Technical Report:
Some Topics in Long Memory Process

Han Wang[1]

Version: September 14, 2019

[1]Email: han.wang.fr@gmail.com

Contents

Chapter 1

Introduction

1.1 Fourier transform

We collect here some conventions and useful results about the Fourier transform:

$$\hat{g}(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(t) e^{it\xi} \, \mathrm{d}t, \quad g(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{g}(\xi) e^{-it\xi} \, \mathrm{d}\xi. \tag{1.1}$$

In the following $\xrightarrow{\mathcal{F}}$ denotes the direct transform and $\xrightarrow{\mathcal{F}^{-1}}$ denotes the inverse transform.

Basic identities For L^2 real functions f, g:

$$f * g \xrightarrow{\mathcal{F}} \sqrt{2\pi} \hat{f}\hat{g}, \quad fg \xrightarrow{\mathcal{F}} \frac{1}{\sqrt{2\pi}} \hat{f} * \hat{g}, \quad \langle f, g \rangle_{L^2(\mathbb{R})} = \left\langle \hat{f}, \hat{g} \right\rangle_{L^2(\mathbb{C})}$$

Poisson summation formula Let δ be the Dirac distribution and $T > 0$ a sampling step, then we have the following Fourier transform pair:

$$\sum_{n \in \mathbb{Z}} \delta\left(\cdot - nT\right) \xrightarrow{\mathcal{F}} \frac{\sqrt{2\pi}}{T} \sum_{k \in \mathbb{Z}} \delta\left(\cdot + \frac{2\pi k}{T}\right) \tag{1.2}$$

Sampling operator Let S^T be the sampling operator with a step T applying on a smooth test function g such that in the sense of distribution

$$S^T g(t) = \sum_{n \in \mathbb{Z}} g(nT) \delta(t - nT) = \left(g * \sum_{n \in \mathbb{Z}} \delta\left(\cdot - nT\right) \right)(t) \tag{1.3}$$

and its Fourier transform reads, using the Poisson summation formula,

$$\widehat{S^T g}(\xi) = \frac{1}{T} \sum_{k \in \mathbb{Z}} \hat{g}\left(\xi + \frac{2\pi k}{T}\right) \tag{1.4}$$

Using (1.4) we can easily establish the following. Let ψ, f be real smooth test functions. Then it holds in the sense of distribution

$$\sum_{n \in \mathbb{Z}} \psi(nT) f(t - nT) = \frac{1}{T} \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}} \hat{\psi}\left(\xi + \frac{2\pi n}{T}\right) \hat{f}(\xi) e^{-it\xi} \, \mathrm{d}\xi \tag{1.5}$$

Fourier transform of $|t|^\alpha$ We recall the following result of Fourier transform of a homogeneous distribution (see e.g. Chapter 2.3 of Generalized Functions, Vol 1, Gel'fand):

Lemma 1.1.1. *The Fourier transform of the function $f(t) = |t|^\alpha$ is, in the sense of distribution, when $\alpha \notin \mathbb{Z}$:*

$$\hat{f}(\xi) = \frac{2\Gamma(\alpha+1)}{\sqrt{2\pi}} \cos\left(\frac{\pi}{2}(\alpha+1)\right) |\xi|^{-(\alpha+1)} \tag{1.6}$$

when $\alpha = n$ for n nonnegative odd integer:

$$\hat{f}(\xi) = \frac{2i^{n+1}}{\sqrt{2\pi}} n! \, |\xi|^{-(n+1)} \tag{1.7}$$

Chapter 2

fBm and related processes

We review in this chapter the fractional Brownian motion and related processes by focusing on some important properties such as their representation and covariance structure.

Notations Let $(X(t))_{t\in\mathbb{R}_+}$ be a real-valued time-continuous zero-mean Gaussian process. Its autocovariance function is denoted by

$$\gamma_X(t,s) := \mathbb{E}\left(X(t)X(s)\right) \tag{2.1}$$

while its autocorrelation function is

$$\rho_X(t,s) := \frac{\gamma_X(t,s)}{\sqrt{\gamma_X(t,t)\gamma_X(s,s)}} \tag{2.2}$$

The increment process of X with step $\delta > 0$ is defined as

$$\Delta_\delta X(t) := X(t+\delta) - X(t) \tag{2.3}$$

Sometimes in order to highlight a time-discrete process (or a sequence), we write $X[n]$ and accordingly $\gamma_X[n,m]$ its autocovariance function. This could be the case for a sampled process $(X(n\delta))_n$.

Self-similar process A process is said *self-similar* for a parameter $H \in (0,1)$ if for any $a > 0$ it holds

$$(X(at))_t \overset{\mathcal{L}}{=} a^H (X(t))_t \tag{2.4}$$

which means $(X(at))_t$ and $a^H (X(t))_t$ have the same finite-dimensional probability distribution. In particular this implies $X(0) = 0$ almost surely. The parameter H is called the *Hurst exponent*.

Long memory process A stationary process X is said to have *long memory* or *long-range dependence* if the autocorrelation (or autocovariance) sequence is not summable, *i.e.*

$$\sum_{n=0}^{+\infty} |\rho_X(n)| = +\infty \tag{2.5}$$

This holds in particular if $\rho_X(n)$ decays slowly as $|n|^{-\alpha}$ for some $0 \leq \alpha \leq 1$.

2.1 fBm and its increments

A stochastic process is said H-sssi if it is self-similar with exponent H and its increment process is stationary whatever the step δ. We call *fractional Brownian motion* $\left(B^{(H)}(t)\right)_{t\in\mathbb{R}_+}$ a non-trivial centered H-sssi Gaussian process.

Proposition 2.1.1. *The covariance function of fBm $B^{(H)}$ is given by*

$$\gamma_{B^{(H)}}(t,s) := \mathbb{E}\left(B^{(H)}(t)B^{(H)}(s)\right) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \tag{2.6}$$

where $\sigma^2 = \mathbb{E}\left(\left|B^{(H)}(1)\right|^2\right)$ is the variance.

Proof. Assume without loss of generality that $\sigma = 1$. By the stationarity of the increments and the self-similarity, it holds for any $t > 0$:

$$\mathbb{E}\left(\Delta_\delta B^{(H)}(t)^2\right) = \mathbb{E}\left(B^{(H)}(\delta)^2\right) = \delta^{2H}.$$

This implies that for $t \geq s > 0$

$$\mathbb{E}\left(\left|B^{(H)}(t) - B^{(H)}(s)\right|^2\right) = s^{2H}\,\mathbb{E}\left(\Delta_{\frac{t-s}{s}} B^{(H)}(1)^2\right) = (t-s)^{2H}.$$

On the other hand this is just $t^{2H} + s^{2H} - 2\gamma_{B^{(H)}}(t,s)$, which allows to establish (2.6). □

The fBm is unique in the sense that all H-sssi Gaussian processes share the same finite-dimensional distribution, *i.e.* the same covariance function (2.6), up to some multiplicative constant. A fBm is almost surely continuous (or sample-continuous) and nowhere differentiable, moreover it is H-Hölder continuous. These facts can be established by studying the increment process $\Delta_\delta B^{(H)}$ which has the autocovariance function

$$\mathbb{E}\left(\Delta_\delta B^{(H)}(t)\Delta_\delta B^{(H)}(s)\right) = \frac{\sigma^2}{2}\left(|t-s+\delta|^{2H} + |t-s-\delta|^{2H} - 2\,|t-s|^{2H}\right) \tag{2.7}$$

In particular $\mathbb{E}\left(\left|B^{(H)}(t) - B^{(H)}(s)\right|^2\right) = \sigma^2\,|t-s|^{2H}$. For a general introduction of fBm, see *e.g.* [Doukhan et al., 2002, Samorodnitsky and Taqqu, 1994] and the references therein.

Fractional Gaussian noise We call *fractional Gaussian noise* (fGn) the stationary process of increment

$$X_\delta^{(H)}(t) := B^{(H)}(t+\delta) - B^{(H)}(t) = \Delta_\delta B^{(H)}(t), \tag{2.8}$$

and denote its covariance sequence by $\gamma_{X_\delta^{(H)}}(t) = \left(\sigma\delta^H\right)^2 \gamma_0(t/\delta)$ with

$$\gamma_0(t) := \rho_{X_\delta^{(H)}}(t) = \frac{1}{2}\left(|t+1|^{2H} + |t-1|^{2H} - 2\,|t|^{2H}\right) \tag{2.9}$$

It can be shown that the autocorrelation function γ_0 is strictly positive if $1/2 < H < 1$ (called *persistant*) and strictly negative if $0 < H < 1/2$ (called *anti-persistant*), and decays as $H(2H-1)\,|t|^{2H-2}$. By (2.5) this shows in the persistant case the fGn is a process of long memory. Note that we could equivalently use a causal filter in the definition (2.8), *i.e.* $X_\delta^{(H)}(t) := B^{(H)}(t) - B^{(H)}(t-\delta)$, which gives the same autocovariance function.

2.1.1 Spectrum analysis

For a stationary process X the power spectrum density (psd) function $\mathscr{P}_X$ is a continuous and positive function defined on $[0, 2\pi)$ such that the identity

$$\gamma_X(n) = \int_0^{2\pi} \mathscr{P}_X(\omega) e^{-in\omega} \, \mathrm{d}\omega \tag{2.10}$$

holds for any $n \in \mathbb{Z}$. Using Poisson summation formula one can establish the psd of fGn

$$\mathscr{P}_{X_\delta^{(H)}}(\omega) \propto (1 - \cos\omega) \sum_{n\in\mathbb{Z}} |2\pi n + \omega|^{-(2H+1)}$$

which behaves as ω^{1-2H} around the origin. Therefore it is continuous at the origin when $0 < H < 1/2$ and blows up when $1/2 < H < 1$. More generally there exists an equivalence between

- the slow decay of the autocovariance sequence at infinity, *e.g.* $\gamma_X(n) \simeq |n|^{-\alpha}$ for some $0 < \alpha < 1$, which is typical to the long memory process [Hosking, 1996];
- the power law behavior of the psd at low frequencies, *e.g.* $\mathscr{P}_X(\omega) \simeq |\omega|^{-(1-\alpha)}$, which is typical to the so-called $1/f$ process [Abry et al., 1995, Wornell, 1990, Wornell, 1993, Shusterman and Feder, 1996, Keshner, 1982].

For a rigourous statement of this equivalence see *e.g.* Theorem 15 of [Samorodnitsky, 2002].

Wavelet based pseudo spectrum of fBm

–**To Do**–Periodogram, Welch...

2.2 Integral representation of fBm

The fBm admits different but equivalent integral representations. We state here the *moving average* and *harmonizable* integral representations which are well studied in the literature and can be easily generalized to the multifractional case. Hereafter we denote by $\Gamma(\cdot)$ the Gamma function and $(x)_+ := \max\{0, x\}$ and $(x)_- := -\min\{0, x\}$. We recall first some basic facts about stochastic integrals.

2.2.1 Stochastic integral

Let $(B(t))_{t\in\mathbb{R}_+}$ be the standard Brownian motion, and $\{\psi_n\}_{n\in\mathbb{N}}$ an orthonormal basis of $L^2(\mathbb{R})$. We define respectively the stochastic integrals w.r.t. the Brownian measure dB and the complex Brownian measure $\widehat{dB}$:

$$\int_{\mathbb{R}} f(x)\, dB(x) = \sum_n \langle f, \psi_n \rangle \, Z_n, \quad \int_{\mathbb{R}} g(\xi)\, \widehat{dB}(\xi) = \sum_n \left\langle g, \hat{\psi}_n \right\rangle Z_n \tag{2.11}$$

for real- and complex-valued $L^2(\mathbb{R})$ function f and g respectively, and $\{Z_n\}_{n\in\mathbb{N}}$ is a fixed collection of i.i.d. standard Gaussian random variables. The definitions above are independent of the choices of $\{\psi_n\}_{n\in\mathbb{N}}$ and $\{Z_n\}_{n\in\mathbb{N}}$ since the equality is in the sense of probability distribution. It is easy to see that

$$\hat{f} = g \text{ a.e. } \iff \int_{\mathbb{R}} f(x)\ \mathrm{d}B(x) = \int_{\mathbb{R}} g(\xi)\ \widehat{\mathrm{d}B}(\xi) \text{ a.s.} \tag{2.12}$$

and we have the following Parseval identities:

$$\begin{aligned} \mathbb{E}\left(\int_{\mathbb{R}} f_1(x)\ \mathrm{d}B(x) \int_{\mathbb{R}} f_2(x)\ \mathrm{d}B(x)\right) &= \int_{\mathbb{R}} f_1(x) f_2(x)\ \mathrm{d}x \\ \mathbb{E}\left(\int_{\mathbb{R}} g_1(\xi)\ \widehat{\mathrm{d}B}(\xi) \overline{\int_{\mathbb{R}} g_2(\xi)\ \widehat{\mathrm{d}B}(\xi)}\right) &= \int_{\mathbb{R}} g_1(\xi)\overline{g_2(\xi)}\ \mathrm{d}\xi \end{aligned} \tag{2.13}$$

Kernels of representation for fBm

A sufficient condition of the kernel giving an integral representation of fBm is as follows.

Proposition 2.2.1. *Let $K(t,\cdot)$ be a real-valued $L^2(\mathbb{R})$ function satisfying*

- *there exists $H > 0$ such that for any $a > 0$ it holds $K(at, ax) = a^{H-1/2}K(t,x)$,*
- *for any $\delta > 0$, the following quantity is a function of $t - s$:*

$$\int_{\mathbb{R}} \left(K(t+\delta,x) - K(t,x)\right)\left(K(s+\delta,x) - K(s,x)\right)\ \mathrm{d}x. \tag{2.14}$$

Then the stochastic process defined by

$$B^{(H)}(t) = \int_{\mathbb{R}} K(t,x)\ \mathrm{d}B(x) \tag{2.15}$$

is a fBm with Hurst exponent H.

Proof. We first show that the stochastic process $B^{(H)}(t)$ is H-self-similar, which means $\left(B^{(H)}(at)\right)_{t\in\mathbb{R}_+}$ and $\left(a^H B^{(H)}(t)\right)_{t\in\mathbb{R}_+}$ follow the same finite-dimensional (multivariate normal) distribution. By noting that $\{\sqrt{a}\psi_n(a\cdot)\}_n$ is an orthonormal basis of $L^2(\mathbb{R})$, this follows easily from the definition (2.11) and from the Parseval identity (2.13) which implies that $\mathbb{E}\left(B^{(H)}(at)B^{(H)}(as)\right) = a^{2H}\,\mathbb{E}\left(B^{(H)}(t)B^{(H)}(s)\right)$. On the other hand, the stationarity of the increments comes directly from (2.16) and the Parseval identity. Therefore the process $B^{(H)}$ is a fBm. □

A parallel result exists for the complex-valued Hermitian kernel:

Proposition 2.2.2. *Let $\tilde{K}(t,\cdot)$ be a complex-valued $L^2(\mathbb{R})$ function satisfying*

- $\overline{\tilde{K}(t,\xi)} = \tilde{K}(t,-\xi)$
- *there exists $H > 0$ such that for any $a > 0$ it holds $\tilde{K}(at, a^{-1}\xi) = a^{H+1/2}\tilde{K}(t,\xi)$,*

- *for any $\delta > 0$, the following quantity is a function of $t - s$:*

$$\int_{\mathbb{R}} \left(\tilde{K}(t+\delta,\xi) - \tilde{K}(t,\xi)\right) \overline{\left(\tilde{K}(s+\delta,\xi) - \tilde{K}(s,\xi)\right)}\, \mathrm{d}\xi. \tag{2.16}$$

Then the stochastic process defined by

$$B^{(H)}(t) = \int_{\mathbb{R}} \tilde{K}(t,\xi)\, \widehat{\mathrm{d}B}(\xi) \tag{2.17}$$

is a real-valued fBm with Hurst exponent H.

Proof. Being Hermitian means $\tilde{K}(t,\cdot)$ is the Fourier transform of some real-valued function $K(t,\cdot) \in L^2(\mathbb{R})$. Using (2.12) this implies that the stochastic process $B^{(H)}(t)$ is a real-valued fBm since it can be checked $K(t,\cdot)$ fulfills the two conditions of Proposition 2.2.1. □

2.2.2 Moving average representation

A fBm admits the following moving average representation

$$B^{(H)}(t) = c_H \int_{\mathbb{R}} (t-x)_+^{H-1/2} - (-x)_+^{H-1/2}\, \mathrm{d}B(x) \tag{2.18}$$

where the normalizing constant is chosen as

$$c_H = \left(\frac{\Gamma(2H+1)\sin(\pi H)}{\Gamma(H+1/2)^2}\right)^{1/2}, \tag{2.19}$$

such that $\mathbb{E}\left(\left|B^{(H)}(1)\right|^2\right) = 1$. The averaging kernel

$$K_+(t,x;H) := (t-x)_+^{H-1/2} - (-x)_+^{H-1/2} \tag{2.20}$$

is a L^2 function for all t and satisfies the conditions in Proposition 2.2.1. Note that the power function here is interpreted as

$$(x)_+^{H-1/2} = \begin{cases} x^{H-1/2}, & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases} \tag{2.21}$$

When $H = 1/2$ the kernel (2.20) becomes simply the indicator function $\mathbb{1}[0,t)(x)$. Another possibility of moving average representation is

$$B^{(H)}(t) = c_H \int_{\mathbb{R}} (t-x)_-^{H-1/2} - (-x)_-^{H-1/2}\, \mathrm{d}B(x) \tag{2.22}$$

with the same normalizing constant as (2.19). Remark that the kernel $K_+(t,x;H)$ is *causal* since $K_+(t,x;H) = 0$ for all $x \ge t \ge 0$ and the kernel

$$K_-(t,x;H) := (t-x)_-^{H-1/2} - (-x)_-^{H-1/2} \tag{2.23}$$

is anti-causal. More generally, any kernel of type

$$K_{(a_+,a_-)} := a_+ K_+ + a_- K_- \tag{2.24}$$

for some constants a_+, a_- defines a fBm up to some multiplicative constant. For example the case $a_+ = 1, a_- = 1$ corresponds to the kernel

$$K_+(t,x;H) + K_-(t,x;H) = |t-x|^{H-1/2} - |x|^{H-1/2} \tag{2.25}$$

and the case $a_+ = 1, a_- = -1$ corresponds to the kernel

$$K_+(t,x;H) - K_-(t,x;H) = \operatorname{sign}(t-x)\,|t-x|^{H-1/2} - \operatorname{sign}(-x)\,|x|^{H-1/2} \tag{2.26}$$

where $\operatorname{sign}(\cdot)$ denotes the sign function. Both kernels decay as $O(|x|^{H-3/2})$ at infinity hence are L^2 functions.

A consequence of Lemma 1.1.1 is the following:

Proposition 2.2.3. *For all $t \in \mathbb{R}$, $H \in (0,1)$, the Fourier transforms of the kernel $K_+(t,\cdot;H)$ and $K_-(t,\cdot;H)$ read*

$$\widehat{K_\pm}(t,\xi;H) = \alpha_H \frac{e^{it\xi}-1}{|\xi|^{H+1/2}} e^{\mp i\frac{\pi}{2}\left(H+\frac{1}{2}\right)\operatorname{sign}(\xi)} \tag{2.27}$$

where $\alpha_H := \Gamma(H+1/2)/\sqrt{2\pi}$.

