WAVELET ANALYSIS FOR THE SOLUTION TO THE WAVE EQUATION WITH FRACTIONAL NOISE IN TIME AND WHITE NOISE IN SPACE*

OBAYDA ASSAAD AND CIPRIAN A. TUDOR**

Abstract. Via Malliavin calculus, we analyze the limit behavior in distribution of the spatial wavelet variation for the solution to the stochastic linear wave equation with fractional Gaussian noise in time and white noise in space. We propose a wavelet-type estimator for the Hurst parameter of the this solution and we study its asymptotic properties.

1. INTRODUCTION

In mathematical statistics, the parameter estimation for stochastic (partial) differential equations constitutes a topic of wide interest (see, among many others, the monographs or surveys [8, 14] or [20]). In the last decades, the statistical inference for stochastic models driven by fractional Brownian motion and related processes also became a popular topic, due to the developments of the stochastic calculus for fractional processes (see, again among many others, [13, 21, 25]). A common characteristic of the above mentioned references is that they analyze estimators for the drift parameter or for the diffusion coefficient for standard fractional stochastic (partial) differential equations and very few works studied the problem of the estimation of the Hurst parameter of the driving noise (see [12, 22, 23]).

In our work, we will consider the linear stochastic wave equation (2.1) driven by a fractional-white Gaussian noise (*i.e.* a Gaussian noise that behaves as a fractional Brownian motion in time and as a white noise in space) and we construct and analyze statistical estimators for the Hurst index of the solution, based on the discrete observations of the solution in space and time. The stochastic partial differential equation (2.1) constitutes a model for an infinite vibrating string (under an ideal context, with uniform mass, neglecting the air resistance, etc.) perturbed by a random force which behaves as a fractional Brownian motion in time and as a Wiener process in space. For related works on the stochastic wave equation, we refer, among many others, to [4, 10, 24]. The value $u(t, x)$ modelizes the vertical displacement from the x-axis of the string at time t and at position x (in a coordinate system with x on the horizontal line and u on the vertical line). The displacement of the string is clearly affected by the random force and in particular by its Hurst parameter H. This influence of the

Université de Lille, CNRS, Laboratoire Paul Painlevé, UMR 8524, 59655 Villeneuve d'Ascq, France.

** Corresponding author: ciprian.tudor@univ-lille.fr

Hurst parameter appears in several aspects, such as the probability distribution of the solution to (2.1) or the regularity of its sample paths. Indeed, for fixed $x \in \mathbb{R}$, the process u is self-similar of order $H + \frac{1}{2}$ in time and its paths are Hölder continuous of order $\delta \in (0, H)$ in space and the same Hölder continuity holds with respect to the time variable (see *e.g.* [24]). The Hurst parameter also characterizes other properties of the solution, such as the hitting times, the Hausdorff dimension or the regularity of its local times (see *e.g.* [9]). Therefore, the estimation of this parameter is of interest.

We propose a wavelet-type estimator defined *via* the decomposition of the observed process in a wavelet basis. The wavelet estimators have been intensively used in order to identify the Hurst paramter of the fractional Brownian motion and related processes (see *e.g.* [1, 5, 7, 11, 15]). Such estimators have in general several advantages: they are robust and computationally efficient, they are based on the log–log regression of the empirical variance onto several scales and this regression is useful for goodness-of-fit of the model, they offer flexibility on the choice of the wavelet basis etc.

Let $(u(t, x), t \geq 0, x \in \mathbb{R})$ be the solution to the wave equation with fractional-white additive noise. Here we used a wavelet decomposition of the solution to the wave equation (2.1) with respect to its space variable by assuming that the time variable is fixed. That is, we consider a "mother wavelet" Ψ with Q vanishing moments ($Q \geq 1$) and we define the wavelet coefficient $d(t, a, i) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \Psi\left(\frac{x}{a}\right) u(t, x) \mathrm{d}x$ with $t > 0$ fixed and the scale $a > 0$. The wavelet variation, denoted $V_N(t, a)$ in the sequel, is defined by (2.10) by taking the sum of the centered and renormalized squared wavelet coefficients. By analyzing the asymptotic behavior of the wavelet variation $V_N(t, a)$ as $N \to \infty$, we are able to construct, *via* a log–log regression of the empirical variance onto several scales, an estimator for the Hurst parameter of the solution to (2.1) and to analyze its asymptotic behavior. The asymptotic behavior of the estimator is strongly connected to the asymptotic behavior of the wavelet variation $V_N(a)$. The time t also plays a role. For practical purposes, it would be convenient to estimate H by assuming that the solution is observed at a fixed time and at discrete points in space. On the other hand, as we will notice later, in the case of fixed time the empirical variance does not behave as a power function whose exponent is a linear function of H and the log–log regression argument cannot be applied. The relation between the wavelet variance and the Hurst index is more complex and we construct our estimator by analyzing this connection.

The techniques that we use to study the limit behavior in distribution of the wavelet variation are based on the Malliavin calculus and Stein method. We employ the recent Stein-Malliavin theory (see *e.g.* [16]) in order to prove that this sequence satisfies a Central Limit Theorem (CLT) and to derive the rate of convergence for this limit theorem. As mentioned above, we distinguish two situations: when the time t varies with N (*i.e.* $t = N^\beta$ with $\beta > 0$) or when the time t is fixed (and in this case we restrict to the case of the Haar wavelet). We will see that in these two situations, the behavior of the wavelet variation is pretty different, although it always satisfies a CLT (with a different rate of convergence). We deduce the limit behavior of the associated Hurst parameter estimators, *via* a log–log regression of the empirical variance. We also notice that we use *spatial* wavelet variation to estimate the Hurst parameter of the solution, although this parameter appears in the time covariance of the noise and it characterizes the self-similarity of the solution in time.

