken. Implications for usefulness of wavelet-based synthesis of fBm are discussed in Section 5. Some guidelines to practical use of wavelet-based synthesis can be found in Section 6. In Section 7, we describe some connections of the wavelet-based synthesis to strong invariance-type principles encountered in Probability and Statistics. In Section 8, we argue that our work is relevant to other contexts, for example, that of the wavelet-based synthesis of the Rosenblatt process of Abry and Pipiras (2003), based on Pipiras (2003*b*).

2 Wavelet decomposition and synthesis of fBm

Fractional Brownian motion (*fBm*, in short) is a zero mean, Gaussian process $\{B_H(t)\}_{t\in\mathbb{R}}$, $H \in (0,1)$, having the covariance function

$$EB_H(u)B_H(v) = \frac{\sigma^2}{2}\Big\{|u|^{2H} + |v|^{2H} - |u-v|^{2H}\Big\}, \quad u, v \in \mathbb{R}, \tag{2.1}$$

where $\sigma^2 = EB_H^2(1)$. It is seen from (2.1) that fBm B_H has *stationary increments* and is H-*self-similar* in the sense that the processes $B_H(ct)$ and $c^H B_H(t)$ have identical finite-dimensional distributions for any fixed $c > 0$. In fact, fBm is the only (up to a multiplicative constant) Gaussian self-similar process with stationary increments. It plays an important role in Probability, Statistics and is widely used in applications. For more information, see Section 7 in Samorodnitsky and Taqqu (1994), a recent collection of articles Doukhan, Oppenheim and Taqqu (2003), or Embrechts and Maejima (2002).

As shown by Sellan (1995), Meyer et al. (1999), fBm admits a wavelet expansion which converges almost surely on compact intervals and decorrelates the high frequencies. More precisely, one has

$$B_H(t) = \sum_{k=-\infty}^{\infty} \Phi_H(t-k) S_k^{(H)} + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} 2^{-jH}\Psi_H(2^j t - k)\epsilon_{j,k} - b_0, \tag{2.2}$$

where $S_k^{(H)}$, $k \in \mathbb{Z}$, is a *Gaussian FARIMA*$(0, H+1/2, 0)$ sequence, that is, a partial sum process of a stationary Gaussian FARIMA$(0, H-1/2, 0)$ sequence with independent Gaussian innovations $\mathcal{N}(0,1)$, $\epsilon_{j,k}$, $j \geq 0, k \in \mathbb{Z}$, are *independent* Gaussian $\mathcal{N}(0,1)$ random variables and b_0 is a random constant such that $B_H(0) = 0$. (FARIMA sequences are studied, for example, in Brockwell and Davis (1991).) The functions Φ_H and Ψ_H are a suitably chosen *biorthogonal scaling function* and a *wavelet*, respectively. These functions are defined as transformations of a scaling function ϕ and a wavelet ψ corresponding to an orthogonal *MultiResolution Analysis* (*MRA*, in short), for example,

$$\widehat{\Phi}_H(x) = \left(\frac{1-e^{-ix}}{ix}\right)^{H+1/2} \widehat{\phi}(x), \tag{2.3}$$

where $\widehat{f}(x) = \int_{\mathbb{R}} e^{-itx} f(t)dt$ denotes the Fourier transform of a function f. The function Φ_H^*, biorthogonal to Φ_H, is defined by

$$\widehat{\Phi}_H^*(x) = \left(\frac{1 - e^{ix}}{-ix} \right)^{-(H+1/2)} \widehat{\phi}(x). \qquad (2.4)$$

For more details on the wavelet expansion (2.2), see Meyer et al. (1999).

Practical implementation of the wavelet decomposition (2.2) to simulate fBm is described by Abry and Sellan (1996), and Pipiras (2003*a*). Let

$$S_k^{(H)}(j) = 2^{j(H+1)} \int_{\mathbb{R}} (B_H(t) + b_0) \Phi_H^*(2^j t - k) dt, \quad k \in \mathbb{Z}, \qquad (2.5)$$

be the conveniently normalized approximation coefficients in the wavelet expansion of fBm at a scale 2^{-j}, $j \in \mathbb{Z}$. These approximation coefficients can be computed from those at a coarser scale by using the fast wavelet transform as

$$S_\cdot^{(H)}(j) = u^{(s)} * (\uparrow_2 S_\cdot^{(H)}(j-1)) + v^{(s)} * (\uparrow_2 \epsilon_{j-1,\cdot}), \qquad (2.6)$$

where $*$ stands for a convolution and the standard operator $(\uparrow_2 x)$ inserts zeros between the elements of a sequence x, and

$$s = H + 1/2.$$

The filters $u^{(s)}$ and $v^{(s)}$, called *fractional low* and *high-pass filters*, respectively, are defined through the z-transformations as

$$u^{(s)}(z) = (1 + z^{-1})^{N+s} u_0(z), \quad v^{(s)}(z) = (1 - z^{-1})^{N-s} v_0(z), \qquad (2.7)$$