Denoting the constants $\beta_+ = 2\alpha_H \cos\left(\left(H+\frac{1}{2}\right)\frac{\pi}{2}\right)$ and $\beta_- = 2\alpha_H \sin\left(\left(H+\frac{1}{2}\right)\frac{\pi}{2}\right)$, we define therefore the following kernels

$$\mathcal{K}_+ := (K_+ + K_-)/\beta_+, \quad \mathcal{K}_- := (K_+ - K_-)/\beta_-. \tag{2.28}$$

Then using Proposition 2.2.3 we obtain their Fourier transform for all $H \in (0,1)$:

$$\widehat{\mathcal{K}_+}(t,\xi;H) = \frac{e^{it\xi}-1}{|\xi|^{H+1/2}}, \quad \widehat{\mathcal{K}_-}(t,\xi;H) = \frac{e^{it\xi}-1}{i\xi\,|\xi|^{H-1/2}} \tag{2.29}$$

Remark 2.2.1. For the special case $H = 1/2$ we obtain the expressions of these kernels by taking the limit as $H \to 1/2$, namely

$$\mathcal{K}_+(t,x) = \sqrt{\frac{2}{\pi}} \ln\left(\frac{|x|}{|t-x|}\right), \quad \mathcal{K}_-(t,x) = \operatorname{sign}(t-x) - \operatorname{sign}(-x), \tag{2.30}$$

in fact the first one comes from the observation

$$\left(\frac{K_+ + K_-}{\beta_+}\right) = \left(\frac{K_+ + K_-}{H-1/2}\right)\left(\frac{H-1/2}{\beta_+}\right).$$

2.2.3 Harmonizable representation

A fBm admits the harmonizable representation:

$$B^{(H)}(t) = \hat{c}_H \int_{\mathbb{R}} \frac{e^{it\xi} - 1}{i\xi \, |\xi|^{H-1/2}} \widehat{\mathrm{d}B}(\xi) \tag{2.31}$$

where the kernel fulfills the conditions in Proposition 2.2.2 and the normalizing constant is given by

$$\hat{c}_H = \left(\frac{\Gamma(2H+1)\sin(\pi H)}{2\pi} \right)^{1/2}, \tag{2.32}$$

such that $\mathbb{E}\left(B^{(H)}(1)^2\right) = 1$. Another possibility is

$$B^{(H)}(t) = \hat{c}_H \int_{\mathbb{R}} \frac{e^{it\xi} - 1}{|\xi|^{H+1/2}} \widehat{\mathrm{d}B}(\xi) \tag{2.33}$$

with the same normalizing constant as (2.32). The corresponding real kernel for (2.31) and (2.33) are given in (2.29).

The harmonizable representation allows simple computation of the covariance function which makes it easier to manipulate when generalized to the multifractional case.

2.3 Structure of covariance

A fBm $\left(B^{(H)}(t)\right)_{t\in\mathbb{R}_+}$ with fixed Hurst exponent H has the covariance function given by (2.6) (up to some multiplicative constant) and this being independent of the choice of the kernel in the integral representation. We want to know if the same thing can be said for the more general covariance function $\mathbb{E}\left(B^{(H)}(t)B^{(H')}(t')\right)$. In order to give a meaning to this expression, we write $B(t;H) = B^{(H)}(t)$ and interpret $(B(t;H))_{t\in\mathbb{R}_+, H\in(0,1)}$ as a two dimensional random field which admits the expansion

$$B(t;H) = \sum_n \left\langle K_{(a_+,a_-)}(t,\cdot;H), \psi_n(\cdot) \right\rangle Z_n \tag{2.34}$$

with a fixed orthonormal basis $\{\psi_n\}_{n\in\mathbb{N}}$ and a collection of i.i.d. standard Gaussian random variables $\{Z_n\}_{n\in\mathbb{N}}$. Then the question is if different kernels define the same random field (up to some multiplicative constant). We state here the results of [Stoev and Taqqu, 2006] which will be used later to establish the covariance function of a multifractional Brownian motion.

From Proposition 2.2.3 we know that the kernel $K_{(a_+,a_-)}(t,\cdot;H)$ of fBm has the Fourier transform

$$\widehat{K_{(a_+,a_-)}}(t,\xi;H) = \alpha(H)\frac{e^{it\xi} - 1}{|\xi|^{H+1/2}} U_{(a_+,a_-)}(\xi;H) \tag{2.35}$$

where $\alpha(H) = \Gamma(H+1/2)/\sqrt{2\pi}$ and $U_{(a_+,a_-)}$ is defined by

$$U_{(a_+,a_-)}(\xi;H) := a_+ e^{-i\,\mathrm{sign}(\xi)(H+1/2)\frac{\pi}{2}} + a_- e^{i\,\mathrm{sign}(\xi)(H+1/2)\frac{\pi}{2}}. \tag{2.36}$$

It is easy to show for $U_{(a_+,a_-)}$ that its modulus is a constant independent of ξ and its argument equals to $\mathrm{sign}\,(\xi)\,\theta(H)$ with

$$\theta(H) := \mathrm{Arg}\left(a_+ e^{-i(H+1/2)\frac{\pi}{2}} + a_- e^{i(H+1/2)\frac{\pi}{2}}\right) \in [0, 2\pi), \tag{2.37}$$

hence we can write $U_{(a_+,a_-)}(\xi;H) = \left|U_{(a_+,a_-)}(1;H)\right| e^{i\,\mathrm{sign}(\xi)\theta(H)}$, with the modulus fulfills

$$\left|U_{(a_+,a_-)}(1;H)\right|^2 = a_+^2 + a_-^2 - 2a_+a_-\sin(\pi H) \tag{2.38}$$

Theorem 2.3.1 ([Stoev and Taqqu, 2006]). *Let $(B(t;H))_{t\in\mathbb{R},H\in(0,1)}$ be given by (2.34) with the kernel $K_{(a_+,a_-)}$. Let $\bar{H} := (H+H')/2$ and $\Delta_\theta(H,H') := \theta(H) - \theta(H')$, and define the function*

$$h(u) = h(u;H,H') := \cos\left(\Delta_\theta(H,H') - \mathrm{sign}\,(u)\,\pi\bar{H}\right). \tag{2.39}$$

and the constant

$$C(H,H') := \pi^{-1}\Gamma(H+1/2)\Gamma(H'+1/2)\left|U_{(a_+,a_-)}(1;H)\right|\left|U_{(a_+,a_-)}(1;H')\right|. \tag{2.40}$$

Then when $\bar{H} \neq 1/2$ the covariance function has the expression

$$\begin{aligned}\mathbb{E}\left(B(t;H)B(t';H')\right) =& C(H,H')\frac{\Gamma(2-2\bar{H})}{2\bar{H}(1-2\bar{H})} \\ &\times\left(h(t)\,|t|^{2\bar{H}} + h(-t')\left|t'\right|^{2\bar{H}} - h(t-t')\left|t-t'\right|^{2\bar{H}}\right),\end{aligned} \tag{2.41}$$

and when $\bar{H} = 1/2$ the covariance function is the limit of the above expression as $\bar{H} \to 1/2$:

$$\begin{aligned}\mathbb{E}\left(B(t;H)B(t';H')\right) =& C(H,H')\left(\cos(\Delta_\theta(H,H'))\frac{\pi}{2}(|t| + \left|t'\right| - \left|t-t'\right|) - \right. \\ & \left.\sin(\Delta_\theta(H,H'))(t\ln|t| - t'\ln\left|t'\right| - (t-t')\ln\left|t-t'\right|)\right).\end{aligned} \tag{2.42}$$

Remark 2.3.1. For H fixed, all kernels $K_{(a_+,a_-)}$ are equivalent up to some multiplicative constant that depends only on a_+, a_-, H. In fact the covariance function in this case reads

$$\mathbb{E}\left(B(t;H)B(s;H)\right) = \frac{\left|U_{(a_+,a_-)}(1;H)\right|^2\Gamma(H+1/2)^2}{\Gamma(2H+1)\sin(\pi H)}\frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) \tag{2.43}$$

For H non fixed, both $C(H,H')$ and $h(u)$ will depend on a_+, a_- and the covariance function $\mathbb{E}\left(B(t;H)B(s;H')\right)$ generally can not be factorized into a kernel-dependent part and a kernel-independent part, as in (2.43). This means the random fields $(B(t;H))_{t\in\mathbb{R}_+,H\in(0,1)}$ represented by the different kernels $K_{(a_+,a_-)}$ are not equivalent in general.

2.3.1 Equivalent kernels with simplified covariance structure

Below are some special kernels that define the same random field:

- $a_+ = a_0, a_- = a_0$, and $H \neq 1/2$ with

$$a_0(H) = \frac{(\Gamma(2H+1)\sin(\pi H))^{1/2}}{2\cos((H+1/2)\pi/2)\Gamma(H+1/2)} \tag{2.44}$$

  which corresponds to the kernel $\hat{c}_H \mathcal{K}_+$ with $\hat{c}_H$ and $\mathcal{K}_+$ being given by (2.32) and (2.28) respectively. For $H = 1/2$ the kernel has to be modified as in (2.28).

- $a_+ = a_0, a_- = -a_0$, with

$$a_0(H) = \frac{(\Gamma(2H+1)\sin(\pi H))^{1/2}}{2\sin((H+1/2)\pi/2)\Gamma(H+1/2)} \tag{2.45}$$

  which corresponds to the kernel $\hat{c}_H \mathcal{K}_-$ with $\hat{c}_H$ and $\mathcal{K}_-$ being given by (2.32) and (2.28) respectively.

- $a_+ = a_0, a_- = 0$, or $a_+ = 0, a_- = a_0$ with

$$a_0(H) = \frac{(\Gamma(2H+1)\sin(\pi H))^{1/2}}{\Gamma(H+1/2)} \tag{2.46}$$

  which corresponds respectively to the kernel $c_H K_+$ and $c_H K_-$ with c_H being given by (2.19).

Using Theorem 2.3.1 it can be checked easily that in any of these cases the covariance function $\mathbb{E}\left(B(t;H)B(t';H')\right)$ has the same expression, namely

$$\mathbb{E}\left(B(t;H)B(t';H')\right) = D(H,H')\frac{1}{2}\left(|t|^{2\bar{H}} + \left|t'\right|^{2\bar{H}} - \left|t-t'\right|^{2\bar{H}}\right) \tag{2.47}$$

with the constant defined by

$$D(H,H') := \frac{\sqrt{\Gamma(2H+1)\Gamma(2H'+1)\sin(\pi H)\sin(\pi H')}}{\Gamma(2\bar{H}+1)\sin(\pi\bar{H})}, \tag{2.48}$$

and clearly $D(H,H) = 1$.

Lemma 2.3.2. *D defined by (2.48) is a smooth function on $(0,1)\times(0,1)$ such that $0 < D(H,H') \leq 1$. Moreover its first-order partial derivatives evaluated on the diagonal vanish,* i.e.

$$\partial_1 D(H,H) = \partial_2 D(H,H) = 0 \tag{2.49}$$

Proof. The smoothness and the positiveness comes from the that of the Gamma and the sinus function. Using Cauchy-Schwartz inequality and by (2.47), it holds for any $t > 0$

$$\left|\mathbb{E}\left(B(t;H)B(t;H')\right)\right| = D(H,H')\,|t|^{H+H'} \leq \sqrt{\mathbb{E}\left(|B(t;H)|^2\right)}\sqrt{\mathbb{E}\left(|B(t;H')|^2\right)} = |t|^{H+H'}$$

which implies $D(H,H') \leq 1$. Now let $F(H) = \log(\Gamma(2H+1)) + \log(\sin(\pi H))$ and define $L(H,H')$ by

$$L(H,H') := \log D(H,H') = \frac{F(H)+F(H')}{2} - F\left(\frac{H+H'}{2}\right)$$

then (2.49) follows by observing that the partial derivatives of L evaluated on the diagonal vanish. □

2.4 Fractional ARIMA process

Fractional autoregressive integrated moving average (FARIMA or ARFIMA) process is a well known discrete time process having similar power law and long memory behavior as the fBm. Let L be the lag operator such that $Lx(n) = x(n-1)$. The operator of fractional differential (finite difference) is defined as analogy of the integer case (see for example [Hosking, 1984])

$$\nabla^d := (1-L)^d = \sum_{k=0}^{\infty} \frac{d(d-1)\cdots(d-k+1)}{k!}(-L)^k = \sum_{k=0}^{\infty} c_k L^k$$

where the coefficient c_k is expressed using Gamma function as

$$c_k = (-1)^k \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} \sim \frac{k^{-(d+1)}}{\Gamma(-d)} \quad \text{as } k \to \infty, \tag{2.50}$$

and it is easy to check that $(1-L)^d L = L(1-L)^d$.

Definition 2.4.1 (FARIMA(p,d,q) process)**.** Let $\{\varepsilon(n)\}_{n\in\mathbb{Z}}$ be a process of Gaussian noise with $\varepsilon(n) \sim \mathcal{N}(0,1)$ being i.i.d Gaussian variables. Given the autoregressive kernel $\{\phi_i\}_{i=1\ldots p}$ and the moving average kernel $\{\theta_j\}_{j=1\ldots q}$, a fractional ARIMA process $S(p,d,q) = \{S(n)\}_{n\in\mathbb{Z}}$ is defined by the equation:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)\nabla^d S(n) = (1 + \sum_{j=1}^{q} \theta_j L^j)\varepsilon(n), \quad \text{for } n = 1, 2\ldots \tag{2.51}$$

The *fractional integrated* process is the process $S(0,d,0)$ and it can be defined through the z-transformation as

$$S(z) = \left(\frac{1}{1-z^{-1}}\right)^d \varepsilon(z) \tag{2.52}$$

For $|d| < 1/2$ this is a stationary process and has the autocovariance function

$$\gamma_S(k) = \begin{cases} \dfrac{\Gamma(1-2d)}{\Gamma(1-d)^2}, & \text{if } k = 0 \\ \gamma_{S_d}(k-1)\dfrac{k-1+d}{k-d}, & \text{if } k = 1, 2\ldots \end{cases} \tag{2.53}$$

However for $|d| > 1/2$ the process $S(0,d,0)$ is not stationary. This can be seen from the z-transformation (2.52) for $d \in (1/2, 3/2)$ in particular, since

$$\left(1-z^{-1}\right)^{-d} = \left(1-z^{-1}\right)^{-1}\left(1-z^{-1}\right)^{-(d-1)},$$

therefore the process can be expressed as the partial sum (or cumulative sum) of a stationary process. The case $d > 3/2$ can be reduced to $d \in (1/2, 3/2)$ by iterating partial sums.

2.5 Multifractional Brownian motion

We generalize fBm by admitting the Hurst exponent to evovle with time. Let $(B(t;H))_{t\in\mathbb{R}_+,H\in(0,1)}$ be the random field described in the section 2.3.1 (using any of the mentioned kernel and the associated normalization constant). We assume in the following that the Hurst exponent H as a function of time is Hölder continuous with exponent η, satisfying

$$0 < \inf_t H(t) \leq \sup_t H(t) < \eta \tag{2.54}$$

Given H we call a *multifractional Brownian motion* (mBm) the centered Gaussian process $(W^{(H)}(t))_{t\in\mathbb{R}_+}$ such that $W^{(H)}(t) = B(t;H(t))$ for any $t \geq 0$. We present here and in the Section2.6 some properties of mBm by studying its autocovariance function, see also [Peltier and Véhel, 1995, Balança, 2014, Ayache et al., 2000] and the references therein.

Processes of increment The increment process of $W^{(H)}$ is

$$\Delta_\delta W^{(H)}(t) := B(t+\delta;H(t+\delta)) - B(t;H(t)) \tag{2.55}$$

i.e. the increment is in both t and the Hurst function. Similarly we define

$$\Delta_\delta B^{(H(t))}(t) := B(t+\delta;H(t)) - B(t;H(t)) \tag{2.56}$$

which is the increment process of a fBm with the constant Hurst exponent $H = H(t)$ at time t.

Asymptotic equivalence of processes of increment For two processes $(X_\delta(t))_t$ and $(Y_\delta(t))_t$ depending on the parameter δ we write

$$(X_\delta(t))_t \overset{\mathcal{L}}{\sim} (Y_\delta(t))_t, \quad \text{as } \delta \to 0 \tag{2.57}$$

if all finite-dimensional distributions of X_δ and Y_δ are asymptotically equivalent. Similarly we write

$$(X_\delta(t))_t \xrightarrow{\mathcal{L}} (X(t))_t, \quad \text{as } \delta \to 0. \tag{2.58}$$

if X_δ converges in all finite-dimensional distributions to X.

2.5.1 Structure of covariance

We establish here the asymptotic autocovariance function of the increment process of mBm and relate it to that of the fBm. These results will be useful for both the simulation and the estimation of (multiple-) fBm trajectory presented later. Unlike the approach in [Jin et al., 2018], here we do not assume any any particular representation kernel.

In the following we denote by $\bar{H} = (H(t) + H(s))/2$ for any given t, s. Using (2.47) we obtain easily the autocovariance function of the mBm

$$\mathbb{E}\left(W^{(H)}(t)W^{(H)}(s)\right) = D(H(t),H(s))\frac{1}{2}\left(|t|^{2\bar{H}} + |s|^{2\bar{H}} - |t-s|^{2\bar{H}}\right) \tag{2.59}$$

and that of the process $\Delta_\delta B^{(H(\cdot))}$ which is defined in (2.56)

$$\mathbb{E}\left(\Delta_\delta B^{(H(t))}(t)\Delta_\delta B^{(H(s))}(s)\right) = D(H(t),H(s))\times \frac{1}{2}\left(|t-s+\delta|^{2\bar{H}} + |t-s-\delta|^{2\bar{H}} - 2\,|t-s|^{2\bar{H}}\right) \tag{2.60}$$

Remark that (2.60) is $\delta^{2H(t)}$ for $t = s$ and of order δ^2 for $t \neq s$. On the other hand the computation of the autocovariance of the increment process (2.55) is more involved, we establish here an asymptotic result.

Proposition 2.5.1. *Under the assumption (2.54) and for any fixed t it holds, as $\delta \to 0$*

$$\mathbb{E}\left(\left(\Delta_\delta W^{(H)}(t)\right)^2\right) = \mathbb{E}\left(\left(\Delta_\delta B^{(H(t))}(t)\right)^2\right) + o\left(\delta^{1+\eta}\right) \tag{2.61}$$

Moreover, if $\eta > 1$ it holds for fixed $s \neq t$:

$$\mathbb{E}\left(\Delta_\delta W^{(H)}(t)\Delta_\delta W^{(H)}(s)\right) = \mathbb{E}\left(\Delta_\delta B^{(H(t))}(t)\Delta_\delta B^{(H(s))}(s)\right) + o\left(\delta^{1+\eta}\right) \tag{2.62}$$

In particular, the sample path of a mBm is almost surely pointwise Hölder continuous with the local exponent at t being $H(t)$.

Proof. (Sketch of proof) Recalling the definition of increment (2.55), the LHS of (2.62) is the sum of four terms expressed using (2.47), which involves only D and exponential functions. The proof is then essentially based on asymptotic expansions in δ for the exponential functions and in $H(t+\delta) - H(t) = O(\delta^\eta)$ for the functions D. For the case $t \neq s$ the expansion is carried for all functions to the first order, except that the expansion for the exponential functions where the base involves $t - s$ is in δ of the exponent only. For the case $t = s$ the expansion is carried for the function D to the second order, using the property (2.49). The Hölder continuity follows by taking $t = s$ in (2.62). □

The following result is a direct consequence of Proposition 2.5.1 and will be useful later for the simulation of mBm.