We organized our paper in the following way: Section 2 contains some preliminaries on the wave equation with fractional-colored noise and on wavelets. In Section 3 we state our main theoretical results. Section 4 contains the proofs of the main results, including the correlation structure of the wavelet coefficients, the magnitude of the L^2 -norm of the wavelet variation and the Central Limit Theorem for this sequence as well as the Berry-Essén bound for this limit theorem. Section 5 is devoted to discretized of the wavelet variation and the construction and the asymptotic study of the wavelet-type estimator for the Hurst parameter of the solution to the stochastic wave equation.

2. Preliminaries

Let us start by presenting some basic facts on the solution to the wave equation with additive fractional-colored noise and on the wavelet analysis.

2.1. The solution to the wave equation

Let $(u(t, x), t \geq 0, x \in \mathbb{R}^n)$ be the solution to the wave equation with fractional-white noise

$$
\begin{cases}
\frac{\partial^2 u}{\partial t^2}(t, x) & = \quad \Delta u(t, x) + \dot{W}^H(t, x), \ \ t \in (0, T], \ T > 0, \ x \in \mathbb{R}^n \\
u(0, x) & = \quad 0, \quad x \in \mathbb{R}^n \\
\frac{\partial u}{\partial t}(0, x) & = \quad 0, \quad x \in \mathbb{R}^n.
\end{cases}
\tag{2.1}
$$

Here Δ is the Laplacian on $\mathbb{R}^n$, $n \geq 1$ and $W^H = \{W_t^H(A); \ t \in [0, T], \ A \in \mathcal{B}_b(\mathbb{R}^n)\}$ is a real valued centered Gaussian field, over a given complete filtered probability space $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \geq 0}, \mathbb{P})$, whose covariance function is

$$
\mathbf{E}(W_t^H(A) W_s^H(B)) = R^H(t, s) \lambda(A \cap B), \text{ for every } t, s \geq 0, A, B \in \mathfrak{B}_b(\mathbb{R}^n),
\tag{2.2}
$$

where λ is the d-dimensional Lebesgue measure, $\mathfrak{B}_b(\mathbb{R}^n)$ is the set of the λ-bounded Borel subsets of $\mathbb{R}^n$ and R^H is the covariance function of the fBm with Hurst parameter $H \in (0, 1)$ given by

$$
R^H(t, s) := \frac{1}{2}\left(t^{2H} + s^{2H} - |t - s|^{2H} \right), \quad s, t \geq 0.
\tag{2.3}
$$

Throughout this work, we will assume $H \in \left(\frac{1}{2}, 1\right)$.

The solution of the equation (2.1) is understood in the mild sense, that is, it is defined as a square-integrable centered field $u = (u(t, x); \ t \in [0, T], x \in \mathbb{R}^n)$ defined by

$$
u(t, x) = \int_0^t \int_{\mathbb{R}^d} G_1(t - s, x - y) W^H(\mathrm{d}s, \mathrm{d}y), \quad t \geq 0, x \in \mathbb{R}^n,
\tag{2.4}
$$

where G_1 is the fundamental solution to the wave equation and the integral in (2.4) is a Wiener integral with respect to the Gaussian process W^H. Recall that for $n = 1$ (we will later restrict to this situation in our work) we have, for every $t \geq 0$ and $x \in \mathbb{R}$,

$$
G_1(t, x) = \frac{1}{2} \mathbb{1}_{\{|x| < t\}}.
\tag{2.5}
$$

We refer to *e.g.* [10] (when $H = \frac{1}{2}$) and to *e.g.* [4] (for $H \in \left(\frac{1}{2}, 1\right)$) for the definition and basic properties of the solution. The solution (2.4) is well-defined in dimension $n = 1$ for every $H \in \left(\frac{1}{2}, 1\right)$ (see *e.g.* [24]) and we have an explicit formula for its spatial covariance which will be a key ingredient in our study (see [12])

$$
\begin{aligned}
\mathbf{E}\left(u(t, x) u(t, y)\right) = {} & \frac{1}{2}\left(c_H |y - x|^{2H+1} - t\frac{|y - x|^{2H}}{2} + \frac{t^{2H+1}}{2H + 1} \right) 1_{\{|y - x| < t\}} \\
& + \frac{(2t - |y - x|)^{2H+1}}{8(2H + 1)} 1_{\{t \leq |y - x| < 2t\}}
\end{aligned}
\tag{2.6}
$$

with $c_H = \frac{4H-1}{4(2H+1)}$. When $t > 1$ and $|x - y| \leq 1$, this expression reduces to

$$
\mathbf{E}\left(u(t, x) u(t, y)\right) = \frac{1}{2}\left(c_H |y - x|^{2H+1} - t\frac{|y - x|^{2H}}{2} + \frac{t^{2H+1}}{2H + 1} \right).
\tag{2.7}
$$

We notice that the solution is stationary in space while it has a scaling property in time (it is actually self-similar in time of order $H + \frac{1}{2}$). The sample paths of the solution are Hölder continuous in time and in time of order $\delta \in (0, H)$ (see *e.g.* [24]).