where N is the number of zero moments of an underlying orthogonal MRA having the low and high-pass filters $u(z) = (1 + z^{-1})^N u_0(z)$ and $v(z) = (1 - z^{-1})^N v_0(z)$, respectively. (Recall that the z-transformation of a sequence $x = \{x_k\}$ is defined by $x(z) = \sum_k x_k z^{-k}$.) As shown in Pipiras (2003*a*), the sequence

$$\{S_k^{(H)}(j)\}_{k \in \mathbb{Z}} \qquad (2.8)$$

is also a Gaussian FARIMA$(0, H + 1/2, 0)$ sequence for each j, and

$$\sup_{t \in K} \left| 2^{-jH} S_{[2^j t]}^{(H)}(j) - (B_H(t) + b_0) \right| \leq C 2^{-j(H-\epsilon)}, \qquad (2.9)$$

where $[x]$ stands for the integer part function of $x \in \mathbb{R}$, $K \subset \mathbb{R}$ is a compact, C is a random variable depending on H,ϵ,K and the scaling function ϕ, and $\epsilon > 0$ is arbitrarily small. In other words, the approximation (2.8) tends to fBm exponentially fast as j increases.

In practice, the wavelet-based synthesis works as follows. First, by using, for example, the popular Circular Matrix Embedding method (Dietrich and Newsam (1997)), one generates a Gaussian FARIMA$(0, H + 1/2, 0)$ sequence (2.8) of finite length as an approximation of fBm at some coarse scale. Then, one applies the fast wavelet transform (2.6) recursively a number of times to obtain another, much longer Gaussian FARIMA$(0, H + 1/2, 0)$ sequence. The obtained FARIMA sequence is properly normalized and taken for an approximation of fBm at some finer scale.

3 Parameters and variables of wavelet-based synthesis

The following parameters make part of the wavelet-based synthesis of fBm and will be used in the next sections. Let

$$T = 2^M \tag{3.1}$$

be the time duration $[0, T]$ of the synthesis of fBm. Let

$$2^{-J} \tag{3.2}$$

be the scale at which a final wavelet-based approximation of fBm is taken. This means that a wavelet-based approximation of fBm is taken at the time points $0, 2^{-J}, 2\cdot 2^{-J}, \ldots, 2^{M+J}\cdot 2^{-J}(= T)$ and hence that its length is $2^{M+J}+1$. We will refer to 2^{-J} in (3.2) as a final approximation scale. To obtain a wavelet-based approximation at the scale 2^{-J}, the wavelet-based synthesis starts with an approximation at a coarser scale and then recursively applies to it the scheme (2.6). Let

$$2^{-L} \tag{3.3}$$

with $L \le J$, be the coarser scale at which the initial wavelet-based approximation is taken. We call 2^{-L} in (3.3) an initial scale.

For $j \ge L$, let

$$S_k^{(H)}(L, j) \tag{3.4}$$

be the wavelet approximation coefficients at the scale 2^{-j} computed in practice using (2.6) and starting with the initial wavelet-based approximation at the scale 2^{-L}. Applying the scheme (2.6) $J - L$ times, one obtains the coefficients $S_k^{(H)}(L, J)$ which can be taken as approximation of fBm at the final approximation scale 2^{-J}. Let also

$$B_H(L, t) + b_0(L) = \lim_{J\to\infty} 2^{-JH} S^{(H)}_{[2^J t]}(L, J) \tag{3.5}$$

be the fBm plus the random variable which appear as the limit of the normalized wavelet approximation coefficients (see (2.9)). The parameter L in the left-hand side of (3.5) indicates dependence on the initial scale 2^{-L}.

In order to study the quality of wavelet-based approximations of fBm, we define the random variables

$$C_K(L, j) = \sup_{t\in K}\Big|2^{-jH} S^{(H)}_{[2^j t]}(L, j) - (B_H(L, t) + b_0(L))\Big|, \quad j \ge L, \tag{3.6}$$

where $K \subset [0, T]$, for example, $K = [0, T]$ or $K = \{0, 2^{-j}, 2\cdot 2^{-j}, \ldots, T\}$. In practice, the variables (3.6) cannot be observed. We shall also consider the variables

$$D_K(L, j) = \sup_{t\in K}\Big|2^{-jH} S^{(H)}_{[2^j t]}(L, j) - 2^{-(j+1)H} S^{(H)}_{[2^{j+1} t]}(L, j+1)\Big|, \quad j \ge L, \tag{3.7}$$

which are observable. The variables $D_K(L, j)$ measure the difference (in the sup-norm over K) between the wavelet-based approximations at two consecutive scales 2^{-j} and $2^{-(j+1)}$. Note that