Corollary 2.5.2. *The process $(\delta^{-H(t)}\Delta_\delta W^{(H)}(t))_t$ is asymptotically equivalent to the process $(\delta^{-H(t)}\Delta_\delta B^{(H(\cdot))}(t))_t$,* i.e.

$$\left(\frac{W^{(H)}(t+\delta) - W^{(H)}(t)}{\delta^{H(t)}}\right)_t \overset{\mathcal{L}}{\sim} \left(\frac{B^{(H(t))}(t+\delta) - B^{(H(t))}(t)}{\delta^{H(t)}}\right)_t \tag{2.63}$$

as $\delta \to 0$.

Using the same technique of asymptotic expansion, we can establish a result similar to Proposition 2.5.1.

Proposition 2.5.3. *Let t be fixed, it holds for any u, v as $\delta \to 0$:*

$$\mathbb{E}\left(\left(W^{(H)}(t+\delta u) - W^{(H)}(t)\right)\left(W^{(H)}(t+\delta v) - W^{(H)}(t)\right)\right) = \delta^{2H(t)}\frac{1}{2}\left(|u|^{2H(t)} + |v|^{2H(t)} - |u-v|^{2H(t)}\right) + o\left(\delta^{2\eta}\right) \tag{2.64}$$

Proof. Omitted. □

Note that the leading term in (2.64) is just the autocovariance of the fBm $\left(B^{(H(t))}(\delta u)\right)_u$, therefore for a fixed t it holds as $\delta \to 0$:

$$\left(\frac{W^{(H)}(t+\delta u) - W^{(H)}(t)}{\delta^{H(t)}}\right)_u \xrightarrow{\mathcal{L}} \left(B^{(H(t))}(u)\right)_u \tag{2.65}$$

In other words, the tangent process of a mBm at t is a fBm with Hurst exponent $H(t)$.

2.6 Wavelet analysis of covariance

Wavelet is an important tool for the analysis of fBm. Here we invest the autocovariance function of the wavelet filtered process, which is the foundation of the wavelet-based parameter estimation techniques developped later. For similar results see [Flandrin, 1992, Abry et al., 1995, Abry et al., 2003, Abry, 2003, Jin et al., 2018, Roueff and Von Sachs, 2011] and the references therein.

Choice of wavelet Let $\psi \in L^2(\mathbb{R})$ be a wavelet function with compact support on $[0,1]$ and $\|\psi\|_{L^2} = 1$. By Paley-Wiener Theorem the compact support implies the Fourier transform $\hat{\psi}$ is a smooth function and

$$\left|\hat{\psi}(\omega)\right| \lesssim (1+|\omega|)^{-N}, \quad \forall N \in \mathbb{N} \text{ when } \omega \to \infty \tag{2.66}$$

We require also ψ has Q vanishing moments, which means

$$\int_{\mathbb{R}} \psi(t) t^k \, \mathrm{d}t = 0, \text{ for } k = 0 \dots Q-1, \quad \text{and} \quad \int_{\mathbb{R}} \psi(t) t^Q \, \mathrm{d}t \neq 0. \tag{2.67}$$

This is equivalent to say

$$\hat{\psi}^{(k)}(0) = 0 \text{ for } k = 0 \dots Q-1, \quad \text{or} \quad \hat{\psi}(\omega) \simeq \omega^Q \text{ around } 0. \tag{2.68}$$

In particular this implies for any $0 \leq p < Q$ the generalized condition of admissibility:

$$C^{\psi}(p) := \int_{\mathbb{R}} \frac{\left|\hat{\psi}(\omega)\right|^2}{|\omega|^{1+2p}} \, \mathrm{d}\omega < +\infty \tag{2.69}$$

More generally, we introduce the following function depending on a parameter $\rho > 0$:

$$C^{\psi}_{\rho}(\tau, p) := \int_{\mathbb{R}} \frac{\hat{\psi}(\sqrt{\rho}\omega)\hat{\psi}(\sqrt{\rho}^{-1}\omega)^*}{|\omega|^{1+2p}} e^{-i\omega\tau} \, \mathrm{d}\omega \tag{2.70}$$

so that $C^{\psi}(p) = C^{\psi}_1(0,p)$, and the admissibility condition above implies the bound

$$\left|C^{\psi}_{\rho}(\tau, p)\right| < +\infty, \quad \text{uniformly in } \tau, \rho. \tag{2.71}$$

One can also prove for real-valued wavelet that C^{ψ}_{ρ} is also a real-valued function which decays in τ as $O\left(|\tau|^{-2(Q-p)}\right)$ when $|\tau|$ tends to infinity, and it holds $C^{\psi}_{\rho}(\tau, p) = C^{\psi}_{1/\rho}(-\tau, p)$, in particular $C^{\psi}_{\rho}(0,p) = C^{\psi}_{1/\rho}(0,p)$.

Wavelet transform We denote the wavelet function at the scale $a > 0$ by

$$\psi_a(t) := \frac{1}{\sqrt{a}}\psi\left(\frac{t}{a}\right) \tag{2.72}$$

such that ψ_a is supported on $[0, a]$ and $\|\psi_a\|_{L^2} = 1$. For a second order process $(X(t))_t$ such that $\mathbb{E}\left(\left|X(t)^2\right|\right)$ has a polynomial growth in t, the coefficient of analysis with ψ_a is well defined as

$$\mathcal{W}_X^{\psi}(t; a) = \mathcal{W}_X^{\psi_a}(t) := \int_{\mathbb{R}} X(t+u)\psi_a(u)\,\mathrm{d}u \tag{2.73}$$

which is also called this the continuous wavelet transform (CWT)[1]. In case of the dyadic scaling $a = 2^{-j}$ it is called the stationary wavelet transform. We call the discrete wavelet transform if the decimation $t = 2^{-j}n$ is applied in addition. In the latter case we denote the coefficient of scale $j \in \mathbb{Z}$ and position $n \in \mathbb{Z}$ by

$$d_X[n; j] = d_X^j[n] := \mathcal{W}_X^{\psi}(2^{-j}n; 2^{-j}) \tag{2.74}$$

Not that the transform in (2.73) is *anti-causal* if ψ is supported on $t \geq 0$. A *causal* transform can be obtained by translating the support of ψ to $t \leq 0$.

2.6.1 Structure of covariance for fBm

Note that if X is a standard fBm with Hurst exponent H, the property (2.4) gives

$$(\mathcal{W}_X^{\psi_a}(at))_t \overset{\mathcal{L}}{=} a^{H+1/2}(\mathcal{W}_X^{\psi}(t))_t. \tag{2.75}$$

This implies immediately the covariance structure

$$\mathbb{E}\left(\mathcal{W}_X^{\psi_a}(at)\,\mathcal{W}_X^{\psi_a}(at')\right) = a^{2H+1}\,\mathbb{E}\left(\mathcal{W}_X^{\psi}(t)\,\mathcal{W}_X^{\psi}(t')\right) \tag{2.76}$$

where the RHS actually depends only on $t - t'$. In case of discrete wavelet transform this expression becomes

$$\mathbb{E}\left(d_X^j[k]d_X^j[k']\right) = 2^{-j(2H+1)}\,\mathbb{E}\left(d_X^0[k]d_X^0[k']\right). \tag{2.77}$$

In the following we generalize these results to the multifractional setting.

2.6.2 Structure of covariance for mBm

We will consider here the process defined by $X(t) = \sigma(t)W^{(H)}(t)$ which is the standard mBm multiplied by a non-negative volatility function $\sigma \in \mathcal{C}^1$. Recall η the Hölder exponent of the Hurst function fulfilling (2.54). Then under the assumption

$$Q > \sup_t H(t) \tag{2.78}$$

[1]Here and after we follow the convention that the wavelet transform (continuous or discrete) is not a convolution but a correlation, *i.e.* $X(t+u)$ rather than $X(t-u)$ in the integrand.

we can establish the asymptotic covariance structure of the wavelet coefficients

$$\mathbb{E}\left(\mathcal{W}_X^{\psi_a}(t)\,\mathcal{W}_X^{\psi_{a'}}(t')\right) = \int_0^{a'}\int_0^{a} \sigma(t+s)\sigma(t'+s')\mathcal{C}(t+s,t'+s')\psi_a(s)\psi_{a'}(s')\,ds\,ds' \quad (2.79)$$

for small but proportional scale parameters a and a'. Recall that $\mathcal{C}$ here is the covariance function (2.59), *i.e.*

$$\mathcal{C}(t,t') = D(H(t),H(t'))\frac{1}{2}\left(|t|^{2\bar{H}(t,t')} + \left|t'\right|^{2\bar{H}(t,t')} - \left|t-t'\right|^{2\bar{H}(t,t')}\right)$$

with D being given by (2.48) and $\bar{H}(t,t') := (H(t)+H(t'))/2$.

Proposition 2.6.1. *Suppose $a' = \rho a$ for some fixed $\rho > 0$ and $t = ak + t_0, t' = ak' + t_0$ for some fixed k, k' and $t_0 \neq 0$. Under the assumption (2.78) and the condition of admissibility (2.69) the following statements hold as $a \to 0$:*

1. *The covariance of the wavelet coefficients is*

$$\mathbb{E}\left(\mathcal{W}_X^{\psi_a}(t)\,\mathcal{W}_X^{\psi_{a'}}(t')\right) = A_\rho^\psi\left(\frac{k-k'}{\sqrt{\rho}}, H(t_0)\right)\sigma(t_0)^2\left(\sqrt{\rho}a\right)^{2H(t_0)+1} + o(a^{2H(t_0)+1}) \quad (2.80)$$

   where the function A_ρ^ψ is

$$A_\rho^\psi(\tau,H) := \Gamma(2H+1)\sin(\pi H)C_\rho^\psi(\tau,H) \quad (2.81)$$

2. *The variance of the wavelet coefficients is*

$$\mathbb{E}\left(\left|\mathcal{W}_X^{\psi_a}(t_0)\right|^2\right) = A^\psi(H(t_0))\sigma(t_0)^2 a^{2H(t_0)+1} + o(a^{2H(t_0)+1}) \quad (2.82)$$

   where the function A^ψ is

$$A^\psi(H) = \Gamma(2H+1)\sin(\pi H)C^\psi(H). \quad (2.83)$$

In particular, if H and σ are both time-independent constants then (2.80) and (2.82) hold for any $a > 0$ without the small o terms.

Remark 2.6.1. Ignoring the small o term, (2.80) shows that $\mathcal{W}_X^{\psi_a}$ for fixed a is a stationary process.

Proof. We use the following asymptotic expansions for $0 < s < a, 0 < s' < a'$:

- on σ around t_0

$$\sigma(t+s)\sigma(t'+s') = \sigma(t_0)^2 + \sigma(t_0)\sigma'(t_0)(s+s'+a(k+k')) + O(a^2) \quad (2.84)$$

- on D around $H(t_0), H(t_0)$

$$D(H(t+s),H(t'+s')) = 1 + O(\Delta^2 + \Delta'^2) = 1 + O(a^{2\eta}) \quad (2.85)$$

  where $\Delta = H(t+s) - H(t_0) = O(a^\eta)$ and $\Delta' = H(t'+s') - H(t_0) = O(a^\eta)$.

- On the function $g(t) := |t|^{2H(t_0)} \ln|t|$ around t_0. Since $t_0 \neq 0$, the asymptotic
$$g(t+s) = g(t_0) + g'(t_0)(s+ak) + O(a^2)$$
holds for any value of $H(t_0)$. In particular $g(a) = o(a^{H(t_0)})$.

- On the function $|t+s|^{H(t+s)+H(t'+s')}$ as
$$|t+s|^{H(t+s)+H(t'+s')} = |t+s|^{2H(t_0)} + (\Delta+\Delta')g(t+s) + O(\Delta^2+\Delta'^2).$$

- By writing $\mathcal{C}(t,t') = D(H(t),H(t'))A(t,t')/2$, then
$$A(t+s,t'+s') = \underbrace{|t+s|^{2H(t_0)}}_{A_1} + \underbrace{\left|t'+s'\right|^{2H(t_0)}}_{A_2} - \underbrace{\left|a(k-k')+s-s'\right|^{2H(t_0)}}_{A_3} + \underbrace{(\Delta+\Delta')(2g(t_0)+g'(t_0)(s+s'+a(k+k'))+g(s-s'+a(k-k')))}_{A_4 = A_{4,1}+A_{4,2}+A_{4,3}} + O(a^{2+\eta}+a^{2\eta}) \tag{2.86}$$
where the big O term is also $o(a^{\eta+1})$, and $A_{4,1} = (\Delta+\Delta')2g(t_0)$ and so on. Note that $A_{4,1} = O(a^\eta)$, $A_{4,2} = O(a^{\eta+1})$ and $A_{4,3} = o(a^{\eta+H(t_0)})$.

The expression (2.79) now can be analyzed (essentially using Cauchy-Schwartz inequality and the vanishing moments of ψ) term-by-term by injecting the asymptotic expansions of σ, D and A:

- Terms involving $\sigma(t_0)^2$, as well as
  - $A_1 + A_2$: these terms vanish.
  - A_3: this term is
$$-\frac{\sigma(t_0)^2}{2}\int_{\mathbb{R}}\int_{\mathbb{R}} \left|a(k-k')+s-s'\right|^{2H(t_0)} \psi_a(s)\psi_{a'}(s')\,dsds' \tag{2.87}$$
    which equals to the leading term of (2.80) by applying Parseval's identity and the Fourier transform of Lemma 1.1.1. Note that (2.70) is well defined thanks to the assumption (2.78) and the condition of admissibility (2.69). In particular it holds $\left|C_\rho^\psi(\tau,H)\right| \leq C_1^\psi(0,H)$ for any τ, H.
  - $A_{4,1} + A_{4,2}$: these terms vanish.
  - $A_{4,3}$: this term is $o(a^{\eta+H(t_0)+1})$.

- Terms involving $\sigma(t_0)\sigma'(t_0)(s+ak+s'+ak')$, as well as
  - $A_1 + A_2$: these terms vanish.
  - A_3: this term is $O(a^{2H(t_0)+2})$.
  - $A_{4,1}$: this term vanishes.
  - $A_{4,2}$: this term is $O(a^{\eta+3})$.
  - $A_{4,3}$: this term is $o(a^{\eta+H(t_0)+2})$.

Besides (2.87), all these terms are $o(a^{2H(t_0)+1})$. On the other hand, it is easy to see that the three big O terms in (2.84), (2.85) and (2.86) always result in $o(a^{2H(t_0)+1})$. This concludes the proof. □

2.7 Discretization and fractional Wavelet noise

2.7.1 Fractional Wavelet Noise

Let $(\psi[n])_{n\in\mathbb{Z}}$ be a real high-pass filter satisfying

$$\sum_{n\in\mathbb{Z}} \psi[n] = 0 \tag{2.88}$$

and the coefficients decay sufficiently fast with $|n|$ such that the associated Fourier series

$$\hat{\psi}_d(\omega) := \sum_n \psi[n] e^{in\omega} \tag{2.89}$$

is well defined. We call fractional Wavelet noise (fWn) the stationary process defined by

$$V_\delta^\psi(t) := \sum_n \psi[n] B^{(H)}(t + n\delta) \tag{2.90}$$

which is the filtration of a pure fBm by $(\psi[n])_{n\in\mathbb{Z}}$ at sampling step δ. The filter is said *causal* if its support is on negative indices, *i.e.* $\psi[n] = 0$ for any $n > 0$. Similarly it is said *anti-causal* if its support is on positive indices, and *non-causal* (or *centered*) if its support is centered around 0. This is somehow contrary to the usual convention (*i.e.* the filter is *causal* if supported on positive indices) due to the definition of (2.90), see also the footnote about the CWT in section 2.6.

Covariance structure

Given two real high-pass filters ψ, ψ' satisfying the conditions mentionned above, define the sequence

$$a[k] = -\sum_n \psi[n+k]\psi'[n] = -\psi * \widetilde{\psi}'[k] \tag{2.91}$$

where $\widetilde{\psi}'[n] = \psi'[-n]$ and $*$ denotes the discrete convolution. We introduce the covariance function

$$\gamma^{\psi,\psi'}(t) = \frac{1}{2}\sum_{k\in\mathbb{Z}} a[k] \left|t + k\right|^{2H} \tag{2.92}$$

The covariance structure of fWn can be easily established using definition. For this let ψ, ψ' be two high pass filters satisfying (2.88), then for any $t, t' \in \mathbb{R}_+$ it holds

$$\begin{aligned}\mathbb{E}\left(V_\delta^\psi(t) V_\delta^{\psi'}(t')\right) &= -\frac{\sigma^2}{2}\sum_n\sum_{n'} \psi[n]\psi'[n'] \left|t - t' + (n-n')\delta\right|^{2H} \\ &= \frac{\sigma^2}{2}\sum_{k\in\mathbb{Z}} a[k] \left|t - t' + k\delta\right|^{2H} = \left(\sigma\delta^H\right)^2 \gamma^{\psi,\psi'}\left(\frac{t-t'}{\delta}\right)\end{aligned} \tag{2.93}$$

Note that the covariance function (2.92) can be rewritten as, by means of Fourier transform

$$\gamma^{\psi,\psi'}(t) = C_H \int_{\mathbb{R}} \frac{\hat{\psi}_d(\omega)\hat{\psi}'_d(\omega)^*}{|\omega|^{2H+1}} e^{-it\omega}\,\mathrm{d}\omega \tag{2.94}$$

with some constant C_H depending on H only.

Wavelet filter bank

Let ψ be a wavelet with Q vanishing moments satisfying the admissible condition (2.69) $C^{\psi}(H) < +\infty$. The discrete filter at the scale j is taken as

$$\psi_j[n] := \psi_j(n) = \frac{1}{\sqrt{j}}\psi\left(\frac{n}{j}\right) \tag{2.95}$$

and we assume for the moment that the condition (2.88) is also fulfilled. Its Fourier series is expressed as, using Poisson summation formula

$$\widehat{(\psi_j)}_d(\omega) = \sqrt{j}\sum_n \hat{\psi}\left(\omega j + 2\pi n j\right)$$

Let $i, j \in \mathbb{N}$ be two scales with the ratio $\rho = j/i$ being fixed. Denote by γ^{ψ_i,ψ_j} the covariance function (2.92) with the discrete filters ψ_i, ψ_j. Then one can establish easily that as i, j increases the covariance function at $t = 0$ is asymptotically

$$\gamma^{\psi_i,\psi_j}(0) \simeq A^{\psi}_{\rho}(0, H)\left(\sqrt{ij}\right)^{2H+1} \tag{2.96}$$

where A^{ψ}_{ρ} is defined as in (2.81). In particular the variance of the wavelet coefficient at a large scale i is asymptotically

$$\mathbb{E}\left(\left|V^i_{\delta}(t)\right|^2\right) \simeq A^{\psi}(H)\left(\sigma\delta^H\right)^2 i^{2H+1} \tag{2.97}$$

where A^{ψ} is defined as in (2.83). In the next section we generalize this result to the multi-fBm setting.

Single scale fGn As an example of fWn consider the filter

$$\psi_j = [-1, \underbrace{0, \ldots 0}_{j \text{ zeros}}, 1] \tag{2.98}$$

then taking $\psi' = \psi_j$ in (2.91) yields

$$a[k] = \begin{cases} 1, & l = \pm(j+1) \\ -2, & l = 0 \\ 0, & \text{otherwise} \end{cases}$$

and (2.93) is the autocovariance function of fGn at the scale j.