2.2. Wavelets

Let Ψ be a continuous function with support in $[0, 1]$ such that its first Q moments vanish *i.e.* there exists and integer $Q \geq 1$ such that

$$\int_{\mathbb{R}} t^p \Psi(t) \mathrm{d}t = 0 \text{ for } p = 0, 1, \ldots, Q - 1 \text{ and } \int_{\mathbb{R}} t^Q \Psi(t) \mathrm{d}t \neq 0. \tag{2.8}$$

The function Ψ is usually called *mother wavelet*. Define for $a > 0$, $i = 1, \ldots, N_a$ (with $N_a = [N/a] - 1$)

$$d(t, a, i) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \Psi\left(\frac{x}{a} - i\right) u(t, x) \mathrm{d}x = \sqrt{a} \int_{\mathbb{R}} \Psi(x) u(t, a(x + i)) \mathrm{d}x \tag{2.9}$$

and

$$\tilde{d}(t, a, i) = \frac{d(t, a, i)}{\left(\mathbf{E}(d(t, a, i))^2\right)^{\frac{1}{2}}}.$$

Also define *the wavelet variation* in space of the solution (2.4) by

$$V_N(t, a) = \frac{1}{N_a} \sum_{i=1}^{N_a} \left(\tilde{d}(t, a, i)^2 - 1\right). \tag{2.10}$$

We will study the asymptotic behavior, as $N_a \to \infty$, of the wavelet variation $V_N(t, a)$. In applications, the parameter a, which is called *scale*, will depend on N and it is usually assumed that $a = a_N \to_{N \to \infty} \infty$.

Given the covariance of the solution to the wave equation (see formula (2.6)), it is clear that the time t will play an important role, depending on its position with respect to the spatial increment $|x - y|$.

We will consider two situations: the *fixed time case, i.e.* the time $t > 0$ is fixed, and the *moving time case,* when the time depends on N and it tends to infinity as $N \to \infty$. The first situation would be more convenient for applications to parameter estimation, since it means that the solution is observed only at a fixed time. Nevertheless, in this case the wavelet variation does not provide an explicit estimator since the usual log–log regression procedure to construct an wavelet estimator based on $V_N(t, a)$ leads to a more complicated equation in H. A slightly different argument is then used for fixed time.

We will start with the moving time situation. We will assume

$$a = a_N = N^\alpha \text{ with } 0 < \alpha < 1 \text{ and } t = t_N = N^\beta \text{ with } \beta \geq 1. \tag{2.11}$$

The choice of such time t will be explained later, it allows to reduce the expression of the correlation of the wavelet coefficients. Then, we will consider the situation when the time is fixed, *i.e.* we suppose

$$a = a_N = N^\alpha \text{ with } 0 < \alpha < 1 \text{ and } t > 0 \text{ is fixed.} \tag{2.12}$$

In this second case, in order to have a precise estimate on the wavelet coefficient and on the empirical variance $\mathbf{E}V_N(t, a)$, we need to restrict to a particular case of wavelet system (the Haar wavelet).

3. Main results

In this section we will state our main theoretical results. Their proofs are postponed to Section 4. These results give the asymptotic behavior as $N \to \infty$ of the wavelet variation $V_N(t, a)$ given by (2.10) as well as the limit behavior in distribution of the renormalized wavelet variation. We will show that, in both moving time and fixed time cases, the magnitude of the variance of $V_N(t, a)$ as $N \to \infty$ is the same and the renormalized wavelet variation satisfies a Central Limit Theorem. We also evaluate the rate of convergence to the normal distribution, which varies in the two cases under consideration.

3.1. The moving time case

Let us start by treating the situation when the time t depends on N, *i.e.* we assume (2.11). In this case, we obtain the following renormalization of the wavelet variation.

Proposition 3.1. *Let $V_N(t, a)$ be given by (2.10). Assume $Q \geq 2$ or $Q = 1, H < \frac{3}{4}$. Let a_N, t_N be given by (2.11). Then*

$$N^{1-\alpha}\mathbf{E}V_N(t_N, a_N)^2 \to_{N\to\infty} \frac{2}{K_{\Psi,H}^2} \sum_{k\in\mathbb{Z}} g_H(k)^2 := K_{0,\Psi,H} \tag{3.1}$$

with g_H given by

$$g_H(k) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y + k|^{2H} \tag{3.2}$$

and $K_{\Psi,H}$ given by, for $H \in \left(\frac{1}{2}, \frac{3}{2}\right)$

$$K_{\Psi,H} = -\int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y|^{2H}. \tag{3.3}$$

Notice that the above integral (3.3) is finite because the support of the mother wavelet Ψ is included in the interval $[0, 1]$ and $2H > 0$. We assume, as in [5], that $K_{\Psi,H} > 0$ (which is satisfied by a large choice of the mother wavelet Ψ). The results in Section 4 show also that the series in the right-hand side of (3.1) is convergent.