$$C_K(L, j) \le \sum_{l=j}^{\infty} D_K(L, l). \tag{3.8}$$

The bound (2.9) shows that the variables $C_K(L,j)$ and $D_K(L,j)$ converge to zero almost surely and exponentially fast as $j \to \infty$. The evidence of this for the variables $D_K(L,j)$ can be seen in practice. Suppose that $L = 0$ and $T = 1$. In the top two subplots of Figure 1, we present 10 realizations of $D_{[0,1]}(0,j)$ and $\ln D_{[0,1]}(0,j)$ as functions of $j = 0, 1, \dots, 18$. Each replication of $D_{[0,1]}(0,j)$ is obtained by (3.7) through the wavelet-based synthesis starting with the initial Gaussian FARIMA$(0, H + 1/2, 0)$ sequence approximation at the scale 2^0. The self-similarity parameter H was chosen as $H = 0.75$. An initial Gaussian FARIMA sequence was generated by a Circulant Matrix Embedding method (Dietrich and Newsam (1997)). We used the Daubechies MRA with $N = 10$ zero moments and truncated fractional low and high-pass filters $u^{(s)}$ and $v^{(s)}$ at the precision level 10^{-9}. In the bottom two subplots of Figure 1, we present the boxplots of the values of $D_{[0,1]}(0,j)$ and $\ln D_{[0,1]}(0,j)$ as functions of $j = 0, 1, \dots, 18$. The boxplots were computed based on 1000 independent Monte Carlo replications of $D_{[0,1]}(0,j)$. Observe from Figure 1 that $D_{[0,1]}(0,j)$ appears to decay exponentially fast as j increases, supporting the theoretical result.

4 What should an initial scale be?

Supposing that M and J are fixed, we want to understand first whether the choice of the initial scale 2^{-L} affects the quality of the approximation of fBm. For example, in Abry and Sellan (1996), it is suggested to use the initial scale 2^0 ($L = 0$). Pipiras (2003*a*) suggests to use the initial scale 2^M ($L = -M$), the difference from Abry and Sellan (1996) being that the initial approximation needs to be taken of much shorter length when M is large. In other words,

$$\textit{Does it matter what the initial scale } 2^{-L} \textit{ is?} \tag{4.1}$$

We next address this question from both practical and theoretical point of views. The following result shows that distributional properties of the variables $C_K(L,j)$ and $D_K(L,j)$ do not depend on the initial scale 2^{-L}.

Proposition 4.1 *With the notation (3.6) and (3.7) above, we have for fixed J and $L_1, L_2 \le J$,*

$$\{C_K(L_1,j)\}_{j\ge J} \stackrel{d}{=} \{C_K(L_2,j)\}_{j\ge J}, \quad \{D_K(L_1,j)\}_{j\ge J} \stackrel{d}{=} \{D_K(L_2,j)\}_{j\ge J}. \tag{4.2}$$

PROOF: The proof is elementary. By (2.5) and since $\int_{\mathbb{R}} \Phi_H^*(s)ds = \widehat{\Phi}_H^*(0) = 1$, we have

$$C_K(L,j) = \sup_{t\in K} \Big| 2^j \int_{\mathbb{R}} B_H(L,s)\Phi_H^*(2^j s - [2^j t])ds - B_H(L,t)\Big|$$

$$\stackrel{d}{=} \sup_{t\in K} \Big| 2^j \int_{\mathbb{R}} B_H(s)\Phi_H^*(2^j s - [2^j t])ds - B_H(t)\Big|, \quad j \ge J, \tag{4.3}$$

where B_H is a fBm. Since the right-hand side of (4.3) does not depend on L, we obtain the first relation in (4.2). The second relation in (4.2) can be proved in a similar way. □

Evidence for the second equality in (4.2) can be seen in practice. For example, Figure 2 examines the relation