2.7.2 Discretization of CWT and High-frequency limit

Let δ be the sampling step of observation, at the high-frequency limit the wavelet transform (2.73) can be approximated using quadrature as

$$\mathcal{W}^{\psi_a}_X(t) \simeq \frac{\delta}{\sqrt{a}}\sum_{n\in\mathbb{Z}} X(n\delta)\psi\left(\frac{n\delta - t}{a}\right) \tag{2.99}$$

On compactly supported ψ this is a finite-length sum. The fine scales $a \ll \delta$ contain essentially aliasing due to the sampling of X hence we should consider only $a \geq \delta$. We can choose $a = j\delta, t = m\delta$ and define the discrete version of CWT on a discrete trajectory $(X[n])_{n\in\mathbb{Z}}$ with $X[n] := X(n\delta)$ to be

$$w_X^{\psi_j}[m] = w_X^{\psi}[m; j] := \frac{1}{\sqrt{j}} \sum_{n\in\mathbb{Z}} X[n]\psi\left(\frac{n-m}{j}\right) \tag{2.100}$$

where $m \in \mathbb{Z}, j \in \mathbb{N}$. Note that a factor $\sqrt{\delta}$ is dropped in (2.99) such that the sampling step δ does not appear formaly in the definition of DCWT.

As a corollary we can establish a discrete version of the asymptotic covariance structure at the high-frequency limit.

Proposition 2.7.1. *Suppose $\hat{\psi}$ has p-th integrable derivative. For a fixed t_0 let $l_0 = \lfloor t_0/\delta \rfloor$. Let $m, n \in \mathbb{Z}$ and $i, j \in \mathbb{N}$ be fixed and $\rho = j/i$. Then under the same assumptions as in Proposition 2.6.1 the following statements hold as $\delta \to 0$:*

1. *The covariance of the wavelet coefficients is*

$$\mathbb{E}\left(w_X^{\psi_i}[l_0+m]\, w_X^{\psi_j}[l_0+n]\right) = A_\rho^\psi\left(\frac{m-n}{\sqrt{ij}}, H(t_0)\right)\left(\sigma(t_0)\delta^{H(t_0)}\right)^2\left(\sqrt{ij}\right)^{2H(t_0)+1} + O((\sqrt{ij})^{-p}) + o(\delta^{2H(t_0)+1}). \tag{2.101}$$

2. *The variance of the coefficients is*

$$\mathbb{E}\left(\left|w_X^{\psi_i}[l_0]\right|^2\right) = A^\psi(H(t_0))\left(\sigma(t_0)\delta^{H(t_0)}\right)^2 i^{2H(t_0)+1} + O(i^{-p}) + o(\delta^{2H(t_0)+1}). \tag{2.102}$$

where the functions A_ρ^ψ, A^ψ above are the same as in Proposition 2.6.1.

Sketch of proof. The discretization in (2.100) will introduce aliasing in the Fourier domain and the aliasing error can be controlled using Paley-Wiener theory. Define $a_1 = i\delta, a_2 = j\delta$ and $t_1 = t_0 + m\delta, t_2 = t_0 + n\delta$. Now in (2.80) replacing a, t by a_1, t_1 and a', t' by a_2, t_2 respectively and using the fact that at t_0 the mBm is $H(t_0)$-Hölder continuous to bound the error of discretization, we obtain the desired result. □

Remark 2.7.1. The presence of the term $\delta^{2H(t_0)}$ in (2.101) and (2.102) seems a suprise since the discretization (2.100) does not depend formly on the sampling step. In fact δ is "hidden" in the discrete trajectory of $(X[n])_{n\in\mathbb{Z}}$ due to the self-similarity of the process X.

It is easy to establish the correlation of the wavelet coefficients

$$\mathrm{Cor}\left(w_X^{\psi_i}[l_0+m]\, w_X^{\psi_j}[l_0+n]\right) \simeq \left(A^\psi(H(t_0))\right)^{-1} A_\rho^\psi\left(\frac{m-n}{\sqrt{ij}}, H(t_0)\right) \tag{2.103}$$

and the decay of the correlation is determined by that of the funtion A_ρ^ψ.

2.7.3 B-Spline wavelet transform

As an example of discretization we consider here the B-Spline wavelet. Recall the Haar wavelet defined on $[0,1)$

$$\psi^{(0)}(t) := \begin{cases} 1, & \text{if } t \in [0,1/2), \\ -1, & \text{if } t \in [1/2,1), \\ 0, & \text{otherwise.} \end{cases} \tag{2.104}$$

and the discrete Haar transform at even scales

$$w_X^{\psi^{(0)}}[m;2j] = \frac{1}{\sqrt{2j}} \left(\sum_{n=m}^{m+j-1} X[n] - \sum_{n=m+j}^{m+2j-1} X[n] \right) \tag{2.105}$$

which corresponds to an anti-causal convolution. We emphasize that only even scale transform is defined due to the half-closed intervals convention adopted in (2.104). More generally we can take the q-fold auto-convolution of $\psi^{(0)}$ as the mother wavelet

$$\psi^{(q)}(t) := \underbrace{\psi^{(0)} * \cdots * \psi^{(0)}}_{q-\text{fold convolutions}}(t) \tag{2.106}$$

which has $q+1$ vanishing moments, and equivalently expressed in the frequency domain as

$$\widehat{\psi^{(q)}}(\omega) = \left(\sqrt{2\pi}\right)^q \left(-\frac{\left(1-e^{i\omega/2}\right)^2}{\sqrt{2\pi} i\omega} \right)^{q+1} \tag{2.107}$$

It is easy to check that for any $j \in \mathbb{N}$ it holds

$$\psi^{(1)}\left(\frac{m}{2j}\right) = \left(\frac{1}{2j}\right) \sum_{n\in\mathbb{Z}} \psi^{(0)}\left(\frac{n}{2j}\right) \psi^{(0)}\left(\frac{m-n}{2j}\right), \quad \text{for any } m \in \mathbb{Z} \tag{2.108}$$

and more generally the computation of discretized CWT (2.100) with $\psi^{(q)}$ can be realized easily thanks to the following result

Proposition 2.7.2. *Let $j \in \mathbb{N}$, then it holds*

$$\psi^{(q)}\left(\frac{m}{2j}\right) = \left(\frac{1}{2j}\right)^q \underbrace{\left(\psi^{(0)}\left(\frac{\cdot}{2j}\right) * \cdots * \psi^{(0)}\left(\frac{\cdot}{2j}\right) \right)}_{q-\textit{fold discrete convolutions}}(m), \quad \textit{for any } m \in \mathbb{Z}, \tag{2.109}$$

where $\psi^{(0)}(\cdot/2j)$ denotes the discrete sequence $\left(\psi^{(0)}(n/2j)\right)_{n\in\mathbb{Z}}$, and the discrete convolution is defined as in (2.108).

Proof. –**To Do**– □

Discrete filter We define the FIR B-Spline filter of order q at scale $2j$ to be the finite-length sequence $(\psi^{(q)}(m/2j))_{m\in\mathbb{Z}}$ in (2.109), which satisfies the condition (2.88) by construction.

Numerical evaluation of $C_\rho^\psi(\tau, H)$

In practice the real-valued function $C_\rho^\psi(\tau, H)$ is evaluated by means of numerical integration and one may run into problems of unstability with a naive implementation due to the (removable) singularities of the integrand. Let $\mathrm{sinc}(t) := \sin(\pi t)/\pi t$ and let $\psi^{(q-1)}$ be the B-Spline wavelet defined originally in (2.106) having q vanishing moments and supported on $[0, q]$. Recall that we defined above the wavelet transform (both continuous and discrete) as a correlation (equivalent to a convolution using the time-reversed ψ), and in practice one may translate the wavelet filter to obtain *causal* or *anti-causal* transform. This will affect the expression of $C_\rho^\psi(\tau, H)$, in particular:

- for non-causal transform or centered wavelet $\psi(t) = \psi^{(q-1)}(t + q/2)$ which has support on $[-q/2, q/2]$, the corresponding function is

$$C_\rho^\psi(\tau, H) := \frac{1}{\pi 16^q} \int_0^{+\infty} \frac{\left(\mathrm{sinc}\left(\omega\sqrt{\rho}/(4\pi)\right) \mathrm{sinc}\left(\omega/(\sqrt{\rho}4\pi)\right)\right)^{2q}}{|\omega|^{2H+1-2q}} \cos(\omega\tau)\, \mathrm{d}\omega. \quad (2.110)$$

- for anti-causal transform or non-centered wavelet $\psi(t) = \psi^{(q-1)}(t)$ which has support on $[0, q]$, the corresponding function is given by a translation of the expression above

$$C_\rho^\psi(\tau - q(\sqrt{\rho} - 1/\sqrt{\rho})/2, H) \quad (2.111)$$

- similarly for causal transform or the left-shifted wavelet $\psi(t) = \psi^{(q-1)}(t + q)$ which has support on $[-q, 0]$, the corresponding function is

$$C_\rho^\psi(\tau + q(\sqrt{\rho} - 1/\sqrt{\rho})/2, H) \quad (2.112)$$

Chapter 3

Simulation of fBm and related process

We study in this chapter the numerical simulation of a stochastic process on some time interval $[T_{\min}, T_{\max}]$. For a fBm we can do this on the interval $[0,1]$, then by stationarity of increments and by self-similarity the samples can be shifted and rescaled to adapt to any desired interval. We will consider the regular sampling grid of size N

$$\mathcal{G}_N := \left\{ (T_{\max} - T_{\min}) \times \left(0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}\right) + T_{\min} \right\}$$

In fact, it is only on the regular grid that the covariance matrix of a stationary process has the Toeplitz structure.

A sampling method of a stochastic process is called *exact* if the samples follow the finite dimensional probability distribution of the stochastic process when restricted on the sampling grid. For a centered Gaussian process this amounts to sample a multivariate Gaussian distribution of some prescribed covariance matrix.

3.1 Overview of sampling methods

For a sampling step δ, the discrete trajectory of a fBm $\left\{B^{(H)}(k\delta)\right\}_{k=1\dots N}$ follow a centered multivariate Gaussian distribution with covariance matrix

$$(\mathbf{\Gamma}_{B^{(H)}})_{n,m} = \gamma_{B^{(H)}}(n\delta, m\delta), \tag{3.1}$$

which is symmetric and definite positive. Here we intentionally excluded the sample $B^{(H)}(0)$ to avoid having a singular covariance matrix. Recall the time-discrete fGn $\left(X_\delta^{(H)}[k]\right)_{k=0\dots N-1}$ is defined as

$$X_\delta^{(H)}[k] := B^{(H)}(k\delta + \delta) - B^{(H)}(k\delta)$$

and its covariance matrix reads

$$\left(\mathbf{\Gamma}_{X_\delta^{(H)}}\right)_{n,m} = \gamma_{X_\delta^{(H)}}\left((n-m)\delta\right) = (\sigma\delta^H)^2\gamma_0(n-m) \tag{3.2}$$

which is symmetric and positive definite Toeplitz matrix. The covariance function $\gamma_{X_\delta^{(H)}}$ was introduced in (2.9).

Relation between covariance matrices Using the fact

$$B^{(H)}(n\delta) = \sum_{k=0}^{n-1} X_\delta^{(H)}[k]$$

we prove the following identity

$$\gamma_{B^{(H)}}(n\delta, m\delta) = \sum_{k=0}^{n-1}\sum_{l=0}^{m-1} \gamma_{X_\delta^{(H)}}((k-l)\delta) \tag{3.3}$$

which can be rewritten in a matrix form as

$$\boldsymbol{\Gamma}_{B^{(H)}} = \boldsymbol{L}\boldsymbol{\Gamma}_{X_\delta^{(H)}}\boldsymbol{L}^\top = \delta^{2H}\boldsymbol{L}\boldsymbol{\Gamma}_{X_1^{(H)}}\boldsymbol{L}^\top \tag{3.4}$$

where $\boldsymbol{L}$ is the lower triangular matrix that all entries (including the diagonal) equals to 1 and it computes the cumulative sum when applied to a vector.

Cholesky method

The simplest exact method for a Gaussian process (stationary or not) consists in computing the Cholesky decomposition of the covariance matrix $\boldsymbol{\Gamma}_{B^{(H)}} = \boldsymbol{W}^\top\boldsymbol{W}$. Then applying $\boldsymbol{W}$ on a N-dimensional Gaussian vector $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 I)$ gives the random vector $\boldsymbol{x} = \boldsymbol{W}\boldsymbol{\epsilon}$ with the desired covariance matrix $\boldsymbol{\Gamma}$. This method actually applies for any Gaussian process on any type of sampling grid and has the time complexity $O(N^3)$, which is prohibitive in practice when $N > 10^4$ (on a laptop with 2.8GHz Intel i7 CPU and 16GB memory).

Exact methods

Generally speaking a stationary process is easier to simulate than a non-stationary one thanks to its Toeplitz structure. On the other hand, we can first simulate a fGn process then compute the cumulative sum to obtain a sample trajectory of the corresponding fBm process, and this method will be exact as long as the simulation of the fGn is so, as shown by the identity (3.4). This strategy is used in particular by the circulant embedding method, the Hosking method and the CRMD method that we present later.

Sequential sampling of a Gaussian process

Let X, Y be two random variables with the joint p.d.f. $\mathrm{d}\,\mathbb{P}_{X,Y}(x, y)$. In order to draw a sample (x, y) from the joint distribution, one can draw first a sample x from the distribution $\mathrm{d}\,\mathbb{P}_X(\cdot)$ then a sample y from the conditional distribution $\mathrm{d}\,\mathbb{P}_{Y|X}(\cdot|x)$. For the Gaussian random vectors this property can also be verified using the formulae of conditional distribution. Let $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu_x}, \Sigma_{\boldsymbol{x}})$ and $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu_y}, \Sigma_{\boldsymbol{y}})$ be two Gaussian random vectors. The random vector $(\boldsymbol{x}, \boldsymbol{y})$ has the same distribution as $(\boldsymbol{x}, \hat{\boldsymbol{y}})$, where

$$\hat{\boldsymbol{y}} = \mathbb{E}(\boldsymbol{y}|\boldsymbol{x}) + \boldsymbol{\epsilon} = \boldsymbol{\mu_y} + \Sigma_{\boldsymbol{y}}\Sigma_{\boldsymbol{x}}^{-1}(\boldsymbol{x} - \boldsymbol{\mu_x}) + \boldsymbol{\epsilon}$$

and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma_{\boldsymbol{y}} - \Sigma_{\boldsymbol{xy}}^\top\Sigma_{\boldsymbol{x}}^{-1}\Sigma_{\boldsymbol{xy}})$. This suggests an iterative way for sampling a (centered) Gaussian process $X(t)$ at the points $\{t_0, t_1, \ldots\}$:

- Initialization: draw a sample $\hat{X}(t_0)$ from the distribution $\mathcal{N}(0, \sigma_0^2)$
- Step $n \geq 1$: suppose the samples $\hat{X}(t_0) \dots \hat{X}(t_{n-1})$ have been generated.
  - Compute the coefficients of linear prediction $\{\varphi_k^n, k = 0 \dots n-1\}$ such that

$$\mathbb{E}\left(X(t_n) | X(t_0), \dots X(t_{n-1})\right) = \sum_{k=0}^{n-1} \varphi_k^n X(t_k)$$

    and the variance

$$\sigma_n^2 = \mathbb{E}\left(\left(X(t_n) - \sum_{k=0}^{n-1} \varphi_k^n X(t_k)\right)^2\right)$$

  - Draw a sample $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_n^2)$ and set the n-th sample to

$$\hat{X}(t_n) = \sum_{k=0}^{n-1} \varphi_k^n \hat{X}(t_k) + \boldsymbol{\epsilon}$$

We note that the procedure above does not depend on the chronological order of the points $\{t_0, t_1 \dots\}$, and can be easily adapted to the case of multivariate Gaussian vector, or to have more than one sample per step. This principle of conditionalized sampling is in particular useful for establishing the Hosking method and the midpoint method as we shall see later.

3.2 Sampling methods for fBm

In this section we present several popular sampling methods for fBm. Excellent literature reviews of the subject can be found for example in [Coeurjolly, 2000, Bardet et al., 2003, Dieker, 2004].

3.2.1 Circulant embedding method

This method is also refered to as the Wood-Chan method and it is exact in sampling fGn [Baraniuk and Crouse, 1999, Dietrich and Newsam, 1997, Chan and Wood, 1998, Craigmile, 2003]. The key observation here is that on the regular grid $\mathcal{G}_N$ the covariance matrix of a fGn can be embedded into a bigger circulant matrix, which can be efficiently diagonalized with Fourier transform. The method can be adapted also to a general stationary process although it may not be exact in that case.

Let $(X(n))_{n=0,1\dots}$ be a discrete-time centered stationary process and denote the sequence of auto-covariance by $c_n = \gamma_X(n)$. We concatenate the sequence $[c_0, \dots c_{N-1}]$ with the reverted sequence $[c_{N-2}, c_{N-3}, \dots c_1]$ as

$$\tilde{\boldsymbol{c}} = [c_0, c_1 \dots c_{N-1}, c_{N-2}, c_{N-3}, \dots c_1] \tag{3.5}$$

which has size $M = 2N - 2$, and construct the symmetric circulant matrix from $\tilde{\boldsymbol{c}}$ as

$$\tilde{\boldsymbol{C}} = \begin{pmatrix} c_0 & c_1 & \dots & c_{N-1} & c_{N-2} & \dots & c_1 \\ c_1 & c_0 & \dots & c_{N-2} & c_{N-1} & \dots & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ c_{N-1} & c_{N-2} & \dots & c_0 & c_1 & \dots & c_{N-2} \\ c_{N-2} & c_{N-1} & \dots & c_1 & c_0 & \dots & c_{N-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \dots & c_{N-2} & c_{N-3} & \dots & c_0 \end{pmatrix}$$

whose the upper left N-by-N submatrix is nothing but the covariance matrix of the samples $(X(n))_{n=0,1\dots N-1}$.

Proposition 3.2.1. *For a fGn* $\left(X_\delta^{(H)}[n]\right)_{n=0,\dots N-1}$*, the Fourier transform of* $\tilde{\boldsymbol{c}}$

$$\left\{ \lambda_k = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} \tilde{c}_m e^{-2\pi i k m/M},\ k = 0, \dots M-1 \right\} \tag{3.6}$$

is real and non negative. In particular, $\tilde{\boldsymbol{C}}$ *is the covariance matrix of some random vector.*

As a circulant matrix, $\tilde{\boldsymbol{C}}$ can be diagonalized as

$$\tilde{\boldsymbol{C}} = \boldsymbol{F}^* \boldsymbol{\Lambda} \boldsymbol{F}$$

where $\boldsymbol{F} = (\boldsymbol{F})_{mn}$ is the Fourier transform matrix with

$$\boldsymbol{F}_{mn} = \frac{1}{\sqrt{M}} \exp(-2\pi i m n/M)$$

and $\boldsymbol{\Lambda}$ is the diagonal matrix with $(\boldsymbol{\Lambda})_{kk} = \lambda_k$.