Let us denote, for every $N \geq 1$

$$F_N = K_{0,\Psi,H}^{-\frac{1}{2}} N^{\frac{1-\alpha}{2}} V_N(t_N, a_N) \tag{3.4}$$

with $V_N(t_N, a_N)$ defined in (2.10), $K_{0,\Psi,H}$ from (3.1) and suppose that assumption (2.11) is verified. From Proposition 3.1

$$\mathbf{E}F_N^2 \to_{N\to\infty} 1.$$

We will obtain the following result. We denote below by c, C generic strictly positive constants that may change from line to line. By d we denote the distance between distributions of random variable and below it can be each of the following distances: Kolmogorov, total variation, Wasserstein or Fortet-Mourier (see [16]).

Theorem 3.2. *Let F_N be given by (3.4). Then the sequence $(F_N)_{N\geq 1}$ converges in distribution to a standard normal random variable $Z \sim N(0, 1)$ and*

$$d(F_N, Z) \leq cN^{\frac{\alpha-1}{2}}.$$

We can also prove a multidimensional central limit theorem for the wavelet variation considered at different scales. This will be used in order to estimate the Hust parameter of the solution to the wave equation in the next section.

Theorem 3.3. *Let $V_N(t, a)$ be given by (2.10) and assume (2.11). Let $d \geq 1$. Then the d-dimensional random vector $\left(N^{\frac{1-\alpha}{2}} V_n(t_N, La_N) \right)_{L=1,\ldots,d}$ converges in distribution, as $N \to \infty$, to a centered d-dimensional Gaussian vector with covariance matrix $(\Gamma_{L_1,L_2})_{L_1,L_2=1,\ldots,d}$ where*

$$\Gamma_{L_1,L_2} = \frac{32}{K_{\Psi,H}^2} \frac{1}{(L_1 L_2)^{2H+1}} C(L_1, L_2, H) \tag{3.5}$$

with $C(L_1, L_2, H)$ given by

$$C(L_1, L_2, H) = \lim_{N \to \infty} N^{1-\alpha} \sum_{i=1}^{N_{L_1 a_N}} \sum_{j=1}^{N_{L_2 a_N}} \left(g_{L_1,L_2,H}(i,j) \right)^2 \tag{3.6}$$

where

$$g_{L_1,L_2,H}(i,j) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x \mathrm{d}y \, \Psi(x) \Psi(y) |L_1 x - L_2 y + L_1 - L_2 j|^{2H}.$$

It follows from our proofs in Section 4 that the limit in the right-hand side of (3.6) exists and is finite.

3.2. The fixed time case

If t is fixed, we can prove the following approximation result for the variance of the wavelet variation. As mentioned, the role of the mother wavelet will be played by the Haar wavelet, *i.e.*

$$\Psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}$$

Proposition 3.4. *If $V_N(t, a)$ is given by (2.10) and (2.12), (3.7) hold true, we have for every $t > 0$*

$$N^{1-\alpha} \mathbf{E} V_N(t, a_N)^2 \to_{N \to \infty} 2. \tag{3.8}$$

By Proposition 3.4, we have the following renormalization of the wavelet variation

$$G_N =: \frac{1}{\sqrt{2}} N^{\frac{1-\alpha}{2}} V(t, a_N), \tag{3.9}$$

i.e. $\mathbf{E} G_N^2 \to_{N \to \infty} 1$. We will show below that the renormalized wavelet variation satisfies a CLT also when the time is fixed.

Theorem 3.5. *The sequence $(G_N)_{N \geq 1}$ given by (3.9) converges in distribution to $Z \sim N(0, 1)$ and for N large enough*

$$d(G_N, Z) \leq C \left(\frac{1}{N^{\frac{1-\alpha}{2}}} + \frac{1}{N^{2\alpha}} \right).$$

Let us make a short discussion around the above statements.

Remark 3.6. • We notice that the renormalization of (2.10) is of the same order in both cases (fixed time or moving time) although, as we will see in Section 4, the correlation structure of the wavelet coefficient is different.

• The wavelet variation (2.10) satisfies a CLT both in the moving or fixed time cases. On the other hand, the behavior of this sequence is pretty different in these two cases. While for fixed time, this sequence basically behaves as a sum of independent random variables (see also Remark 3.6), in the moving time case there is a non-trivial correlation between all the summands that compose $V_N(t,a)$.

• The rate of convergence of the sequence (3.9) to the normal distribution varies upon $\alpha \in (0,1)$: when $\alpha \in \left(0, \frac{1}{5}\right)$, we have $d(G_N, Z) \leq c\frac{1}{N^{2\alpha}}$ while for $\alpha \in \left(\frac{1}{5}, 1\right)$, one has $d(G_N, Z) \leq c\frac{1}{N^{\frac{1-\alpha}{2}}}$. Theorem 3.5 also suggests that if the scale a is constant (*i.e.* $\alpha = 0$) the sequence $V_N(t,a)$ does not satisfy a CLT.

4. Proofs

This part contains the proofs of the theoretical results stated in Section 3.

4.1. The correlation structure of the wavelet coefficient

The behavior of the wavelet variation (2.10) will depend on the behavior of the variance of the wavelet coefficient $\mathbf{E}d(t,a,i)^2$ and of the correlation between the wavelet coefficients, *i.e.* $\mathbf{E}d(t,a,i)d(t,a,j)$ with $i \neq j$. We will start by analyzing the behavior of these quantities in both cases (2.11) and (2.12).