$$D_{[0,1]}(L_1, 18) \stackrel{d}{=} D_{[0,1]}(L_2, 18) \tag{4.4}$$

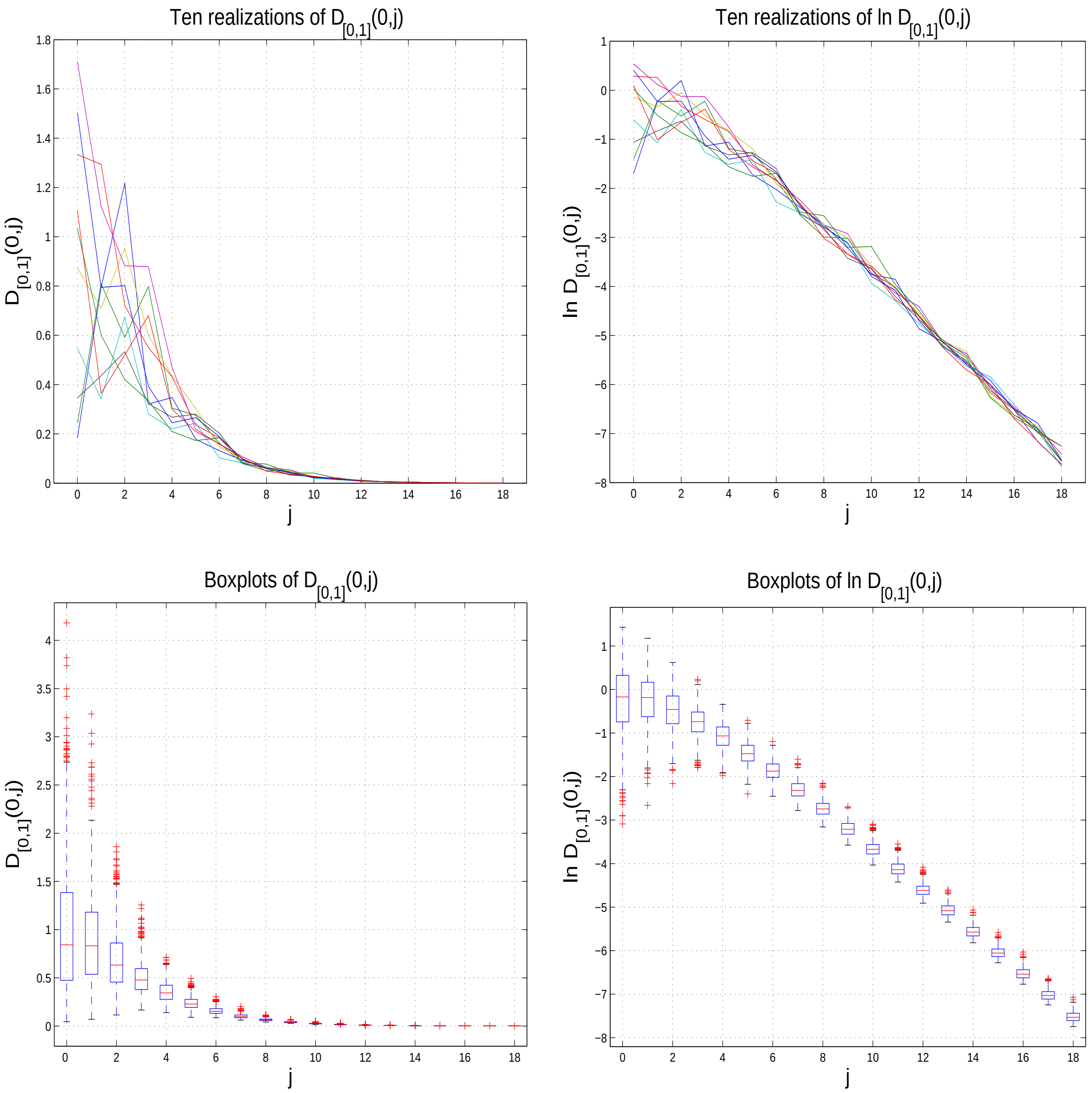

Figure 1: Realizations and boxplots of the variables $D_{[0,1]}(0,j)$ and $\ln D_{[0,1]}(0,j)$ for $j = 0, 1, \ldots, 18$.

for $L_1, L_2 \leq 18$. We present boxplots of the values of $D(L, 18)$ as a function of $L = 0, 1, \ldots, 18$, computed from 1000 Monte Carlo replications of $D_{[0,1]}(L, 18)$ for the specified values of L. Each replication $D_{[0,1]}(L, 18)$ is obtained through (3.7) by using the same parameter H, N, etc., values as at the end of Section 3. Observe from Figure 2 that the boxplots appear identical for different values of L, supporting the theoretical result (4.4). Together with Proposition 4.1, this shows that it does not matter what the starting scale 2^{-L} in the wavelet-based synthesis of fBm is, thus also answering the question raised in (4.1) above. It is important to keep in mind, however, that the length of the initial approximation depends on the initial scale 2^{-L}.

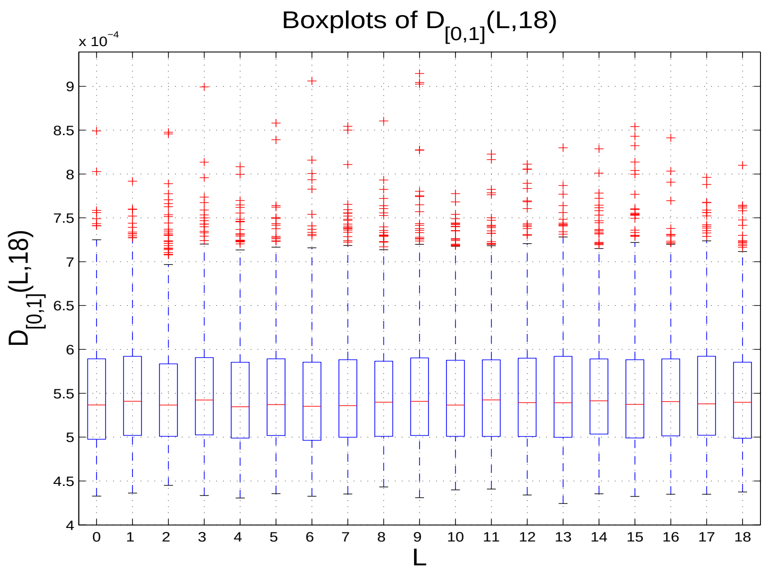

Figure 2: Boxplots of the variables $D_{[0,1]}(L, 18)$ for $L = 0, 1, \ldots, 18$.