Proposition 3.2.2. *Let* $\boldsymbol{u}_1, \boldsymbol{u}_2 \sim \mathcal{N}(0, I_M)$ *be two iid real Gaussian random vectors, and*

$$\boldsymbol{x} := \boldsymbol{F}^* \boldsymbol{\Lambda}^{1/2} (\boldsymbol{u}_1 + i\boldsymbol{u}_2). \tag{3.7}$$

Let $\boldsymbol{x}_R, \boldsymbol{x}_I$ *be the real and imaginary part of the complex vector* $\boldsymbol{x}$*. Then it holds*

$$\mathbb{E}\left(\boldsymbol{x}_R \boldsymbol{x}_R^\top\right) = \mathbb{E}\left(\boldsymbol{x}_I \boldsymbol{x}_I^\top\right) = \tilde{\boldsymbol{C}}$$

Circulant embedding method

This method has the time complexity $O(N \log N)$ and is efficient for the simulation of long trajectory, *e.g.* $N = 10^6$.

- Input: length of the trajectory N, or $M = 2N - 2$
- Step 1: given $c_n = \gamma_{X_H^\delta}(n\delta)$, construct the vector $\tilde{\boldsymbol{c}}$ as in (3.5) and compute the vector $\{\lambda_k, k = 0 \dots M\}$ as in (3.6) via FFT;
- Step 2: compute the vector $\boldsymbol{x}$ as in (3.7): generate a sample of complex Gaussian vector $\boldsymbol{u}_1 + i\boldsymbol{u}_2$, reweight it by $\sqrt{\lambda_k}$ and take the inverse FFT;
- Step 3: take the first N coefficients of $\boldsymbol{x}_R$ (or $\boldsymbol{x}_I$) as the output.

3.2.2 Hosking method

The Hosking method, also known as the Levinson-Durbin method, is exact in sampling any stationary process [Hosking, 1984]. Let $(X(n))_{n=0,1,\dots}$ be a discrete centered stationary process. This method operates iteratively and generates a sample of $X(n)$ conditionned on the historical samples of $(X(0), \dots X(n-1))$. According to the principle of sequential sampling explained in section 3.1, this amounts to compute the coefficients of linear prediction and the error of prediction (or the variance). Since the process is stationary these quantities can be computed efficiently using the Levinson-Durbin iteration.

This method has time complexity $O(N^2)$ and can be summarized as follows.

- Input is the autocovariance function $\gamma(\cdot)$, or equivalently the partial correlation $\rho(\cdot)$.
- Initialization: set $\sigma_0^2 = \gamma(0)$ and draw a sample of $\hat{X}(0) \sim \mathcal{N}(0, \sigma_0^2)$
- Step $n \geq 1$: suppose $\hat{X}(0), \dots \hat{X}(n-1)$ have been generated, repeat the procedure below as long as $n < N$:
  - Using Levinson-Durbin iteration to update the coefficients of linear prediction $\boldsymbol{\varphi}^n = [\varphi_1^n, \dots \varphi_n^n]$

$$\begin{cases} \varphi_n^n = \sigma_{n-1}^{-2} \left(\gamma(n) - \sum_{k=1}^{n-1} \varphi_{n-k}^{n-1} \gamma(k) \right) \\ \varphi_k^n = \varphi_k^{n-1} - \varphi_n^n \varphi_{n-k}^{n-1}, \quad \text{for } k = 1, \dots n-1 \end{cases} \tag{3.8}$$

    and the variance of residual σ_n^2

$$\sigma_n^2 = (1 - (\varphi_n^n)^2)\sigma_{n-1}^2 = \prod_{k=1}^{n} \left(1 - (\varphi_k^k)^2\right) \sigma_0^2 \tag{3.9}$$

  - Draw a sample $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ and set the n-th sample to

$$\hat{X}(n) = \sum_{k=1}^{n} \varphi_k^n \hat{X}(n-k) + \epsilon$$

  - $n \leftarrow n + 1$

Remark 3.2.1. Note that for the step $n = 1$ the equation (3.8) is reduced to compute the vector $\boldsymbol{\varphi}^1 = [\gamma(1)/\gamma(0)]$. The quantity $\rho(n) = \varphi_n^n$ is also known as the partial correlation and it holds $|\rho(n)| \leq 1$ for all $n \geq 1$.

This method also provides a diagonalization of the covariance matrix.

Proposition 3.2.3. *Let* $\{\boldsymbol{\varphi}^n\}_{n=1,\dots N}$ *and* $\left\{\sigma_n^2\right\}_{n=0,\dots N-1}$ *be the sequence generated by the Levinson-Durbin iteration* (3.8) *and* (3.9)*, then it holds*

$$\boldsymbol{\Gamma} = \boldsymbol{A}^{-1} \boldsymbol{\Sigma} \boldsymbol{A}^{-\top} \tag{3.10}$$

where $\mathbf{\Sigma}$ *is a diagonal matrix with* $(\mathbf{\Sigma})_{nn} = \sigma_n^2$ *for* $n = 0, \dots N-1$*, and the matrices* $\mathbf{\Gamma}$*,* $\boldsymbol{A}$ *are given by*

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(N-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(N-1) & \gamma(N-2) & \dots & \gamma(0) \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\varphi_1^1 & 1 & 0 & \dots & 0 \\ -\varphi_2^2 & -\varphi_1^2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\varphi_{N-1}^{N-1} & -\varphi_{N-2}^{N-1} & \dots & -\varphi_1^{N-1} & 1 \end{pmatrix} \tag{3.11}$$

Proof. Omitted. □

Sampling of a fractional ARIMA(p, d, q) process

The Hosking method allows to simulate a FARIMA$(0, d, 0)$ process which is stationary for $d \in (-1/2, 1/2)$. In the case $d \in ((n-1)/2, (n+1)/2), n \neq 0$ one can simulate first a stationary FARIMA$(0, d-n, 0)$ process then filter the sample trajectory properly, see [Hosking, 1984].

A general FARIMA(p, d, q) process with $p > 0$ or $q > 0$ can be simulated as follows. First we draw a sample trajectory $\{x(n)\}_{n=0,1\dots}$ from a FARIMA$(0, d, 0)$ process:

$$\nabla^d x(n) = \varepsilon(n),$$

then filter it via

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) S(n) = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right) x(n), \tag{3.12}$$

or equivalently via

$$S(n) = \sum_{i=1}^{p} \phi_i S(n-i) + x(n) + \sum_{j=1}^{q} \theta_j x(n-j)$$

In fact, taking the fractional differential ∇^d on the two sides of (3.12) and commuting ∇^d with L^j shows that S is indeed an FARIMA(p, d, q) process.

3.2.3 Conditionalized random midpoint displacement method

The CRMD (or midpoint for short) method is a non-exact method for the simulation of fGn [Norros et al., 1999, Lau et al., 1995]. The same trajectory is generated by dyadic refinement from coarse to fine scale. This method exploits also the idea of sequential sampling, but unlike the Hosking method where the whole history is used and the samples are generated in chronological order, CRMD only keep a truncated history of fixed length and the samples are not drawn in chronological order. This method has time complexity $O(N)$.

Dyadic grid We consider the fGn process of increment 2^{-j} on the interval $[0,1]$:

$$X_k^j := X_{2^{-j}}^{(H)}(k), \quad \text{for } k = 0, \ldots, 2^j - 1 \tag{3.13}$$

with the autocovariance function

$$\gamma_j(n) := \gamma_{X_{2^{-j}}^{(H)}}(n) = 2^{-j2H}\gamma_0(n) \tag{3.14}$$

where $\gamma_0(\cdot)$ is given by (2.9). Note that $X_k^{j-1} = X_{2k}^j + X_{2k+1}^j$ so

$$\begin{aligned} \mathbb{E}\left(X_k^{j-1} X_{2k'}^j\right) &= \gamma_j(2(k-k')) + \gamma_j(2(k-k')+1), \quad \text{and} \\ \gamma_{j-1}(n) &= 2\gamma_j(2n) + \gamma_j(2n+1) + \gamma_j(2n-1) \end{aligned}$$

CRMD method

Input: length of the desired trajectory N, index of the coarsest scale j_0, and window size w.

Step 1: (Coarse scale sampling) Generate a sample trajectory at the coarsest scale j_0 using Cholesky or the circulant embedding method.

Step 2: (Dyadic refinement) Suppose the finest scale now is $j-1$, and the samples X^{j-1} have been generated. If $N \leq 2^{j-1}$, keep only the first N samples and terminate, otherwise refine X^{j-1} to the next scale j by conditionalized sampling. This consists of

2a. the initial sampling of w points of even index at the leftmost;

2b. the forward propagation (from left to right) of all points of even index;

2c. the computation of all points of odd index;

and this step has to be repeated as long as necessary. The reason for step 2a is for better handling the points near the bord since the errors made on these points will be propagated in step 2b. We note that other scheme is possible, *e.g.* , initial sampling at the center followed by a forward and backward propagation.

perhaps define more precisely what you mean with forward propagation

Step 2a: (Initial sampling) Generate the w leftmost points of even index of the process X^j conditioned on the w leftmost points of the upper process X^{j-1}.

$$\tilde{X}_{2k}^j := \mathbb{E}\left(X_{2k}^j | X_0^{j-1}, \ldots X_{w-1}^{j-1}\right) = \sum_{n=0}^{w-1} \varphi_n^k X_n^{j-1} \tag{3.15}$$

By definition of the orthogonal projection, it holds for any $n' = 0, \ldots w-1$ that

$$\mathbb{E}\left((X_{2k}^j - \tilde{X}_{2k}^j) X_{n'}^{j-1}\right) = 0 \tag{3.16}$$

which gives the following Yule-Walker equation

$$\sum_{n=0}^{w-1} \gamma_{j-1}(n'-n)\varphi_n^k = \gamma_j(2(n'-k)) + \gamma_j(2(n'-k)+1) \tag{3.17}$$

Using (3.14) this can also be written as

$$2^{2H}\sum_{n=0}^{w-1} \gamma_0(n'-n)\varphi_n^k = \gamma_0(2(n'-k)) + \gamma_0(2(n'-k)+1) \tag{3.18}$$

or in matrix form as $\boldsymbol{\Gamma_0\varphi^k = v^k}$. The conditionalized sampling of initialization is expressed as

$$X_{2k}^j = \tilde{X}_{2k}^j + U_{2k}^j \tag{3.19}$$

with U_{2k}^j being the residual of projection. U^j is a centralized Gaussian vector with the covariance matrix

$$(\boldsymbol{\Gamma_j^U})_{k,k'} = \mathbb{E}\left(U_{2k}^j U_{2k'}^j\right) = \mathbb{E}\left(U_{2k}^j X_{2k'}^j\right) = 2^{-j2H}\left(\gamma_0(2(k-k')) - \boldsymbol{(v^k)^\top \Gamma_0^{-1}(v^{k'})}\right) \tag{3.20}$$

The sampling of U^j is done using the Cholesky method.

Step 2b: (Forward propogation) For $k \geq w$, draw sample of X_{2k}^j conditioned on the $2w$ points around X_k^{j-1} of the process X^{j-1} together with the w points of even index to the left of the process X^j:

$$\begin{aligned}\tilde{X}_{2k}^j &:= \mathbb{E}\left(X_{2k}^j | X_{k-w}^{j-1} \dots X_{k-1}^{j-1}, X_k^{j-1}, \dots X_{k+w-1}^{j-1}, X_{2(k-w)}^j \dots X_{2(k-1)}^j\right) \\ &= \sum_{n=0}^{2w-1} \varphi_n^- X_{k-w+n}^{j-1} + \sum_{m=0}^{w-1} \varphi_m^+ X_{2(k-w+m)}^j \end{aligned} \tag{3.21}$$

The Yule-Walker equations read: for $n' = 0 \dots 2w-1$

$$\begin{aligned}\sum_{n=0}^{2w-1} \gamma_{j-1}(n'-n)\varphi_n^- + \sum_{m=0}^{w-1} \left(\gamma_j(2(n'-m)) + \gamma_j(2(n'-m)+1)\right)\varphi_m^+ = \\ \gamma_j(2(n'-w)) + \gamma_j(2(n'-w)+1)\end{aligned} \tag{3.22}$$

and for $m' = 0 \dots w-1$

$$\begin{aligned}\sum_{n=0}^{2w-1} \left(\gamma_j(2(m'-n)) + \gamma_j(2(m'-n)+1)\right)\varphi_n^- + \sum_{m=0}^{w-1} \gamma_j(2(m'-m))\varphi_m^+ = \\ \gamma_j(2(m'-w))\end{aligned} \tag{3.23}$$

which are rewritten also as $\boldsymbol{G_0\varphi = u}$. For the same reason as in (3.18) the system matrix $\boldsymbol{G}_0 \sim 3w \times 3w$ and the RHS vector $\boldsymbol{u}$ are actually independent of k and j, hence the absence of these indices in the vector $\boldsymbol{\varphi := [\varphi^-, \varphi^+]^\top}$. The conditionalized sampling of forward propagation is expressed in the same way as (3.19), while the variance of the residual now becomes

$$(\sigma_j^U)^2 := \mathbb{E}\left((U_{2k}^j)^2\right) = 2^{-j2H}\left(\gamma_0(0) - \boldsymbol{u^\top G_0^{-1} u}\right) \tag{3.24}$$

Step 2c: (Odd indices) We obtain the points of odd index simply by

$$X_{2k+1}^{j} = X_{k}^{j-1} - X_{2k}^{j} \tag{3.25}$$

3.2.4 Wavelet method

The wavelet method for the simulation of fBm is based on the result that a fBm can be approximated uniformly on a compact interval by a normalized FARIMA$(0, H+1/2, 0)$ process, and the simulation of a long FARIMA process can be efficiently done with the Mallat-type algorithm. This method is not exact and has time complexity $O(N)$. We resume here the main results of [Pipiras, 2005, Pipiras, 2004] which improves the original method of [Meyer et al., 1999, Abry and Sellan, 1996].

Asymptotic expansion and approximation

Denote hereafter the constants

$$s = H + 1/2,\ d = H - 1/2$$

Let ϕ, ψ be respectively the real-valued orthonormal scaling and wavelet function associated to a MRA with regularity $r > s$. We define respectively the fractional scaling and wavelet function Φ_H^+, Ψ_H^+ as well as the biorthogonal functions Φ_H^-, Ψ_H^- in the frequency domain as

$$\hat{\Phi}_H^{\pm}(\xi) = \left(\frac{1-e^{-i\xi}}{i\xi}\right)^{\pm s} \hat{\phi}(\xi),\ \hat{\Psi}_H^{\pm}(\xi) = (i\xi)^{\mp s}\hat{\psi}(\xi) \tag{3.26}$$

which are also real-valued functions.

Proposition 3.2.4. *Let $\left\{B^{(H)}(t)\right\}_{t\geq 0}$ be a fBm. The following asymptotic expansion holds almost surely on compact intervals:*

$$B^{(H)}(t) = \sum_{k=-\infty}^{\infty} \Phi_H^+(t-k) S_0^{(H)}(k) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} 2^{-jH}\Psi_H^+(2^j t - k)\varepsilon_j(k) - b_0 \tag{3.27}$$

where b_0 is a random variable such that $B^{(H)}(0) = 0$, $\left\{S_0^{(H)}(k)\right\}_{k\in\mathbb{Z}}$ is a FARIMA$(0, s, 0)$ process starting at $S_0^{(H)}(0) = 0$ such that

$$S_j^{(H)}(k) = 2^{j(H+1)} \int_{\mathbb{R}} B^{(H)}(t)\Phi_H^-(2^j t - k)\, dt \tag{3.28}$$

and $\left\{\varepsilon_j(k) \overset{\text{iid}}{\sim} \mathcal{N}(0,1)\right\}_{k\in\mathbb{Z}}$ is a Gaussian noise being independent of the $\left\{S_0^{(H)}(k)\right\}_{k\in\mathbb{Z}}$ such that

$$\varepsilon^j(k) = 2^{j(H+1)} \int_{\mathbb{R}} B^{(H)}(t)\Psi_H^-(2^j t - k)\, dt \tag{3.29}$$

Moreover, we have the following uniform bound of approximation error on $[0,1]$. *For any* $0 < \epsilon < H$ *and* $j \in \mathbb{Z}$ *:*

$$\sup_{t\in[0,1]} \left| 2^{-jH} S_j^{(H)}([2^j t]) - B^{(H)}(t) \right| \leq C 2^{-j(H-\epsilon)} \tag{3.30}$$

almost surely, where $C = C(H, \epsilon, \phi)$ *is a random variable and* $[2^j t]$ *stands for the integer part of* $2^j t$.

Remark 3.2.2. The conclusions still hold if the condition $S_0^{(H)}(0) = 0$ is dropped. In this case just replace $B^{(H)}(t)$ by $B^{(H)}(t) + b_0$ in (3.28), (3.29) and (3.30).

Remark 3.2.3. Using the self-similarity of fBm it is easy to see that

$$S_j^{(H)}(k) \triangleq \int_{\mathbb{R}} B^{(H)}(t) \Phi_H^-(t-k)\, dt$$

which shows that the process $\left\{ S_j^{(H)}(k) \right\}_{k\in\mathbb{Z}}$ follows the same distribution for any scale $j \geq 0$. In other words they are all FARIMA$(0, s, 0)$ process.

Fast sampling of FARIMA process by dyadic wavelet synthesis

As shown by the uniform bound (3.30), one can take the truncated and normalized process $\left\{ 2^{-jH} S_0^{(H)}(k) \right\}_{k=0\ldots 2^j-1}$ as approximation of $\left\{ B^{(H)}(t), t \in [0,1] \right\}$, and the longer is the truncated process the better is the approximation. The Hosking method allows to sample exactly the FARIMA$(0, s, 0)$ process with the complexity $O(N^2)$, which is still too slow for large N. The Mallat-type algorithm presented here allows to obtain a sample trajectory of FARIMA$(0, s, 0)$ of length $2N$ from a trajectory of length N. Therefore from a short initialization one can dyadically refine the trajectory to a desired length. The global time complexity of this method is $O(N)$.

Let $u = \{u_n\}_n, v = \{v_n\}_n$ be respectively the low- and high-pass filter (of finite length) associated to the MRA. We denote by $\uparrow_2 x$ the up-sampling operation which inserts zero between two consecutive elements of the sequence x, and by $\downarrow_2 x$ the down-sampling operation by a factor 2.

Proposition 3.2.5. *Let* $\{\xi_n\}_n$ *be i.i.d.* $\mathcal{N}(0,1)$ *variables. Suppose that* $\{X_n\}_n$ *is a Gaussian process given by*

$$X_n = \sum_{k=-\infty}^{\infty} h_k \xi_{n-k} \tag{3.31}$$

or in the z-transformation by

$$X(z) = h(z)\xi(z) \tag{3.32}$$

Define the filters u^h, v^h *via the z-transformation as*

$$u^h(z) = \frac{h(z)}{h(z^2)} u(z), \quad v^h(z) = h(z)v(z) \tag{3.33}$$

Let $\{\varepsilon_n\}_n$ *be i.i.d.* $\mathcal{N}(0,1)$ *variables, then the process* $\{Y_n\}_n$ *defined by*

$$Y = u^h * (\uparrow_2 X) + v^h * (\uparrow_2 \varepsilon) \tag{3.34}$$

is a Gaussian process having the representation

$$Y(z) = u^h(z)X(z^2) + v^h(z)\varepsilon(z^2) = h(z)\varepsilon(z) \tag{3.35}$$

In other words, Y *has the same linear representation, or the same probability distribution, as* X*.*

Recall that a FARIMA$(0, s, 0)$ process is described via the z-transform as (**??**). Then if we define the fractional filters $u^{(s)}, v^{(s)}$ as

$$u^{(s)} = f^{(s)} * u,\ v^{(s)} = g^{(s)} * v \tag{3.36}$$

where the filters $f^{(s)}, g^{(s)}$ are defined through the z-transformations as

$$\begin{aligned} f^{(s)}(z) &= (1+z^{-1})^s = \sum_{n=0}^{\infty} f_n^{(s)} z^{-n} \\ g^{(s)}(z) &= (1-z^{-1})^{-s} = \sum_{n=0}^{\infty} g_n^{(s)} z^{-n} \end{aligned} \tag{3.37}$$

As a consequence of Proposition 3.2.5 we obtain the following Mallat-type synthesis algorithm

$$S_j^{(H)} = u^{(s)} * (\uparrow_2 S_{j-1}^{(H)}) + v^{(s)} * (\uparrow_2 \varepsilon) \tag{3.38}$$

Remark 3.2.4. On causal signals the z-transformation $(1+z^{-1})$ corresponds to the FIR filter $[1,1]$ and its inverse $(1+z^{-1})^{-1}$ corresponds to the IIR filter $[1,-1,1,-1\ldots]$. The z-transformation $(1-z^{-1})^{-1}$ corresponds to the cumulative sum and its inverse $(1-z^{-1})$ corresponds to the FIR filter $[1,-1]$.