Let $d(t,a,i)$ be given by (2.9) with $t > 0, a > 0$ and $i = 1, \ldots, N_a$. We will use the following notation throughout our work

$$D(t,a) := \mathbf{E}d(t,a,i)^2 \tag{4.1}$$

for every $t > 0, a > 0$ and $i = 1, \ldots, N_a$. Notice that, due to the stationarity of the process $(u(t,x), x \in \mathbb{R})$, the quantity $\mathbf{E}d(t,a,i)^2$ does not depend on i.

Let $t > 0, a > 0$. For every $i, j = 1, \ldots, N_a$ we have from the covariance formula (2.6)

$$
\begin{aligned}
\mathbf{E}d(t,a,i)d(t,a,j) &= a \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y\Psi(x)\Psi(y)\mathbf{E}u(t,a(x+i))u(t,a(x+j)) \\
&= a \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y\Psi(x)\Psi(y)\left[\frac{c_H}{2}a^{2H+1}|x-y+i-j|^{2H+1}\right. \\
&\quad \left. -\frac{t}{4}a^{2H}|x-y+i-j|^{2H} + \frac{t^{2H+1}}{2(2H+1)}\right]1_{\{|x-y+i-j|<\frac{t}{a}\}} \\
&\quad + a \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y\Psi(x)\Psi(y)\frac{(2t-a|x-y+i-j|)^{2H+1}}{8(2H+1)}1_{\{\frac{t}{a}\leq|x-y+i-j|<2\frac{t}{a}\}}.
\end{aligned}
\tag{4.2}
$$

We will see below that the above expression will simplify under assumption (2.11).

4.2. The moving time case

First, we assume that we work under the assumption (2.11). We start by studying the variance of the wavelet coefficient. Let us recall the notation $K_{\Psi,H}$ from (3.3).

We have the following result.

Lemma 4.1. *Assume* (2.11). *Consider the wavelet coefficient* $d(t, a, i)$ *defined by* (2.9) *and its variance* $D(t, a)$ *given by* (4.1). *Then*

$$\frac{1}{N^{\beta + (2H+1)\alpha}} D(t_N, a_N) \to_{N\to\infty} \frac{1}{4} K_{\Psi, H}$$

with $K_{\Psi, H}$ *from* (3.3).

Proof. From the assumption (2.11) and the property (2.8) of the function Ψ, using also $|x - y| \leq 1$ (which implies that $|x - y + i - j| \leq t_N = N^\beta$ by (2.11)), the last two summands in (4.2) vanish and we obtain

$$\begin{aligned} & \mathbf{E}d(t_N, a_N, i)d(t_N, a_N, j) \\ & = a_N \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) \left[\frac{c_H}{2} a_N^{2H+1} |x - y + i - j|^{2H+1} - a_N^{2H} \frac{t_N}{4} |x - y + i - j|^{2H} \right]. \end{aligned} \tag{4.3}$$

Let us take $i = j$ in (4.3). We have

$$D(t_N, a_N) = \mathbf{E}d(t_N, a_N, i)^2 = -\frac{c_H}{2} K_{\Psi, H+\frac{1}{2}} N^{(2H+2)\alpha} + \frac{1}{4} K_{\Psi, H} N^{\beta + (2H+1)\alpha}. \tag{4.4}$$

Since $\beta + (2H + 1)\alpha > (2H + 2)\alpha$ (because $\beta > \alpha$) we obtain the conclusion. $\qquad\square$

Let us now study the correlation (4.3) with $i \neq j$. We can write

$$\mathbf{E}d(t_N, a_N, i)d(t_N, a_N, j) = \frac{c_H}{2} a_N^{2H+2} g_{H+\frac{1}{2}}(i - j) - \frac{t_N}{4} a_N^{2H+1} g_H(i - j) \tag{4.5}$$

with the notation $g_H(k)$ from (3.2). Notice that for every $k \in \mathbb{Z}$ we have $g_H(k) = g_H(-k)$ for $k \in \mathbb{Z}$. The analysis of the quantity $g_H(k)$ for k large, will give the asymptotics of the correlation (4.5). Recall that the integer $Q \geq 1$ is fixed by (2.8).

Lemma 4.2. *Let* g_H *be given by* (3.2). *Then for k large enough, we have for every* $H \in \left(\frac{1}{2}, \frac{3}{2} \right)$

$$|g_H(k)| \leq C_{\Psi, H, Q} k^{4H - 4Q}$$

where $C_{\Psi, H, Q}$ *is a strictly positive constant not depending on k.*

Proof. Using the following asymptotic expansion at $z = 0$

$$(1 + z)^{2H} = 1 + 2Hz + \dots + \frac{2H(2H - 1)\dots(2H - 2Q)}{(2Q - 1)!} z^{2Q - 1} + C_{H, Q}(1 + \theta_z)^{2H - 2Q} z^{2Q}$$

where θ_z is a point located between 0 and z, we can write, for k large enough, if $C_{H, Q}$ is a constant depending only of H and Q,

$$\begin{aligned} g_H(k) &= k^{2H} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) \left(1 + \frac{x - y}{k} \right)^{2H} \\ &= C_{H, Q} k^{2H} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) \left(\frac{x - y}{k} \right)^{2Q} (1 + \theta_{x, y, k})^{2H - 2Q} \end{aligned}$$

where we used (2.8) and we denoted by $\theta_{x,y,k}$ a point located between 0 and $\frac{x-y}{k}$. Since $|x-y| \leq 1$, we have for $k \geq 2$

$$\frac{1}{2} \leq |1 + \theta_{x,y,k}| \leq \frac{3}{2}.$$

We deduce that, for k large

$$|g_H(k)| \leq C_{H,Q} 2^{2Q-2H} k^{2H-2Q} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x \mathrm{d}y |\Psi(x)\Psi(y)| |x-y|^{2Q} = C_{\Psi,H,Q} k^{2H-2Q}$$

using the fact that the support of Ψ is included in the interval $[0,1]$.