In the next section, we look into usefulness of the wavelet-based synthesis of fBm since it is related to irrelevance of the initial scale. Further discussion on the practical use of the synthesis can be found in Section 6.

Remark. Observe that the wavelet decomposition (2.2) of fBm uses the functions Φ_H and Ψ_H as a wavelet basis which makes the approximation coefficients in the wavelet decomposition of fBm a Gaussian FARIMA sequence. Other wavelet bases are possible as well, for example, the so-called optimal wavelet basis studied in Section 8 of Meyer et al. (1999). While Proposition 4.1 states that an initial scale 2^{-L} is irrelevant for the distribution of the error terms $C_K(L, J)$ and $D_K(L, j)$, it might be interesting to examine and compare these variables over different wavelet bases.

5 Usefulness of wavelet-based synthesis

In view of the bound (2.9), one may be led to conclude that the wavelet-based synthesis of fBm is a unique method which provides highly accurate approximations to fBm on compact intervals.

It should therefore be preferred over other simulation methods when such approximations are needed. We argue here that this conclusion should be taken with a reservation.

Suppose that one wants to evaluate the distribution of a functional involving a continuous time fBm, for example, $\int_0^1 F(B_H(s))ds$. In this and many other situations, one needs to have independent Monte Carlo replications of accurate approximation of fBm on an interval. The bound (2.9) suggests that such approximations can be obtained by using the wavelet-based synthesis. Moreover, the Figure 1 type plots would help one in choosing the final scale 2^{-J} for the final wavelet-based approximation. What should one take for an initial scale 2^{-L} in the wavelet-based synthesis? Since the initial scale is irrelevant by the results of Section 4, one may as well choose the initial scale to be the desired scale 2^{-J} itself. In this case, it is not necessary to perform any fast wavelet transform (2.6) iterations to achieve the same accuracy. In other words, consider a simple-minded approximation of fBm defined by

$$Y_J(t) = 2^{-JH} \sum_{k=1}^{[2^J t]} X_k, \tag{5.1}$$

where X_k is a Gaussian FARIMA$(0, H-1/2, 0)$ sequence generated by some method, for example, the Circulant Matrix Embedding method (Dietrich and Newsam (1997)) or the Durbin-Levison algorithm (Brockwell and Davis (1991)). Our results indicate that generating the approximations Y_J would be as good as generating the approximations

$$2^{-JH} S^{(H)}_{[2^J t]}(L, J) \tag{5.2}$$

obtained by using the wavelet-based synthesis with some $L < J$ for the initial scale 2^{-L}. On the other hand, this perhaps should not be very surprising because the approximations (5.2) are Gaussian FARIMA$(0, H + 1/2, 0)$ sequences for all J.

These observations raise the following question: if one can use the simple-minded approximation (5.1), then:

Is the wavelet-based synthesis of fBm useful?

The wavelet-based synthesis of fBm is useful or superior over other simulation methods for at least the following reasons:

1. It provides approximations which converge to fBm almost surely and uniformly on compact intervals.

2. It provides information on the approximation error when using other simulation methods.

3. It is computationally fast.

The wavelet method is also novel and interesting in ideas, and can be potentially generalized to other processes (see, for example, Pipiras (2003*b*), Abry and Pipiras (2003)). Parts 1–3 given above require further explanations. We consider Part 2 to be new and quite interesting.

Part 1 can be useful in at least two ways. First, it allows to visualize the convergence to fBm on an interval. We illustrate this in Figure 3 where eight consecutive approximations to fBm are plotted. It is quite astonishing in practice how fast approximations collapse as the number of iterations increases. The MATLAB code (written jointly with Patrice Abry) which allows to visualize the convergence to fBm, is available from the author upon request. Second, Part 1 allows to obtain accurate approximations in an adaptive way. The discussion above shows that the approximation (5.1) is as good as the approximation (5.2) *on average* (or in distribution). On the other hand, by increasing J, one can in principle make sure that each realization of (5.2) is arbitrarily close to fBm uniformly on an interval. One cannot do so for the approximation (5.1). We know of only one other result which provides almost sure and uniform on compact intervals approximations to fBm, see Szabados (2001).