Computation of fractional filters

The difficulty however resides in the computation of the fractional filters $u^{(s)}$ and $v^{(s)}$. In fact the coefficients of $f^{(s)}$ and $g^{(s)}$ are given by

$$\begin{aligned} f_n^{(s)} &= \frac{\Gamma(s+1)}{\Gamma(n+1)\Gamma(s-n+1)} \\ g_n^{(s)} &= (-1)^n f_n^{(-s)} = \frac{\Gamma(n+s)}{\Gamma(n+1)\Gamma(s)} \sim \frac{n^{s-1}}{\Gamma(s)}, \quad \text{as } n \to \infty, \end{aligned} \tag{3.39}$$

therefore the coefficients either diverge (when $H > 1/2$) or decay very slowly (when $H < 1/2$), making the computation of (3.36) difficult.

We can circumvent this issue by changing the original filters to u_0, v_0 which are defined by z-transformation as

$$u_0(z) = (1+z^{-1})^{-r}u(z),\ v_0(z) = (1-z^{-1})^{-r}v(z) \tag{3.40}$$

where $r > s$ is some fixed integer, such that the fractional filters are expressed as

$$u^{(s)} = f^{(s+r)} * u_0, \;\; v^{(s)} = g^{(s-r)} * v_0 \tag{3.41}$$

The asymptotic in (3.39) shows that now one can achieve any desired decay for the coefficients of the filter $f^{(s+r)}$ and $g^{(s-r)}$ by choosing accordingly the value of r.

Note that the modified filters u_0, v_0 are FIR as long as u, v are so. In fact u_0 is the r-times repeated convolution of u with the IIR $[1, -1, 1, -1 \ldots]$, while v_0 is the r-times repeated cumulative sum of v.

Wavelet synthesis of fBm

We use the Daubechies wavelet of $r = 5$ vanishing moment here, in which the filters u, v have finite length. The overall algorithm of wavelet synthesis can be resumed as follows.

- Input: the desired length N and the initial scale j_0, the fractional filters $u^{(s)}, v^{(s)}$ computed off-line via (3.41) with truncated filters $f^{(s+r)}$ and $g^{(s-r)}$ up to some precision level fixed by user.
- Initialization: at scale $j = j_0$ generate a sample trajectory of a FARIMA$(0, s, 0)$ process $S_j^{(H)}$ of length 2^j by some exact method, *e.g.* the Hosking method. Repeat the next step as long as $N > 2^j$.
- Scale $j + 1$: apply (3.38) on $S_j^{(H)}$ and to obtain $S_{j+1}^{(H)}$. Set $j \leftarrow j + 1$.
- Output: $\left\{2^{-jH}\left(S_j^{(H)}(k) - S_j^{(H)}(0)\right)\right\}_k$ which is an approximation of $\left\{B^{(H)}(2^{-j}k)\right\}_k$, for $k = 0, \ldots 2^j - 1$.

3.3 Sampling methods for mBm

We aim here at generating a sample trajectory of the mBm $(W^{(H)}(t))_{t\in\mathbb{R}}$ on the regular grid $\mathcal{G}_N$ of the interval $[T_{\min}, T_{\max}] = [0, 1]$, with a prescribed function of Hurst exponent $H(\cdot)$ which satisfies the condition (2.54) and is bounded in $[H_{\min}, H_{\max}]$. References to this subject can be found among others in [Chan and Wood, 1998, Helgason et al., 2011]. Besides the discretization of time, we will use the following discretization in H of size M:

$$\mathcal{H}_M := \left\{(H_{\max} - H_{\min}) \times \left(0, \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}\right) + H_{\min}\right\}$$

then a pair $(t, H(t))$ is converted to a point (t_j, H_{m_j}) of the grid $\mathcal{G}_N \times \mathcal{H}_M$, with

$$j = \left\lfloor \left(\frac{t - T_{\min}}{T_{\max} - T_{\min}}\right) \times N \right\rfloor, m_j = \left\lfloor \left(\frac{H(t) - H_{\min}}{H_{\max} - H_{\min}}\right) \times M \right\rfloor \tag{3.42}$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x. In this section unless specified, the notation (t_j, H_i) stands for the (j, i)-th point in the grid $\mathcal{G}_N \times \mathcal{H}_M$.

As for the fBm case, we can use the Cholesky method described in section 3.1 which consists in evaluating the covariance matrix on the grid points by (2.59) then decomposing it to obtain a sample trajectory. However other more efficient sampling methods for fBm can not be applied to mBm in general, essentially due to the loss of the stationarity of the increment process.

A simple idea is to first draw a sample of the field $(B(t;H))_{t,H}$ for $t \in [0,1]$ and $H \in [H_{\min}, H_{\max}]$ then set, according to definition, $W^{(H)}(t_j) = B(t_j; H(t_j))$. The simulation of the this random field is however a challenging problem, since its covariance function given by (2.47) seemingly lacks stationary structure (up to some transformation though). We propose here two approaches to handle this problem.

3.3.1 Methods of increment process

The idea of the first approach is to generate the increment process $\Delta_\delta W^{(H)}$ with $\delta = 1/N$ then obtain a sample trajectory of $W^{(H)}$ by accumulation:

$$W^{(H)}(t_j) = \sum_{i=0}^{j-1} \Delta_\delta W^{(H)}(t_i)$$

More precisely, we define the Gaussian vectors

$$Y = (\delta^{-H(t_j)} \Delta_\delta W^{(H)}(t_j))_{n=0\dots N-1}, \quad \text{and} \quad X = (\delta^{-H(t_j)} X_\delta^{(H(t_j))}(j)))_{j=0\dots N-1}$$

where $X_\delta^{(H(t_j))}$ is the fGn of Hurst exponent $H(t_j)$ being defined in (2.8). Then Corollary 2.5.2 says that for δ small the distribution of Y is well approximated by that of X. This suggests the following scheme of sampling.

Scheme of sampling For each $j = 0 \dots N-1$,

- first convert $H(t_j)$ to a grid point $H_{m_j} \in \mathcal{H}_M$ using (3.42).
- draw in some manner (see the remark below) a sample path of length N of $X_1^{(H_{m_j})}$, *i.e.* the standard fGn with unit step and Hurst exponent H_{m_j}, take the j-th value and set $Y_j = N^{-H_{m_j}} X_1^{(H_{m_j})}(j)$.

The vector $(Y_j)_j$ obtained in this way is then an approximation of the sample path of $\Delta_\delta W^{(H)}$ on $[0,1]$.

This scheme however is not complete and needs some explications here. Suppose that we have drawn $X_1^{(H_{m_i})}$ for $i = 0 \dots j-1$, then at step j the sample path of $X_1^{(H_{m_j})}$ must be drawn conditionalized on all the historical samples (see section 3.1), otherwise what we generated at the end (after a cumulative sum in time) would not be a "true" sample of the random field $(B(t,H))_{t\in(0,1),H\in(H_{\min},H_{\max})}$ since it would fail to have the correct covariance structure (2.47). This covariance structure implies, intuitively speaking, that the "horizontal lines" of the field are correlated (we treat t as the horizontal axis and H as the vertical axis). In particular, if the sample path $X_1^{(H_{m_j})}$ is drawn independently for all j then the resulting

random field will inherit this independence so that the "horizontal lines" are independent between them, violating in this way the true covariance structure. We propose here a conditionalized sampling method for the sequence $\left(X_1^{(H_{m_j})}\right)_j$.

Conditionalized sampling of a sequence of stationary process

For ease of notation, we denote by $\left(X^j\right)_{j=0,1\dots N-1}$ the sequence $\left(X_1^{(H_{m_j})}\right)_{j=0,1\dots N-1}$. We denote also the covariance function by $\Gamma^{j,j'}$:

$$\mathbb{E}\left(X_n^j X_{n'}^{j'}\right) = \Gamma^{j,j'}(n-n'). \tag{3.43}$$

From (2.60) it follows

$$\Gamma^{j,j'}(n-n') = D(H(t_j), H(t_{j'}))\frac{1}{2}\left(\left|n-n'+1\right|^{2\bar{H}} + \left|n-n'-1\right|^{2\bar{H}} - 2\left|n-n'\right|^{2\bar{H}}\right) \tag{3.44}$$

where $\bar{H} = (H(t_j) + H(t_{j'}))/2$, and in particular $\Gamma^{j,j}(0) = 1$. Remark that if $j > j'$ is such that $H(t_j) = H(t_{j'})$, then from (3.44) it follows

$$\mathbb{E}\left(\left\|X^j - X^{j'}\right\|^2\right) = 0$$

therefore the sample of X^j must be the replicate of the sample of $X^{j'}$.

Linear prediction Let $\mathcal{V}_j$ be the linear space spanned by $\left\{X^0 \dots, X^j\right\}$ and $\Pi_{\mathcal{V}_{j-1}}(X^j)$ be the orthogonal projection of X^j onto $\mathcal{V}_{j-1}$. The error of prediction is

$$E^j := X^j - \Pi_{\mathcal{V}_{j-1}}(X^j) \tag{3.45}$$

with the exception that $E^0 = X^0$. Whenever it is meaningful we write

$$E^j = \rho_j \varepsilon^j, \quad \text{with } \rho_j = \sqrt{\mathbb{E}\left(\left\|E^j\right\|^2\right)} \tag{3.46}$$

where $\left\|E^j\right\|^2 = \left\langle E^j, E^j\right\rangle$ and $\langle\cdot,\cdot\rangle$ denotes the inner product between two vectors. For the computation of $\Pi_{\mathcal{V}_{j-1}}(X^j)$ we have to handle the singular case. For this let $(j_n)_{n=1,2,\dots}$ be the increasing sequence of all j such that $\rho_j \neq 0$ (in particular $j_1 = 0$), and $d(j)$ be the dimension of $\mathcal{V}_j$ defined by

$$d_j := \#\left\{i | 0 \leq i \leq j \wedge \rho_i \neq 0\right\}. \tag{3.47}$$

Clearly it holds $d_{j_n} = n$ and $\mathcal{V}_j = \mathcal{V}_{j_{n-1}}$ for all $j_{n-1} \leq j < j_n$. This implies

$$d_{j_n - 1} = d_{j_{n-1}} = n - 1 \quad \text{for all } n \geq 2, \tag{3.48}$$

and $\mathcal{V}_j$ can be expressed as

$$\mathcal{V}_j = \left\{X^{j_1}, \dots X^{j_{d_j}}\right\}$$

which allows to write the orthogonal projection as:

$$\Pi_{\mathcal{V}_{j-1}}(X^j) = \sum_{n=1}^{d_{j-1}} \phi_n^j X^{j_n} \tag{3.49}$$

The coefficients of linear prediction $\boldsymbol{\phi}^j := (\phi_n^j)_{n=1\dots d_{j-1}}$ is characterized by

$$\min_{\boldsymbol{\phi}\in\mathbb{R}^{d_{j-1}}} \mathbb{E}\left(\left\| X^j - \sum_{n=1}^{d_{j-1}} \phi_n X^{j_n} \right\|^2\right).$$

and it is the unique solution to the following Yule-Walker equation:

$$\sum_{n=1}^{d_{j-1}} \underbrace{\mathbb{E}\left(\left\langle X^{j_{n'}}, X^{j_n}\right\rangle\right)}_{(\boldsymbol{G}^{d_{j-1}})_{n',n}} \phi_n^j = \underbrace{\mathbb{E}\left(\left\langle X^{j_{n'}}, X^{j}\right\rangle\right)}_{(\boldsymbol{\gamma}^j)_{n'}}, \quad \text{for } n' = 1\dots d_{j-1} \tag{3.50}$$

This is rewritten as

$$\boldsymbol{G}^{d_{j-1}}\boldsymbol{\phi}^j = \boldsymbol{\gamma}^j \tag{3.51}$$

and the symmetric matrix $\boldsymbol{G}^{d_{j-1}}$ and the vector $\boldsymbol{\gamma}^j$ are given explicitly by

$$\begin{aligned}(\boldsymbol{G}^{d_{j-1}})_{n',n} &= N\Gamma^{j_{n'},j_n}(0), \quad \text{for } n', n = 1\dots d_{j-1}\\ (\boldsymbol{\gamma}^j)_{n'} &= N\Gamma^{j_{n'},j}(0), \quad \text{for } n' = 1\dots d_{j-1}\end{aligned} \tag{3.52}$$

Change of basis On the other hand since $\left\{\varepsilon^{j_n}\right\}_{n=1,\dots d_j}$ is an orthonormal basis of $\mathcal{V}_j$, it holds

$$\mathcal{V}_j = \left\{\varepsilon^{j_1}, \dots \varepsilon^{j_{d_j}}\right\}$$

which allows to write the orthogonal projection as:

$$\Pi_{\mathcal{V}_{j-1}}(X^j) = \sum_{n=1}^{d_{j-1}} \psi_n^j \varepsilon^{j_n}, \quad \text{with } \psi_n^j := \mathbb{E}\left(\left\langle X^j, \varepsilon^{j_n}\right\rangle\right) \tag{3.53}$$

Note that

$$\mathbb{E}\left(\left\langle X^{j_l}, \varepsilon^{j_{l'}}\right\rangle\right) = \begin{cases} 0, & \text{if } l' > l \\ \rho_{j_l}, & \text{if } l' = l \\ \psi_{l'}^{j_l}, & \text{if } l' < l \end{cases}$$

and by definition of the residual it holds

$$\varepsilon^{j_n} = \rho_{j_n}^{-1}\left(X^{j_n} - \sum_{m=1}^{n-1} \psi_m^{j_n}\varepsilon^{j_m}\right), \tag{3.54}$$

injecting this into $\psi_n^j = \mathbb{E}\left(\left\langle X^j, \varepsilon^{j_n}\right\rangle\right)$ gives the following linear system

$$\begin{cases} \rho_{j_1}\psi_1^j = N\Gamma^{j,j_1}(0), \\ \sum_{m=1}^{n-1} \psi_m^{j_n}\psi_m^j + \rho_{j_n}\psi_n^j = N\Gamma^{j,j_n}(0), \qquad \text{for } n = 2, \dots d_{j-1} \end{cases}$$

where the RHS comes from (3.43). We rewrite this as

$$\boldsymbol{F}^{d_{j-1}}\boldsymbol{\psi}^j = \boldsymbol{\gamma}^j \tag{3.55}$$

where $\boldsymbol{F}^{d_{j-1}} \sim d_{j-1} \times d_{j-1}$ is a lower triangular matrix. Moreover it follows by identifying (3.49) and (3.53) the formula of change of basis

$$\boldsymbol{\psi}^j = (\boldsymbol{F}^{d_{j-1}})^\top \boldsymbol{\phi}^j \tag{3.56}$$

Covariance of error Using the coefficients of linear prediction, the covariance matrix

$$\boldsymbol{W}^j = \left(\mathbb{E}\left(E_l^j E_{l'}^j\right)\right)_{l=1\dots N, l'=1\dots N} \tag{3.57}$$

can be computed explicitly with

$$\mathbb{E}\left(E_l^j E_{l'}^j\right) = \Gamma^{j,j}(l-l') - 2\sum_{n=1}^{d_{j-1}} \phi_n^j \Gamma^{j,j_n}(l-l') + \sum_{n=1}^{d_{j-1}}\sum_{n'=1}^{d_{j-1}} \phi_n^j \phi_{n'}^j \Gamma^{j_n,j_{n'}}(l-l'). \tag{3.58}$$

Note that $\boldsymbol{W}^j$ is Toeplitz so it can be efficiently diagonalized. In particular for $j \geq 1$ the variance of the residual is

$$\rho_j^2 = \mathbb{E}\left(\left\|X^j\right\|^2\right) - (\boldsymbol{\phi}^j)^\top \boldsymbol{\gamma}^j = \mathbb{E}\left(\left\|X^j\right\|^2\right) - \left\|\boldsymbol{\psi}^j\right\|^2 \tag{3.59}$$

LU decomposition For $k \geq 1$ the vectors $\boldsymbol{\phi}^{j_{k+1}}, \boldsymbol{\psi}^{j_{k+1}} \in \mathbb{R}^k$ correspond to the projection of $X^{j_{k+1}}$ onto the space $\mathcal{V}_{j_k}$ and they are the solution to

$$\boldsymbol{G}^k \boldsymbol{\phi}^{j_{k+1}} = \boldsymbol{\gamma}^{j_{k+1}}, \quad \boldsymbol{F}^k \boldsymbol{\psi}^{j_{k+1}} = \boldsymbol{\gamma}^{j_{k+1}} \tag{3.60}$$

respectively. Define the matrix

$$\boldsymbol{A}^k := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_1^{j_2} & 1 & 0 & \dots & 0 \\ -\phi_1^{j_3} & -\phi_2^{j_3} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_1^{j_k} & -\phi_2^{j_k} & \dots & -\phi_{k-1}^{j_k} & 1 \end{pmatrix} \quad \text{for } k \geq 2, \quad \text{and } \boldsymbol{A}^1 = (1). \tag{3.61}$$

and recall that

$$\boldsymbol{F}^k = \begin{pmatrix} \rho_{j_1} & 0 & 0 & \dots & 0 \\ \psi_1^{j_2} & \rho_{j_2} & 0 & \dots & 0 \\ \psi_1^{j_3} & \psi_2^{j_3} & \rho_{j_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_1^{j_k} & \psi_2^{j_k} & \dots & \psi_{k-1}^{j_k} & \rho_{j_k} \end{pmatrix} \quad \text{for } k \geq 2, \quad \text{and } \boldsymbol{F}^1 = (\rho_{j_1}). \tag{3.62}$$

then we have the following LU decomposition of the system matrix.