$\square$

Lemma 4.3. *Let g_H be given by (3.2). Denote, for $a > 0$ and $N \geq 1$*

$$g_{N,H}(a) = \sum_{i,j=1}^{N_a} g_H(i-j)^2. \tag{4.6}$$

Then, for every $H \in \left(\frac{1}{2}, \frac{3}{4}\right)$ (if $Q = 1$) and for every $H \in \left(\frac{1}{2}, 1\right)$ (if $Q \geq 2$)

$$\frac{1}{N_a} g_{N,H}(a) \to_{N_a \to \infty} \sum_{k \in \mathbb{Z}} g_H(k)^2. \tag{4.7}$$

Moreover, for every $H \in \left(\frac{1}{2}, 1\right)$ and for every $Q \geq 1$, for N large enough

$$\frac{1}{N_a} |g_{N,H+\frac{1}{2}}(a)| \leq \begin{cases} C_{\Psi,H,Q} & \text{if } Q \geq 2, \\ C_{\Psi,H,Q} N_a^{4H-1} & \text{if } Q = 1 \end{cases} \tag{4.8}$$

and

$$\frac{1}{N_a} \left| \sum_{i,j=1}^{N_a} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \right| \leq \begin{cases} C_{\Psi,H,Q} & \text{if } Q \geq 2, \\ C_{\Psi,H,Q} N_a^{4H-2} & \text{if } Q = 1. \end{cases} \tag{4.9}$$

Proof. We can write

$$\frac{1}{N_a} g_{N,H}(a) = \sum_{k \in \mathbb{Z}} g_H(k)^2 1_{\{|k| \leq N_a\}} \frac{N_a - |k|}{N_a}.$$

By the dominated convergence theorem and Lemma 4.2 we clearly have

$$\frac{1}{N_a} g_{N,H}(a) \to_{N_a \to \infty} \sum_{k \in \mathbb{Z}} g_H(k)^2.$$

Note that the series $\sum_{k \in \mathbb{Z}} g_H(k)^2$ is convergent due to Lemma 4.2. Now

$$\frac{1}{N_a} |g_{N,H+\frac{1}{2}}(a)| = \left| \sum_{k \in \mathbb{Z}} g_{H+\frac{1}{2}}(k)^2 1_{\{|k| \leq N_a\}} \frac{N_a - |k|}{N_a} \right|$$

$$\leq \sum_{|k| \leq N_a} g_{H+\frac{1}{2}}(k)^2 \leq C \sum_{|k| \leq N_a} k^{4H+2-4Q}$$

again by Lemma 4.2. The series $\sum_{k \in \mathbb{Z}} k^{4H+2-4Q}$ is convergent when $Q \geq 2$ and for $Q = 1$ and $H > \frac{1}{2}$, the sequence $\sum_{|k| \leq N_a} k^{4H+2-4Q}$ behaves as $C_{H,Q} N_a^{4H-1}$. This implies the estimate (4.8). A similar argument gives (4.9), because from Lemma 4.2

$$\frac{1}{N_a} \left| \sum_{i,j=1}^{N_a} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \right| = \left| \sum_{k \in \mathbb{Z}} g_H(k) g_{H+\frac{1}{2}}(k) 1_{\{|k| \leq N_a\}} \frac{N_a - |k|}{N_a} \right|$$
$$\leq C_{\Psi,H,Q} \sum_{|k| \leq N_a} k^{4H+1-4Q}.$$

$\square$

4.2.1. The fixed time case

Let us assume $t > 0$ is fixed, *i.e.* we assume (2.12). As before, we use the notation

$$D(t, a_N) = \mathbf{E} d(t, a_N, i)^2$$

for $i = 1, \ldots, N_{a_N}$, with $a_N = N^\alpha$, $0 < \alpha < 1$. We start by estimating the behavior of $D(t, a_N)$ as $N \to \infty$. It is impossible to get the exact behavior of this quantity for an arbitrary function Ψ. Therefore, in the sequel we will choose the function Ψ to be the mother wavelet of the Haar system, see (3.7).