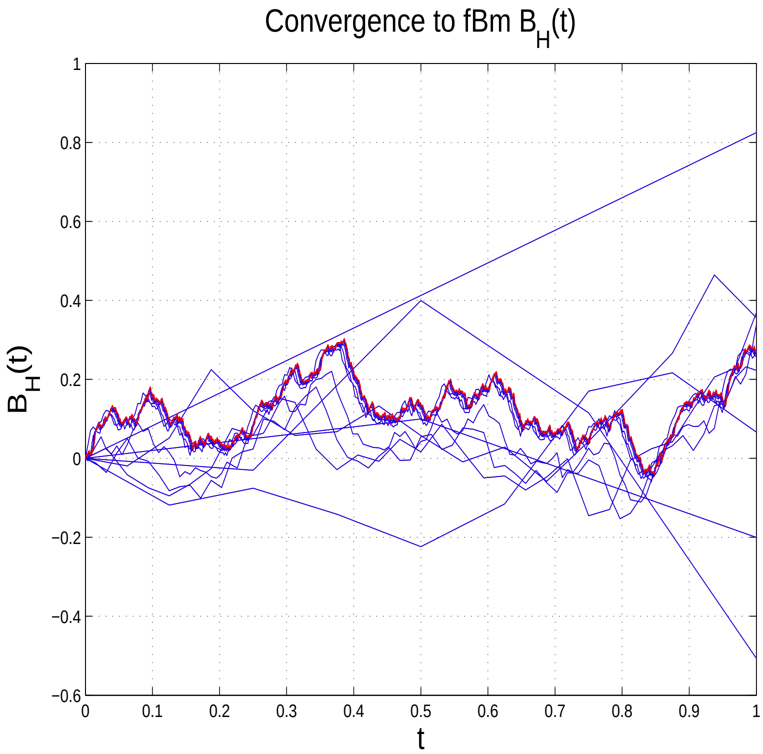

Figure 3: Eight consecutive wavelet-based approximations of fBm B_H with $H = 0.75$.

Part 2 refers to the following. Suppose that X_k, $k \in \mathbb{Z}$, is a Gaussian FARIMA$(0, H - 1/2, 0)$ sequence. Since the partial sum process of this sequence can be viewed as a wavelet-based approximation $S_k^{(H)}(J, J)$, for any $J \in \mathbb{Z}$, there is fBm $B_H(J, t)$, $t \in \mathbb{R}$, such that

$$\sup_{t \in K} \Big| 2^{-JH} \sum_{k=1}^{[2^J t]} X_k - B_H(J, t) \Big| \stackrel{d}{=} C_K(0, J), \tag{5.3}$$

where K is a subset of $\mathbb{R}$ and $C_K(0, J)$ is defined by (3.6) (its distribution does not depend on L by Proposition 4.1). By using the wavelet-based synthesis and the relation (3.8), one can gain information on the random variable $C_K(0, J)$ and hence on the left-hand side of (5.3). Indeed, Figure 1 discussed at the end of Section 3 indicates that $D_K(0, j)$ decays exponentially fast when

j increases. One can generate $D_K(0,j)$ for a finite number of j and extrapolate $D_K(0,j)$ for larger j by supposing that the decay continues exponential. Substituting $D_K(0,j)$ into (3.8), one would obtain an estimate of $C_K(0,J)$. One can then obtain, for example, a boxplot of the values of $C_K(0,J)$. We do not know of other ways to learn about the left-hand side of (5.3) other than through the wavelet-based synthesis of fBm.

Concerning Part 3, suppose that T observations of fBm are needed. If the wavelet-based synthesis of fBm is used, then the number of operations needed to compute a wavelet-based approximation of fBm of length T is of the order $O(T)$, modulo the truncation of the fractional filters (2.7) (see, for example, Mallat (1998), p. 259). The truncation of the fractional filters should not be of big concern because, by taking large number of zero moments N in (2.7), these filters are still short at very small truncation levels (see Pipiras (2003*a*), Section 4 and, in particular, Table 1). Observe that the order $O(T)$ is smaller than the order $O(T^2)$ of the Durbin-Levinson algorithm, the order $O(T\log_2 T)$ of the Circulant Matrix Embedding method and the orders of many other algorithms to simulate fBm (see Bardet et al. (2003)). This means that the wavelet-based method is computationally faster than most methods used to simulate fBm.

6 Use and misuse of wavelet-based synthesis

In practice, starting with an initial FARIMA sequence of finite length, the wavelet-based synthesis of fBm produces approximation coefficients $S_k^{(H)}$ which form another finite but longer FARIMA sequence. With the notation (3.4), one may set

$$S_k^{(H)} = S_k^{(H)}(L,J)$$

for some $L \le J$ but the choice of L and J is subjective. The suitably normalized coefficients $S_k^{(H)}$ can be used as an approximation of fBm at any scale 2^{-J} and a suitable time duration $[0,T]$ of the synthesis. However, it is important to understand that the errors associated with various approximations might be different.