Proposition 3.3.1. *For $k \geq 1$ it holds*

$$(\boldsymbol{G}^k)^{-1} = (\boldsymbol{A}^k)^\top (\boldsymbol{\Sigma}^k)^{-1} \boldsymbol{A}^k, \quad \textit{and} \quad \boldsymbol{G}^k = \boldsymbol{F}^k (\boldsymbol{F}^k)^\top \tag{3.63}$$

where $\boldsymbol{\Sigma}^k$ is the diagonal matrix with $\boldsymbol{\Sigma}^k_{nn} = \rho^2_{j_n}$ for $n = 1, \dots k$. Moreover,

$$(\boldsymbol{\Sigma}^k)^{1/2} (\boldsymbol{F}^k)^{-1} = \boldsymbol{A}^k \tag{3.64}$$

Proof. By definition of $\boldsymbol{\phi}^{j_k}$ we have the identity

$$[E^{j_1} \; E^{j_2} \; \dots E^{j_k}] = [X^{j_1} \; X^{j_2} \; \dots X^{j_k}] (\boldsymbol{A}^k)^\top$$

where the LHS is the concatenation of columns vectors and the RHS is a matrix product. Since E^j are mutually orthogonal, multiplying the LHS by its transpose at left and taking expectation yields $\boldsymbol{\Sigma}^k$. The same operation on the RHS yields $(\boldsymbol{A}^k)\boldsymbol{G}^k(\boldsymbol{A}^k)^\top$. Finally the first identity of (3.63) follows by taking inverse. For the second identity of (3.63) we write

$$[X^{j_1} \; X^{j_2} \; \dots X^{j_k}] = [\varepsilon^{j_1} \; \varepsilon^{j_2} \; \dots \varepsilon^{j_k}] (\boldsymbol{F}^k)^\top$$

and proceed in the same way. For (3.64) we observe that $\boldsymbol{F}^k$ and $\boldsymbol{A}^k$ both have a recursive structure and applying the block-inversion formula and (3.56) gives

$$(\boldsymbol{F}^k)^{-1} = \left(\begin{array}{c|c} (\boldsymbol{F}^{k-1})^{-1} & 0 \\ \hline -\rho_{j_k}^{-1} (\boldsymbol{\phi}^{j_k})^\top & \rho_{j_k}^{-1} \end{array} \right) \tag{3.65}$$

Therefore (3.64) follows by comparing $(\boldsymbol{F}^k)^{-1}$ with $\boldsymbol{A}^k$. □

Conditionalized sampling of $(X^j)_{j=0,\dots N-1}$ Assembling the elements above gives the following algorithm.

- Initialization: Generate a sample $\hat{X}^0$ from X^0 via Circulant embedding, and set
  - $k = 1$, $j_1 = 0$, $\rho_0 = \sqrt{N}$
  - sequence $\mathcal{J} = \{j_1\}$, $\mathcal{X} = \left\{\hat{X}^0\right\}$
  - matrix $(\boldsymbol{F}^1)^{-1} = (\rho_0^{-1})$
- Iteration $j > 0$: Given the samples $\mathcal{X} = \left\{\hat{X}^0, \dots \hat{X}^{j-1}\right\}$,
  - if there exists $j' < j$ such that $H_{m_j} = H_{m_{j'}}$, set $\hat{X}^j = \hat{X}^{j'}$
  - otherwise, compute by order

$$\boldsymbol{\gamma}^j = (N \Gamma^{j_n, j}(0))_{n=1,\dots k} \tag{3.66}$$

$$\boldsymbol{\psi}^j = (\boldsymbol{F}^k)^{-1} \boldsymbol{\gamma}^j \tag{3.67}$$

$$\boldsymbol{\phi}^j = (\boldsymbol{F}^k)^{-\top} \boldsymbol{\psi}^j \tag{3.68}$$

$$\boldsymbol{\mu}^j = \sum_{n=1}^{k} \phi_n^j \hat{X}^{j_n} \tag{3.69}$$

$$\rho_j = \left(N - \left\| \boldsymbol{\psi}^j \right\|^2 \right)^{1/2} \tag{3.70}$$

* if $\rho_j = 0$, set $\hat{X}^j = \boldsymbol{\mu}^j$
* otherwise construct the matrix $\boldsymbol{W}^j$ using (3.58) (where the term d_{j-1} is replaced by k) and draw a sample $\hat{E}^j$ from the distribution $\mathcal{N}(0, \boldsymbol{W}^j)$ via Circulant embedding and set $\hat{X}^j = \boldsymbol{\mu}^j + \hat{E}^j$. Update by order

$$k \leftarrow k + 1 \tag{3.71}$$

$$j_k = j, \mathcal{J} \leftarrow \mathcal{J} \cup \{j_k\} \tag{3.72}$$

$$(\boldsymbol{F}^k)^{-1} \text{ according to (3.65)} \tag{3.73}$$

Update $\mathcal{X} \leftarrow \mathcal{X} \cup \left\{\hat{X}^j\right\}$ and set $j \leftarrow j + 1$. If $j < M$ jump to the next iteration otherwise terminate.

- Output: sequence of samples $\mathcal{X} = \left\{\hat{X}^0, \dots \hat{X}^{N-1}\right\}$.

Remark 3.3.1. The innovation equations for $\boldsymbol{\psi}^j, \boldsymbol{\phi}^j$ are not unique. For example we could use

$$\boldsymbol{\phi}^j = (\boldsymbol{A}^k)^\top (\boldsymbol{\Sigma}^k)^{-1} \boldsymbol{A}^k \boldsymbol{\gamma}^j \tag{3.74}$$

$$\boldsymbol{\psi}^j = (\boldsymbol{F}^k)^\top \boldsymbol{\phi}^j \tag{3.75}$$

and update $\boldsymbol{A}^k, \boldsymbol{F}^k$ rather than $(\boldsymbol{F}^k)^{-1}$. Note that in this case the variable $\boldsymbol{\psi}^j$ becomes unnecessary.

Final algorithm

Now one can inject the conditionalized sampling into the scheme mentioned at the beginning of section 3.3.1. Since $H(t_j)$ is converted to a grid point and the conditional sampling does not depend on the chronological order, we can use the algorithm of conditionalized sampling to draw from the sequence $(X_1^{(H_i)})_{i=0,\dots M-1}$, *i.e.* with strictly increasing Hurst exponent of the grid $\mathcal{H}_M$, then recover the desired increment process $(Y_j)_j$ as before.

This algorithm has the time complexity $O(MN^2)$ which is faster than Cholesky for small M, and in practice it may diverge when the grid $\mathcal{H}_M$ becomes too fine, *e.g.* with a step ≤ 0.025. On a coarse grid it allows to simulate sample paths of length $N = 10^5$ for which Cholesky may fail due to limitation of resources.

3.3.2 Series expansion method

Another approach is to simulate the random field $(B(t; H))_{t \in \mathbb{R}_+, H \in (0,1)}$ via the series expansion (2.34), which amounts to compute the coefficients of the representation kernel under some fixed $L^2(\mathbb{R})$ basis. Consider for example an orthonormal wavelet basis: the series expansion then has to be truncated up to some scale j_0 and some position N_0 and we need to characterize the truncation error, furthermore for the simulation of long trajectory the truncated series has to be evaluated efficiently. In the following we will first explain the Fourier-Wavelet method which allows to establish bounds on the truncation error. Then we propose two efficient methods to evaluate approximately the coefficients of series expansion.

In this section we follow the convention (1.1) of the Fourier transform. Given an arbitrary $f \in L^2(\mathbb{R})$ we will write $f^j(\cdot) = 2^{j/2} f(2^j \cdot)$ its dilation at the scale $j \in \mathbb{Z}$ and $f_n^j(\cdot) = f^j(\cdot - 2^{-j} n)$ its translation at the position indexed by n.

Fourier-Wavelet method

We consider an orthonormal wavelet basis generated by a scaling function ϕ and a mother wavelet function ψ fulfilling the partition of unity: for any $j_0 \in \mathbb{Z}$ it holds

$$\left|\widehat{\phi^{j_0}}(\xi)\right|^2 + \sum_{j \geq j_0} 2^j \left|\hat{\psi}^j(\xi)\right|^2 = \frac{1}{2\pi}, \quad \text{for almost every } \xi \in \mathbb{R} \tag{3.76}$$

and the family

$$\left\{\phi_n^{j_0}, \text{ for } n \in \mathbb{Z}\right\} \bigcup \left\{\psi_n^j, \text{ for } j \geq j_0, n \in \mathbb{Z}\right\} \tag{3.77}$$

constitutes an orthonormal basis of $L^2(\mathbb{R})$. Moreover we suppose that for some integer p the first p derivatives of $\hat{\phi}$ and $\hat{\psi}$ are integrable, so that $\phi(t)$ and $\psi(t)$ both decay at infinity as $O(1+|t|)^{-p}$ by Paley-Wiener theorem.

Meyer wavelets We will consider the Meyer wavelet for reason of simpleness. Its scaling function ϕ is defined in frequency domain as:

$$\hat{\phi}(\xi) = \frac{1}{\sqrt{2\pi}} \begin{cases} 1, & \text{if } |\xi| \leq \frac{2\pi}{3} \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\xi| - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3} \\ 0 & \text{otherwise} \end{cases} \tag{3.78}$$

where the function ν satisfies

$$\nu(\xi) = \begin{cases} 0, & \text{if } \xi \leq 0 \\ 1, & \text{if } \xi \geq 1 \\ \text{transition from 0 to 1}, & \text{otherwise} \end{cases} \tag{3.79}$$

and $\nu(\xi) + \nu(1-\xi) = 1$. Its mother wavelet ψ is constructed from ϕ and defined in frequency domain as:

$$\hat{\psi}(\xi) = \frac{1}{\sqrt{2\pi}} \begin{cases} e^{i\xi/2} \sin\left(\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\xi| - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3} \\ e^{i\xi/2} \cos\left(\frac{\pi}{2}\nu\left(\frac{3}{4\pi}|\xi| - 1\right)\right) & \text{if } \frac{4\pi}{3} \leq |\xi| \leq \frac{8\pi}{3} \\ 0 & \text{otherwise} \end{cases} \tag{3.80}$$

Note that both ϕ, ψ are real, even and smooth function.

Kernel of representation We will explain the method with the first kernel mentioned in section 2.3.1, *i.e.*

$$K(x) = K(t, x; H) := a_0 \left(|t-x|^{H-1/2} - |x|^{H-1/2}\right) \tag{3.81}$$

where a_0 is defined in (2.44), and recall its Fourier transform

$$\hat{K}(\xi) = \left(\frac{\Gamma(2H+1)\sin(\pi H)}{2\pi}\right)^{1/2} \frac{e^{it\xi} - 1}{|\xi|^{H+1/2}}. \tag{3.82}$$

Bounds on the truncation error Using the family (3.77) the field $B(t,H)$ can be expressed as

$$B(t,H) = \sum_{n\in\mathbb{Z}} \langle K, \phi_n^{j_0}\rangle Z_n + \sum_{j\geq j_0}\sum_{n'\in\mathbb{Z}} \left\langle K, \psi_{n'}^{j}\right\rangle Z_{n'}^{j}$$

where Z_n and $Z_{n'}^j$ are all i.i.d. standard Gaussian random variables. As an approximation we consider the following field

$$\tilde{B}(t,H) := \sum_{n=-N_0}^{N_0} \langle K, \phi_n^{j_0}\rangle Z_n \tag{3.83}$$

for some fixed integer N_0, with the error of truncation

$$\mathbb{E}\left(\left|B(t,H)-\tilde{B}(t,H)\right|^2\right) = \sum_{|n|>N_0} \left|\langle K, \phi_n^{j_0}\rangle\right|^2 + \sum_{j\geq j_0}\sum_{n'\in\mathbb{Z}} \left|\left\langle K, \psi_{n'}^{j}\right\rangle\right|^2 = E_1 + E_2. \tag{3.84}$$

We establish here the bounds on this error which will guide the choice of j_0 and N_0.

Sampling operator Let S^j be the sampling operator applying on a smooth test function g such that in the sense of distribution

$$S^j g(x) = \sum_{n\in\mathbb{Z}} g(2^{-j}n)\delta(x-2^{-j}n) \tag{3.85}$$

and its Fourier transform reads, using the Poisson summation formula,

$$\widehat{S^j g}(\xi) = 2^j \sum_{k\in\mathbb{Z}} \hat{g}(\xi - 2\pi 2^j k) \tag{3.86}$$

Bound on E_2 By the property of orthonormal basis, it holds

$$E_2 = \|K\|^2 - \sum_{n\in\mathbb{Z}} \left|\langle K, \phi_n^{j_0}\rangle\right|^2$$

Remark that ϕ is even and the convolution $K * \phi^{j_0}$ is smooth, so it holds in the L^2 sense

$$\sum_{n\in\mathbb{Z}} \langle K, \phi_n^{j_0}\rangle \phi_n^{j_0} = (S^{j_0}(K*\phi^{j_0}))*\phi^{j_0} \tag{3.87}$$

and the Parseval's identity together with (3.86) gives

$$\sum_{n\in\mathbb{Z}} \left|\langle K, \phi_n^{j_0}\rangle\right|^2 = \left\|\sum_{n\in\mathbb{Z}} \langle K, \phi_n^{j_0}\rangle \phi_n^{j_0}\right\|_{L^2}^2 = 4\pi^2 2^{2j_0} \left\|\widehat{\phi^{j_0}} \sum_{|k|\leq 1} (\hat{K}\widehat{\phi^{j_0}})(\cdot - 2\pi 2^{j_0}k)\right\|_{L^2}^2$$

where the finite sum in the last term is due to the compact support of $\widehat{\phi^{j_0}}$ and does not introduce aliasing on the interval $[-2\pi 2^{j_0}/3, 2\pi 2^{j_0}/3]$. This gives

$$\sum_{n\in\mathbb{Z}} \left|\left\langle K, \phi_n^{j_0}\right\rangle\right|^2 \geq (4\pi)^2 \int_0^{\frac{2\pi}{3}2^{j_0}} \frac{\left|\sin(\frac{1}{2}t\xi)\right|^2}{|\xi|^{2H+1}}\, d\xi$$

and finally we obtain the bound

$$E_2 \lesssim \int_{\frac{2\pi}{3}2^{j_0}}^{+\infty} \frac{\left|\sin(\frac{1}{2}t\xi)\right|^2}{|\xi|^{2H+1}}\, d\xi \leq C_2 2^{-2j_0 H} \tag{3.88}$$

where the constant factor C_2 depends only on H.

Bound on E_1 First note that $\hat{K}\widehat{\phi^{j_0}}$ is not integrable for $H > 1/2$ and $\hat{K}$ is not differentiable at 0. The last statement comes from the observation that for any $t \neq 0$

$$\lim_{\xi\to 0} \frac{\left|e^{it\xi} - 1\right|}{|\xi|^h} = 2 \lim_{\xi\to 0^+} \frac{\sin(\frac{1}{2}t\xi)}{\xi^h} = \begin{cases} +\infty & \text{if } h > 1, \\ t & \text{if } h = 1, \\ 0 & \text{if } h < 1. \end{cases}$$

Therefore the Payley-Wiener theory can not be applied directly here to characterize the asymptotic decay of the coefficient $\left\langle K, \phi_n^{j_0}\right\rangle$. We proceed hence by direct computation for the case $H \neq 1/2$ (the case $H = 1/2$ can be treated similarly). Note that for x large

$$K(t, x-y; H) = (1/2 - H)t \operatorname{sign}(x) |x|^{H-3/2} \left(1 - \frac{y + \theta t}{x}\right)^{H-3/2}$$

where the constant $\theta \in (0,1)$ depends on y and t, and

$$\lim_{x\to\infty} \int_{\mathbb{R}} \left(1 - \frac{y+\theta t}{x}\right)^{H-3/2} \phi^{j_0}(y)\, dy = \sqrt{2\pi} 2^{-j_0/2}$$

therefore for x large $K * \phi^{j_0}(x) = O(|x|^{H-3/2})$. Finally we obtain

$$E_1 = \sum_{|n|>N_0} \left|K * \phi^{j_0}(2^{-j_0}n)\right|^2 \lesssim 2^{-2j_0(H-1)} \sum_{|n|>N_0} |n|^{2H-3} \leq C_1 (2^{-j_0} N_0)^{2(H-1)}. \tag{3.89}$$

where the constant factor C_1 depends on t and H.

Computation of basis coefficients

–To Do–Suppose j_0, N_0 are set according to the error bounds, now we aim at computing the coefficients $\left\langle K, \phi_n^{j_0}\right\rangle$. First write

$$\left\langle \hat{K}(t, \cdot; H), \widehat{\phi_n^{j_0}}(\cdot)\right\rangle = \frac{1}{2^{j_0 H}} \frac{1}{\sqrt{2\pi}} (I_1^\phi + I_2^\phi)$$

with the integrals

$$I_1^{\phi} := \int_{|\xi| \in [0, \frac{2\pi}{3}]} \frac{e^{i(2^{j_0}t-n)\xi} - e^{-in\xi}}{|\xi|^{H+1/2}}\, d\xi = 2\int_0^{\frac{2\pi}{3}} \frac{\cos\left((2^{j_0}t-n)\xi\right) - \cos(n\xi)}{\xi^{H+1/2}}\, d\xi,$$

$$I_2^{\phi} := 2\int_{\frac{2\pi}{3}}^{\frac{4\pi}{3}} \frac{\cos\left((2^{j_0}t-n)\xi\right) - \cos(n\xi)}{\xi^{H+1/2}} \cos\left(\frac{\pi}{2}\nu\left(\frac{3}{2\pi}|\xi| - 1\right)\right)\, d\xi$$

and note that the last integrand in I_1^{ϕ} is well defined at 0. Although these integrals can be evaluated efficiently via some numerical integration method, such a method is impracticable for the simulation of long trajectory.

Chapter 4

Parameter estimation in fBm

We use the same notations as in Chapter 2. Recall that $B^{(H)}$ is a pure fBm with the volatility $\sigma^2 = \mathbb{E}\left(\left|B^{(H)}(1)\right|^2\right)$. Given discrete observations $(B^{(H)}(n\delta))_{n\in\mathbb{Z}}$ with the sampling step $\delta > 0$, we aim to make estimations of the Hurst exponent H as well as the volatility σ^2. However the volatility σ can not be extracted without knowing the sampling step δ, therefore in the following we will consider the set of parameters $\boldsymbol{\theta} = \{H, v\}$ with

$$v := \sigma\delta^H \tag{4.1}$$

being the scaled volatility. The estimation of v can be easily obtained from that of H, see e.g. [Brouste and Iacus, 2013, Brouste and Fukasawa, 2016] for detailed results on the property of joint estimator.

4.1 Overview

Many popular methods have been proposed for estimation of Hurst exponent, see *e.g.* [] and references therein. Generally speaking the information about H can be extracted by exploiting 1) the time-dependency which is a consequence of the autocovariance structure, or 2) the scale-dependency which is a consequence of the self-similarity. Corresponding to the two types are two classes of estimators:

1. maximum likelihood-type which uses time information only
2. the scalogram-type which uses scale information only

We will focus on these two classes of estimators and propose some generalizations.

4.2 Maximum-Likelihood Estimator

We present first the general framework of MLE, then adapt it to different type observations constructed from the same sample trajectory which gives different variations of MLE. For general properties of MLE, see *e.g.* [Lundahl et al., 1986].