Proposition 4.4. *Let Ψ be given by (3.7) and assume (2.12). For every $t > 0$ and for N large enough*

$$D(t, a_N) = K_{1,t}(H) + K_{2,t}(H) \frac{1}{N^\alpha}$$

with

$$K_{1,t}(H) = \frac{1}{2(H+1)} t^{2H+2}. \tag{4.10}$$

and $K_{2,t}(H) = \sum_{j=1}^{4} K_{j,2,t}(H)$ for $K_{j,2,t}(H)$, $j = 1, \ldots, 4$ given by (4.17), (4.18), (4.19) and (4.20). In particular,

$$D(t, a_N) \to_{N \to \infty} K_{1,t}(H).$$

Proof. From (4.2) we have

$$D(t, a_N) = I_{1,t,N} + I_{2,t,N} + I_{3,t,N} + I_{4,t,N}$$

with

$$I_{1,t,N} = \frac{c_H}{2} N^{\alpha(2H+2)} A_{H+\frac{1}{2},N}, \quad I_{2,t,N} = -\frac{t}{4} N^{\alpha(2H+1)} A_{H,N}, \tag{4.11}$$

$$I_{3,t,N} = \frac{t^{2H+1}}{2(2H+1)} N^\alpha \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x \mathrm{d}y \Psi(x) \Psi(y) 1_{\{|x-y| < \frac{t}{N^\alpha}\}}, \tag{4.12}$$

$$I_{4,t,N} = N^\alpha \frac{1}{8(2H+1)} B_{H,N} \tag{4.13}$$

where we used the notation

$$A_{H,N} := \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x-y|^{2H} 1_{\{|x-y|<\frac{t}{N^\alpha}\}} \tag{4.14}$$

and

$$B_{H,N} = \int_0^1 \int_0^1 \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) \left(2t - N^\alpha|x-y|\right)^{2H+1} 1_{\{\frac{t}{N^\alpha}\leq|x-y|<2\frac{t}{N^\alpha}\}}. \tag{4.15}$$

To obtain the speed of convergence of $I_{1,t,N}$ and $I_{2,t,N}$, we need to study the sequence $A_{H,N}$ defined by (4.14). Clearly, $A_{H,N}$ converges to zero as $N \to \infty$ but we need to analyze how fast this sequence goes to zero. We have

$$\begin{aligned}
A_{H,N} &= 2 \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}y \Psi(x)\Psi(y)(x-y)^{2H} 1_{\{x-y<\frac{t}{N^\alpha}\}} \\
&= 2 \int_0^1 \mathrm{d}x \int_{(x-tN^{-\alpha})\vee 0}^x \mathrm{d}y \Psi(x)\Psi(y)(x-y)^{2H} \\
&= 2 \int_0^{tN^{-\alpha}} \mathrm{d}x \int_0^x \mathrm{d}y \Psi(x)\Psi(y)(x-y)^{2H} + 2 \int_{tN^{-\alpha}}^1 \mathrm{d}x \int_{x-tN^{-\alpha}}^x \mathrm{d}y \Psi(x)\Psi(y)(x-y)^{2H}.
\end{aligned}$$

Let us chose N large enough such that

$$\frac{t}{N^\alpha} < \frac{1}{2}.$$

We will have, with Ψ from (3.7),

$$\begin{aligned}
A_{H,N} &= 2 \int_0^{tN^{-\alpha}} \mathrm{d}x \int_0^x \mathrm{d}y(x-y)^{2H} + 2 \int_{tN^{-\alpha}}^{\frac{1}{2}} \mathrm{d}x \int_{x-tN^{-\alpha}}^x \mathrm{d}y(x-y)^{2H} \\
&\quad - 2 \int_{\frac{1}{2}}^1 \mathrm{d}x \int_{x-tN^{-\alpha}}^x \mathrm{d}y \Psi(y)(x-y)^{2H}
\end{aligned}$$

and by separating the integral $\mathrm{d}y$ in the last term above upon $x = tN^{-\alpha}$ less or bigger than one-half we will obtain

$$\begin{aligned}
A_{H,N} &= 2 \int_0^{tN^{-\alpha}} \mathrm{d}x \int_0^x \mathrm{d}y(x-y)^{2H} + 2 \int_{tN^{-\alpha}}^{\frac{1}{2}} \mathrm{d}x \int_{x-tN^{-\alpha}}^x \mathrm{d}y(x-y)^{2H} \\
&\quad - 2 \int_{\frac{1}{2}}^{\frac{1}{2}+tN^{-\alpha}} \mathrm{d}x \int_{x-tN^{-\alpha}}^{\frac{1}{2}} \mathrm{d}y(x-y)^{2H} + 2 \int_{\frac{1}{2}}^{\frac{1}{2}+tN^{-\alpha}} \mathrm{d}x \int_{\frac{1}{2}}^x \mathrm{d}y(x-y)^{2H} \\
&\quad + 2 \int_{\frac{1}{2}+tN^{-\alpha}}^1 \mathrm{d}x \int_{x-tN^{-\alpha}}^x \mathrm{d}y(x-y)^{2H}.
\end{aligned}$$

This gives

$$
\begin{aligned}
A_{H,N} &= \frac{2}{2H+1}\left[\frac{1}{2H+2}\left(\frac{t}{N^\alpha}\right)^{2H+2} + \left(\frac{t}{N^\alpha}\right)^{2H+1}\left(\frac{1}{2}-\frac{t}{N^\alpha}\right) - \left(\frac{t}{N^\alpha}\right)^{2H+1}\frac{t}{N^\alpha}\right. \\
&\quad \left. +\frac{1}{2H+2}\left(\frac{t}{N^\alpha}\right)^{2H+2} + \frac{1}{2H+2}\left(\frac{t}{N^\alpha}\right)^{2H+2} + \left(\frac{t}{N^\alpha}\right)^{2H+1}\left(\frac{1}{2}-\frac{t}{N^\alpha}\right)\right] \\
&= \frac{2}{2H+1}\left[3\frac{1}{2H+2}\left(\frac{t}{N^\alpha}\right)^{2H+2} + \left(\frac{t}{N^\alpha}\right)^{2H+1}\left(1-3\frac{t}{N^\alpha}\right)\right] \\
&= -\frac{6}{2H+2}\left(\frac{t}{N^\alpha}\right)^{2H+2} + \frac{2}{2H+1}\left(\frac{t}{N^\alpha}\right)^{2H+1}.
\end{aligned}
\tag{4.16}
$$