Proposition 6.1 *With the notation (3.6) and (3.7), we have for $J, J' \in \mathbb{Z}$,*

$$C_K(0,J) \stackrel{d}{=} 2^{H(J'-J)} C_{2^{J-J'}K}(0,J'), \quad D_K(0,J) = 2^{H(J'-J)} D_{2^{J-J'}K}(0,J'). \tag{6.1}$$

PROOF: The first identity in (6.1) follows by using the H-self-similarly of fBm as

$$C_K(0,J) = \sup_{t\in K}\left|2^{-JH}S^{(H)}_{[2^J t]}(0,J) - B_H(0,t)\right| = \sup_{2^{J'-J}t\in K}\left|2^{-JH}S^{(H)}_{[2^{J'}t]}(0,J) - B_H(0,2^{J'-J}t)\right|$$

$$\stackrel{d}{=} 2^{H(J'-J)} \sup_{t\in 2^{J-J'}K}\left|2^{-J'H}S^{(H)}_{[2^{J'}t]}(0,J') - B_H(0,t)\right| = 2^{H(J'-J)} C_{2^{J-J'}K}(0,J').$$

The second identity in (6.1) can be obtained in a similar way. □

In other words, by taking the coefficients $S_k^{(H)}$ as an approximation on a larger scale and a larger interval, the approximation errors increase according to the formulas (6.1). Scaling of wavelet-based approximations should therefore be used with great care. When choosing an approximation scale 2^{-J} for a wavelet-based approximation of fBm, one should think about the corresponding approximation error. When the synthesis interval is $[0,T]=[0,1]$, for example, one may gain information about $D_{[0,1]}(0,J)$ directly from Figure 1 and about $C_{[0,1]}(0,J)$ from the same Figure 1 by arguing as in Section 4 (see (5.3)). When $[0,T]$ is larger or smaller, the formula (6.1) could be used in addition.

A consequence of the above discussion is that the sequence $S_k^{(H)}$ should not be used as an approximation of fBm $B_H(k)$ at integer points. This can be seen as follows. Supposing that the length of $S_k^{(H)}$ is 2^J+1, the error of the approximation of $B_H(k)$ by $S_k^{(H)}$ has the same distribution as

$$C_{\{0,1,2,\dots,2^J\}}(0,0) \stackrel{d}{=} 2^{JH} C_{\{0,2^{-J},2\cdot 2^{-J},\dots,1\}}(0,J), \tag{6.2}$$

where the equality in distribution follows from (6.1). Even though the error C on the right-hand side of (6.2) is small for large J, when multiplied by 2^{JH}, we expect it to become large. The evidence of this can be seen from Figure 1, where the related errors $D_{[0,1]}(0,j)$ are reported. When multiplied by 2^{jH}, these small errors become large.

In order to have an accurate approximation of fBm at integer points, it is necessary to downsample the sequence $S_k^{(H)}$, that is, consider the sequence $S_{2^R k}^{(H)}$ for some integer $R>0$ as an approximation of fBm at integer points. This is easy to understand because downsampling chooses in effect a finer final approximation scale 2^{-R} and hence the approximation error decreases according to (2.9).

Since it is necessary to use downsampling to obtain accurate approximations of fBm at integer points, and since there are other exact and computationally fast simulation methods for fBm (e.g. the Circulant Matrix Embedding method), we do not recommend to use the wavelet-based synthesis to generate fBm at integer points. The importance of downsampling in the wavelet-based synthesis of fBm has been observed by Bardet, Lang, Oppenheim, Philippe and Taqqu (2003) who also investigated the wavelet-based synthesis of fBm in practice.

7 Connections to strong invariance principles

Let X_k, $k\in\mathbb{Z}$, be a Gaussian FARIMA$(0,H-1/2,0)$ sequence. Suppose $J\in\mathbb{Z}$ is fixed. In view of the bound (2.9), there is fBm $B_H(J,\cdot)$ such that

$$C_K(J) := \sup_{t\in K}\Big|2^{-JH}\sum_{k=1}^{[2^J t]} X_k - B_H(J,t)\Big| \le C 2^{-J(H-\epsilon)}, \tag{7.1}$$

where $\epsilon>0$ is arbitrarily small, K is a compact and C is a random variable. By the H–self-similarity of fBm, relation (7.1) is equivalent to the condition

$$\sup_{t\in 2^J K}\Big|\sum_{k=1}^{[t]} X_k - B_H(J,t)\Big| \le C 2^{J\epsilon}. \tag{7.2}$$

Observe that this follows from the stronger condition that, for $t \geq 0$,

$$\Big|\sum_{k=1}^{[t]} X_k - B_H(t)\Big| \leq Ct^p, \tag{7.3}$$

where B_H is a fBm and $p = \epsilon > 0$.

Results of the type (7.3) are known in Probability and Statistics as *strong invariance principles* (strong approximations, almost sure invariance principles) or *SIP's*, in short. There is a vast literature on SIP's when $H = 1/2$ (that is, B_H is the usual Bm) and X_k are i.i.d. or, more generally, weakly dependent random variable with finite variance. See, for example, Philipp and Stout (1975), Csörgö and Révész (1981), Philipp (1986) and references therein. The exponent p in (7.3) is given by $p = 1/2 - \eta$ with some $\eta > 0$, when X_k are independent, not necessarily Gaussian, random variables. The case of Gaussian, weakly dependent random variables has been also considered. See, for example, Section 5 in Philipp and Stout (1975). When $H \neq 1/2$, SIP's are not many. See, for example, Wang, Lin and Gulati (2003).