4.2.1 General framework

Let $\boldsymbol{x} \in \mathbb{R}^d$ be a centered multivariate Gaussian vector with the invertible covariance matrix $\mathbf{C} = \mathbf{C}(\boldsymbol{\theta})$ that depends on a set of parameters $\boldsymbol{\theta}$. The pdf of $\boldsymbol{x}$ reads

$$\mathbf{p}\left(\boldsymbol{x}|\mathbf{C}\right) = (2\pi)^{-d/2} \left|\det \mathbf{C}\right|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{x}^\top \mathbf{C}^{-1}\boldsymbol{x}\right). \tag{4.2}$$

and the log-pdf is

$$\ell(\boldsymbol{\theta}; \boldsymbol{x}) := \log \mathbf{p}\left(\boldsymbol{x}|\mathbf{C}\right) = -\frac{1}{2}\left(\boldsymbol{x}^\top \mathbf{C}^{-1}\boldsymbol{x} + \log\left|\det \mathbf{C}\right| + d\log(2\pi)\right) \tag{4.3}$$

Given independent observations $\boldsymbol{X} = \{\boldsymbol{x}_1, \dots \boldsymbol{x}_N\}$, we define the MLE of $\boldsymbol{\theta}$ as

$$\hat{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; \boldsymbol{X}). \tag{4.4}$$

with the sample log-likelihood

$$L(\boldsymbol{\theta}; \boldsymbol{X}) := \sum_{n=1}^{N} \ell(\boldsymbol{\theta}; \boldsymbol{x}_n) \tag{4.5}$$

Partial derivatives Numerical solution of (4.4) requires the (first and second order) partial derivatives of $\ell(\boldsymbol{\theta}; \boldsymbol{x})$ which can be obtained using the rules

$$\begin{cases} \dfrac{\partial}{\partial\theta}\left(\mathbf{C}^{-1}\right) = -\mathbf{C}^{-1}\left(\dfrac{\partial}{\partial\theta}\mathbf{C}\right)\mathbf{C}^{-1} \\ \dfrac{\partial}{\partial\theta}\left(\log\det\mathbf{C}\right) = \operatorname{trace}\left(\mathbf{C}^{-1}\dfrac{\partial}{\partial\theta}\mathbf{C}\right) \end{cases} \tag{4.6}$$

Namely, the first order partial derivative reads

$$\frac{\partial\ell}{\partial\theta_i}(\boldsymbol{\theta}; \boldsymbol{x}) = -\frac{1}{2}\operatorname{trace}\left(\left(\mathbf{C}^{-1} - \boldsymbol{\xi}\boldsymbol{\xi}^\top\right)\frac{\partial\mathbf{C}}{\partial\theta_i}\right) \tag{4.7}$$

and the second order partial derivative reads

$$\frac{\partial^2\ell}{\partial\theta_i\partial\theta_j}(\boldsymbol{\theta}; \boldsymbol{x}) = -\frac{1}{2}\operatorname{trace}\left(\left(\mathbf{C}^{-1} - \boldsymbol{\xi}\boldsymbol{\xi}^\top\right)\left(\frac{\partial^2\mathbf{C}}{\partial\theta_i\partial\theta_j} - \mathbf{A}_{ji}\right) + \boldsymbol{\xi}\boldsymbol{\xi}^\top\mathbf{A}_{ij}\right) \tag{4.8}$$

where we introduced the vector $\boldsymbol{\xi} := \mathbf{C}^{-1}\boldsymbol{x}$ and the matrix

$$\mathbf{A}_{ij} := \frac{\partial\mathbf{C}}{\partial\theta_i}\left(\mathbf{C}^{-1}\right)\frac{\partial\mathbf{C}}{\partial\theta_j}. \tag{4.9}$$

It can be checked for (4.8) that $\partial_{\theta_i}\partial_{\theta_j}\ell = \partial_{\theta_j}\partial_{\theta_i}\ell$, although in general $\mathbf{A}_{ij} \neq \mathbf{A}_{ji}$.

Joint estimation This general framework can be simplified for the joint estimation of the scaled volatility v and the Hurst exponent H. The covariance matrix $\mathbf{C}$ above now reads

$$\mathbf{C} = v^2 \mathbf{\Gamma}$$

where the matrix $\mathbf{\Gamma} = \mathbf{\Gamma}(H)$ depends on the type of observations constructed from the sample trajectory. The sample log-likelihood now reads

$$L(H, v; \boldsymbol{X}) = -\frac{1}{2} \sum_{n=1}^{N} \left(v^{-2} \boldsymbol{x}_n^\top \mathbf{\Gamma}^{-1} \boldsymbol{x}_n + d \log v^2 + \log |\det \mathbf{\Gamma}| + d \log(2\pi) \right)$$

and the optimal v is first obtained by setting $\partial_v L(H, v; \boldsymbol{X}) = 0$:

$$v^2 = \frac{1}{Nd} \sum_{n=1}^{N} \boldsymbol{x}_n^\top \mathbf{\Gamma}^{-1} \boldsymbol{x}_n. \tag{4.10}$$

Injecting this back gives the H-dependent log-likelihood

$$L(H; \boldsymbol{X}) = -\frac{1}{2} \left(Nd \log \left(\sum_{n=1}^{N} \boldsymbol{x}_n^\top \mathbf{\Gamma}^{-1} \boldsymbol{x}_n \right) + N \log |\det \mathbf{\Gamma}| \right) - \frac{Nd}{2} \log \left(\frac{2\pi e}{Nd} \right) \tag{4.11}$$

and the optimal estimation of H is defined as

$$\hat{H} := \underset{H \in (0,1)}{\arg\max}\ L(H; \boldsymbol{X}). \tag{4.12}$$

In practice the problem (4.12) can be easily solved using for example Brent's method.

4.2.2 fWn bank-MLE

In this approach the observation $\boldsymbol{x} = (x_j)_j \in \mathbb{R}^J$ is a bank of fractional Wavelet noise as defined in Section 2.7.1, *i.e.* the filtration of fBm by a filter bank $\{\psi_1, \dots \psi_J\}$ satisfying (2.88). Namely, at time t the j-th dimension of the coefficient vector is

$$x_j(t) := \sum_{k \in \mathbb{Z}} \psi_j[k] B^{(H)}(t + k\delta) \tag{4.13}$$

From (2.93) the covariance between two scales i, j and two positions $m\delta, n\delta$ is given by:

$$\mathbb{E}\left(x_i((n+l)\delta) x_j(n\delta)\right) = v^2 \gamma^{\psi_i, \psi_j}(l) \tag{4.14}$$

where γ^{ψ_i, ψ_j} denotes the covariance function (2.92). Using (2.92) we only need to compute the coefficients $(a[k])_k$ in (2.91) in order to evaluate this expression. We will consider l adjacent positions and introduce the J-by-J covariance matrix $\mathbf{\Gamma}^l$ with the (i, j)-th coefficient

$$\mathbf{\Gamma}^l_{i,j} := \gamma^{\psi_i, \psi_j}(l). \tag{4.15}$$

Then the concatenation of l adjacent observation vectors $\boldsymbol{x}$ follows a multivariate Gaussian distribution $\mathcal{N}(0, v^2\boldsymbol{\Gamma})$ with $\boldsymbol{\Gamma}$ being constructed blockwisely as

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}^0 & \boldsymbol{\Gamma}^1 & \dots & \boldsymbol{\Gamma}^{l-1} \\ \boldsymbol{\Gamma}^{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{\Gamma}^1 \\ \boldsymbol{\Gamma}^{1-l} & \dots & \boldsymbol{\Gamma}^{-1} & \boldsymbol{\Gamma}^0 \end{pmatrix} \tag{4.16}$$

which is symmetric thanks to the relation $\boldsymbol{\Gamma}^l = \left(\boldsymbol{\Gamma}^{-l}\right)^\top$.

The fWn bank-framework presented above is general, and as particular cases we have

- fGn-MLE: we take the filter bank $\{\psi_j\}$ with ψ_j being the filter in (2.98) which computes the increment process (fGn) at the scale j.
- B-Spline-MLE: we use the filter bank of the order q B-Spline wavelet $\psi^{(q)}$ as defined in section 2.7.3.

Remark 4.2.1. An important issue concerning the practical use of the fWn bank-MLE method is the cross-scale interference. The framework presented here has been established on a pure fBm which is scale invariant, however in practice the long memory behavior of data is often scale dependent. This means in particular the theoretical covariance function (4.14) will fail to hold if 1) the scales i, j are too distinct 2) on a large scale the distinct positions m, n are too close to each other. Therefore it is better to work only with adjacent scales and downsample the coefficients (4.13) in order to avoid cross scale interference.

4.3 Power law method

Recall the p-th moment ($p \geq 1$) of fGn:

$$\mathbb{E}\left(\left|B^{(H)}(t+\delta) - B^{(H)}(t)\right|^p\right) = c_p\left(\sigma\delta^H\right)^p \tag{4.17}$$

where the constant c_p is

$$c_p := \sqrt{\frac{2^p}{\pi}}\,\Gamma\left(\frac{p+1}{2}\right) \tag{4.18}$$

with Γ being the Gamma function. For the fGn at the step $k\delta$ for integer $k = 1, \dots K$ the p-th logarithmic moment reads,

$$\mathbf{y}_k := \log\mathbb{E}\left(\left|B^{(H)}((n+k)\delta) - B^{(H)}(n\delta)\right|^p\right) = \underbrace{\left(\log c_p + p\log(\sigma\delta^H)\right)}_{=:\eta} + \underbrace{(p\log k)}_{=:\mathbf{x}_k} H \tag{4.19}$$

which does not depend on the index n since fGn is stationary. We then form the K-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ such that

$$H\mathbf{x} + \eta\mathbf{1} = \mathbf{y} \tag{4.20}$$

By definition of η it holds

$$v = \sigma \delta^H = e^{(\eta - \log c_p)/p} \tag{4.21}$$

In practice $\mathbf{y}$ can be computed by e.g. the sample mean. Once some estimation of H is obtained, we compute v via (4.21).

Estimator for H A standard way for the estimation of H is via linear regression of (4.20), which does not guarantee $0 < H < 1$. As a remede, we propose to replace the linear regression step by a constrained optimization problem

$$\min_{H \in (0,1), \eta} \sum_{k=1}^{K} |\mathbf{y}_k - (H\mathbf{x}_k + \eta)|^2$$

Note that the optimizer of η satisfies

$$\eta = \frac{1}{K} \sum_{k=1}^{K} (\mathbf{y}_k - H\mathbf{x}_k)$$

hence the joint optimization above is reduced to

$$\min_{H \in (0,1)} \sum_{k=1}^{K} |\tilde{\mathbf{y}}_k - H\tilde{\mathbf{x}}_k|^2 \tag{4.22}$$

with $\tilde{\mathbf{x}}_k := \mathbf{x}_k - K^{-1} \sum_{k=1}^{K} \mathbf{x}_k$ and similarly for the vector $\tilde{\mathbf{y}}_k$.

4.4 Scalogram method

Let $\{s_1, s_2, \dots s_J\}$ be distinct integer scales used in DCWT. In particular for the B-Spline DCWT introduced in section 2.7.2 the scales are even number, *e.g.* $s_i = 2i$. According to Proposition 2.7.1 the covariance function at time-lag l and at the scale index i, j is independent of the time index n and reads

$$\mathbb{E}\left(w_X^{\psi}[n; s_i]\, w_X^{\psi}[n + l; s_j] \right) \simeq v^2 \mathbf{\Gamma}^l_{i,j} \tag{4.23}$$

where $\mathbf{\Gamma}^l_{i,j}$ is the (i,j)-th coefficient of the matrix $\mathbf{\Gamma}^l$:

$$\mathbf{\Gamma}^l_{i,j} := A^{\psi}_{\rho}\left(\frac{l}{\sqrt{s_i s_j}}, H \right) \left(\sqrt{s_i s_j}\right)^{2H+1}, \quad \text{with } \rho = s_j/s_i \tag{4.24}$$

The *scalogram* refers to the diagonal coefficients of the covariance matrix Γ^0, *i.e.*

$$\mathbf{\Gamma}^0_{j,j} = A^{\psi}(H) s_j^{2H+1} \tag{4.25}$$

and represent the energy of the signal at different scales. More generally the p-th moment of the wavelet coefficients reads

$$\mathbb{E}\left(\left| w_X^{\psi}[n; s_j] \right|^p \right) \simeq c_p v^p \left(\mathbf{\Gamma}^0_{j,j} \right)^{p/2}$$

with the same constant c_p of (4.18). In order to extract H we define the vector $\mathbf{y}$ as

$$\mathbf{y}_j := \log \mathbb{E}\left(\left|w_X^{\psi}[n; s_j]\right|^p\right) \simeq (H+1/2)\underbrace{\log(s_j^p)}_{=:\mathbf{x}_j} + \underbrace{\log c_p + (p/2)\log A^{\psi}(H) + p\log v}_{=:\eta}$$

Then form the vectors $\mathbf{x}$ such that

$$(H+1/2)\mathbf{x} + \eta\mathbf{1} = \mathbf{y} \tag{4.26}$$

By definition of η it holds

$$v = e^{\left(\eta - (p/2)\log A^{\psi}(H) - \log c_p\right)/p} \tag{4.27}$$

The vector $\mathbf{y}$ is constructed from data by computing the sample covariance matrix and the estimation of H and v is obtained via linear regression of (4.26), or using a similar optimization procedure as in Power law estimator.

The above scalogram estimator can be easily generalized by observing that oblique off-diagonals in (4.24) also allows to extract H, we skip the detail here.

Bibliography

[Abry, 2003] Abry, P. (2003). Scaling and Wavelets: an introductory Walk. In *Processes with Long-Range Correlations*, pages 34–60. Springer. 17

[Abry et al., 2003] Abry, P., Flandrin, P., Taqqu, M. S., and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens. *Theory and applications of long-range dependence*, pages 527–556. 17

[Abry et al., 1995] Abry, P., Gonçalvés, P., and Flandrin, P. (1995). Wavelets, spectrum analysis and 1/f processes. In *Wavelets and statistics*, pages 15–29. Springer. 7, 17

[Abry and Sellan, 1996] Abry, P. and Sellan, F. (1996). The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation. *Applied and computational harmonic analysis*, 3(4):377–383. 34

[Ayache et al., 2000] Ayache, A., Cohen, S., and Véhel, J. L. (2000). The covariance structure of multifractional Brownian motion, with application to long range dependence. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, volume 6, pages 3810–3813. IEEE. 15

[Balança, 2014] Balança, P. (2014). Some sample path properties of multifractional Brownian motion. *arXiv:1408.0317 [math]*. arXiv: 1408.0317. 15

[Baraniuk and Crouse, 1999] Baraniuk, R. and Crouse, M. (1999). Fast, Exact Synthesis of Gaussian and nonGaussian Long-Range-Dependent Processes. 28

[Bardet et al., 2003] Bardet, J.-M., Lang, G., Oppenheim, G., Philippe, A., and Taqqu, M. S. (2003). Generators of long-range dependent processes: a survey. *Theory and applications of long-range dependence*, pages 579–623. 28

[Brouste and Fukasawa, 2016] Brouste, A. and Fukasawa, M. (2016). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *arXiv:1610.03694 [math, stat]*. arXiv: 1610.03694. 48

[Brouste and Iacus, 2013] Brouste, A. and Iacus, S. M. (2013). Parameter estimation for the discretely observed fractional Ornstein–Uhlenbeck process and the Yuima R package. *Computational Statistics*, 28(4):1529–1547. 48

[Chan and Wood, 1998] Chan, G. and Wood, A. T. A. (1998). Simulation of Multifractional Brownian Motion. In *COMPSTAT*, pages 233–238. Physica, Heidelberg. 28, 37

[Coeurjolly, 2000] Coeurjolly, J.-F. (2000). Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study. 28

[Craigmile, 2003] Craigmile, P. F. (2003). Simulating a class of stationary Gaussian processes using the Davies–Harte algorithm, with application to long memory processes. *Journal of Time Series Analysis*, 24(5):505–511. 28

[Dieker, 2004] Dieker, T. (2004). Simulation of fractional Brownian motion. *MSc theses, University of Twente, Amsterdam, The Netherlands*. 28

[Dietrich and Newsam, 1997] Dietrich, C. and Newsam, G. (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. 28

[Doukhan et al., 2002] Doukhan, P., Oppenheim, G., and Taqqu, M. (2002). *Theory and Applications of Long-Range Dependence*. Springer Science & Business Media. Google-Books-ID: xIrDf3DAnKMC. 6

[Flandrin, 1992] Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Transactions on information theory*, 38(2):910–917. 17

[Helgason et al., 2011] Helgason, H., Pipiras, V., and Abry, P. (2011). Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding. *Signal Processing*, 91(5):1123–1133. 37

[Hosking, 1984] Hosking, J. R. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water resources research*, 20(12):1898–1908. 14, 30, 31

[Hosking, 1996] Hosking, J. R. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics*, 73(1):261–284. 7

[Jin et al., 2018] Jin, S., Peng, Q., and Schellhorn, H. (2018). Estimation of the pointwise Hölder exponent of hidden multifractional Brownian motion using wavelet coefficients. *Statistical Inference for Stochastic Processes*, 21(1):113–140. 15, 17

[Keshner, 1982] Keshner, M. S. (1982). 1/f noise. *Proceedings of the IEEE*, 70(3):212–218. 7

[Lau et al., 1995] Lau, W.-C., Erramilli, A., Wang, J. L., and Willinger, W. (1995). Self-similar traffic generation: The random midpoint displacement algorithm and its properties. In *Communications, 1995. ICC'95 Seattle,'Gateway to Globalization', 1995 IEEE International Conference on*, volume 1, pages 466–472. IEEE. 31

[Lundahl et al., 1986] Lundahl, T., Ohley, W. J., Kay, S. M., and Siffert, R. (1986). Fractional Brownian motion: A maximum likelihood estimator and its application to image texture. *IEEE transactions on Medical Imaging*, 5(3):152–161. 48

[Meyer et al., 1999] Meyer, Y., Sellan, F., and Taqqu, M. S. (1999). Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *Journal of Fourier Analysis and Applications*, 5(5):465–494. 34

[Norros et al., 1999] Norros, I., Mannersalo, P., and Wang, J. L. (1999). Simulation of fractional Brownian motion with conditionalized random midpoint displacement. *Advances in Performance Analysis*, 2(1):77–101. 31

[Peltier and Véhel, 1995] Peltier, R.-F. and Véhel, J. L. (1995). Multifractional Brownian Motion : Definition and Preliminary Results. report, INRIA. 15

[Pipiras, 2004] Pipiras, V. (2004). On the usefulness of wavelet-based simulation of fractional Brownian motion. *Preprint. Available at http://www. stat. unc. edu/faculty/pipiras*. 34

[Pipiras, 2005] Pipiras, V. (2005). Wavelet-based simulation of fractional Brownian motion revisited. *Applied and Computational Harmonic Analysis*, 19(1):49–60. 34

[Roueff and Von Sachs, 2011] Roueff, F. and Von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Processes and their Applications*, 121(4):813–844. 17

[Samorodnitsky, 2002] Samorodnitsky, G. (2002). Long range dependence, heavy tails and rare events. Technical report, Cornell University Operations Research and Industrial Engineering. 7

[Samorodnitsky and Taqqu, 1994] Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. CRC Press. Google-Books-ID: wTTUfYwjksAC. 6

[Shusterman and Feder, 1996] Shusterman, E. and Feder, M. (1996). Analysis of 1/f processes. In *Time-Frequency and Time-Scale Analysis, 1996., Proceedings of the IEEE-SP International Symposium on*, pages 201–204. IEEE. 7

[Stoev and Taqqu, 2006] Stoev, S. A. and Taqqu, M. S. (2006). How rich is the class of multifractional Brownian motions? *Stochastic Processes and their Applications*, 116(2):200–221. 11, 12

[Wornell, 1990] Wornell, G. W. (1990). A Karhunen-Loeve-like expansion for 1/f processes via wavelets. *IEEE Transactions on Information Theory*, 36(4):859–861. 7

[Wornell, 1993] Wornell, G. W. (1993). Wavelet-based representations for the 1/f family of fractal processes. *Proceedings of the IEEE*, 81(10):1428–1450. 7