We have indicated these connection to SIP's for a number of reasons. First, SIP's have been used, for example, to deduce the law of the iterated logarithm or the asymptotic behavior of the maximum of the partial sum process of the sequence X_k. The discussion following (7.1) point to another application where the interest is in the approximation of fBm on an interval. Because of a practical interest of such approximations, it is important to study SIP's for other stochastic processes, for example, stable fractional motions or the so-called Hermite processes. If the approximation quality is of interest, wavelet-based synthesis for such processes should also be studied because they can provide information on the approximation error (see Section 5). Second, wavelet decompositions may sometimes provide alternative proofs for SIP's. For example, suppose that X_k is a Gaussian stationary sequence which is either weakly or strongly dependent. By changing the basis in the decomposition (2.2), one may expect to have

$$\sum_{k=1}^{[t]} X_k = \int_{\mathbb{R}} B_H(s) g(s - [t]) ds$$

as the approximation coefficients for some scaling function g. Then,

$$\Big|\sum_{k=1}^{[t]} X_k - B_H(t)\Big| \leq \int_{\mathbb{R}} |B_H([t] + s) - B_H(s)||g(s)| ds$$

and one should be able to bound the latter integral by using the properties of fBm.

8 Beyond fBm

Ideas of this work should be relevant for wavelet-based synthesis of other stochastic processes as well. For example, Pipiras (2003*b*) established a wavelet-type decomposition of the so-called Rosenblatt process. The Rosenblatt process is self-similar, has stationary increments but, contrary to fBm, its finite-dimensional distributions are no longer Gaussian (however, they still have all their moments finite). Practical implementation issues of the wavelet-based synthesis of the Rosenblatt process are discussed in Abry and Pipiras (2003).

References

Abry, P. & Pipiras, V. (2003), Wavelet-based synthesis of the Rosenblatt process, Preprint.

Abry, P. & Sellan, F. (1996), 'The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation', *Applied and Computational Harmonic Analysis* **3**(4), 377–383.

Ayache, A. & Taqqu, M. S. (2002), Rate optimality of wavelet series approximations of fractional brownian motion, Preprint.

Bardet, J.-M., Lang, G., Oppenheim, G., Philippe, A. & Taqqu, M. S. (2003), Generators of long-range dependent processes: a survey, *in* P. Doukhan, G. Oppenheim & M. S. Taqqu, eds, 'Long-Range Dependence: Theory and Applications', Birkhäuser, Boston, pp. 579–623.

Brockwell, P. J. & Davis, R. A. (1991), *Time Series: Theory and Methods*, 2nd edn, Springer-Verlag, New York.

Csörgő, M. & Révész, P. (1981), *Strong Approximations in Probability and Statistics*, Probability and Mathematical Statistics, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York.

Dietrich, C. R. & Newsam, G. N. (1997), 'Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix', *SIAM Journal on Scientific Computing* **18**(4), 1088–1107.

Doukhan, P., Oppenheim, G. & Taqqu, M. S., eds (2003), *Long-Range Dependence: Theory and Applications*, Birkhäuser, Boston.

Embrechts, P. & Maejima, M. (2002), *Selfsimilar Processes*, Academic Press.

Mallat, S. (1998), *A Wavelet Tour of Signal Processing*, Academic Press, Boston.

Meyer, Y., Sellan, F. & Taqqu, M. S. (1999), 'Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion', *The Journal of Fourier Analysis and Applications* **5**(5), 465–494.

Philipp, W. (1986), Invariance principles for independent and weakly dependent random variables, *in* 'Dependence in probability and statistics (Oberwolfach, 1985)', Vol. 11 of *Progr. Probab. Statist.*, Birkhäuser Boston, Boston, MA, pp. 225–268.

Philipp, W. & Stout, W. (1975), 'Almost sure invariance principles for partial sums of weakly dependent random variables', *Mem. Amer. Math. Soc. 2*.

Pipiras, V. (2003*a*), Wavelet-based simulation of fractional Brownian motion revisited, Preprint. Available at http://www.stat.unc.edu/faculty/pipiras/.

Pipiras, V. (2003*b*), Wavelet-type expansion of the Rosenblatt process, Preprint. Available at http://www.stat.unc.edu/faculty/pipiras/. To appear in *The Journal of Fourier Analysis and Its Applications.*

Samorodnitsky, G. & Taqqu, M. S. (1994), *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, London.

Sellan, F. (1995), 'Synthèse de mouvements browniens fractionnaires à l'aide de la transformation par ondelettes', *Comptes Rendus de l'Académie des Sciences de Paris* **321**, 351–358. Série I.

Wang, Q., Lin, Y.-X. & Gulati, C. M. (2003), 'Strong approximation for long memory processes with applications', *Journal of Theoretical Probability* **16**(2), 377–389.

Vladas Pipiras
Department of Statistics and Operations Research
New West, CB #3260
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599, USA
pipiras@email.unc.edu