Consequently, we obtain from (4.16) the following behavior for the summand $I_{1,t,N}$ in (4.11)

$$
I_{1,t,N} = \frac{c_H}{2}N^{\alpha(2H+2)}A_{H+\frac{1}{2},N} = K_{1,1,t}(H) + K_{1,2,t}(H)\frac{1}{N^\alpha}
\tag{4.17}
$$

with

$$
K_{1,1,t}(H) = \frac{c_H}{2H+2}t^{2H+2} \text{ and } K_{1,2,t}(H) = \frac{-3c_H}{2H+3}t^{2H+3}.
$$

The second summand $I_{2,t,N}$ gives, using (4.16)

$$
I_{2,t,N} = -\frac{t}{4}N^{\alpha(2H+1)}A_{H,N} = K_{2,1,t}(H) + K_{2,2,t}(H)\frac{1}{N^\alpha}
\tag{4.18}
$$

with

$$
K_{2,1,t}(H) = -\frac{1}{2(2H+1)}t^{2H+2} \text{ and } K_{2,2,t}(H) = \frac{3}{2(2H+2)}t^{2H+3}.
$$

Let us now calculate the term $I_{3,t,N}$ defined in (4.13). We can write

$$
I_{3,t,N} = N^\alpha\frac{t^{2H+1}}{2(2H+1)}2\int_0^1 dx\int_0^x dy\Psi(x)\Psi(y)1_{\{x-y<\frac{t}{N^\alpha}\}}
$$

and since (this is the same calculation as for $A_{H,N}$ without the factor $(x-y)^{2H}$)

$$
2\int_0^1 dx\int_0^x dy\Psi(x)\Psi(y)1_{\{x-y<\frac{t}{N^\alpha}\}} = \frac{2t}{N^\alpha} - 3\left(\frac{t}{N^\alpha}\right)^2
$$

we obtain

$$
I_{3,t,N} = K_{3,1,t}(H) + K_{3,2,t}(H)\frac{1}{N^\alpha}
\tag{4.19}
$$

with

$$K_{3,1,t}(H) = \frac{1}{(2H+1)}t^{2H+2} \text{ and } K_{3,2,t}(H) = \frac{-3}{2(2H+1)}t^{2H+3}.$$

Let us regard the last summand $I_{4,t,N}$ in (4.13). With $B_{H,N}$ given by (4.15)

$$
\begin{aligned}
B_{H,N} &= \int_0^1 \int_0^1 dxdy\Psi(x)\Psi(y)\left(2t - N^\alpha|x-y|\right)^{2H+1} 1_{\{\frac{t}{N^\alpha} \leq |x-y| < 2\frac{t}{N^\alpha}\}} \\
&= 2\int_0^1 \int_0^x dxdy\Psi(x)\Psi(y)\left(2t - N^\alpha(x-y)\right)^{2H+1} 1_{\{\frac{t}{N^\alpha} \leq x-y < 2\frac{t}{N^\alpha}\}} \\
&= 2\int_0^1 dx \int_{(x-2tN^{-\alpha})\vee 0}^{x-tN^{-\alpha}} dy\Psi(x)\Psi(y)\left(2t - N^\alpha(x-y)\right)^{2H+1} \\
&= 2\int_0^{2tN^{-\alpha}} dx \int_0^{x-tN^{-\alpha}} dy\Psi(x)\Psi(y)\left(2t - N^\alpha(x-y)\right)^{2H+1} \\
&\quad + 2\int_{2tN^{-\alpha}}^1 dx \int_{x-2tN^{-\alpha}}^{x-tN^{-\alpha}} dy\Psi(x)\Psi(y)\left(2t - N^\alpha(x-y)\right)^{2H+1} := B_{1,H,N} + B_{2,H,N}.
\end{aligned}
$$

We estimate separately the summands $B_{1,H,N}$ and $B_{2,H,N}$. First, notice that we can choose N large enough so that $\frac{t}{N^\alpha} < \frac{1}{4}$ and therefore $\frac{2t}{N^\alpha} < \frac{1}{2}$. We then get

$$B_{1,H,N} = 2\frac{t^{2H+3}}{(2H+2)N^{2\alpha}}\left(2 - \frac{2^{2H+3}}{2H+3}\right)$$

while for $B_{2,H,N}$ we have

$$
\begin{aligned}
B_{2,H,N} &= 2\int_{2tN^{-\alpha}}^{1/2} dx \int_{x-2tN^{-\alpha}}^{x-tN^{-\alpha}} dy\left(2t - N^\alpha(x-y)\right)^{2H+1} \\
&\quad - 2\int_{1/2}^{1/2+tN^{-\alpha}} dx \int_{x-2tN^{-\alpha}}^{x-tN^{-\alpha}} dy\left(2t - N^\alpha(x-y)\right)^{2H+1} \\
&\quad + 2\int_{1/2+tN^{-\alpha}}^1 dx \int_{x-2tN^{-\alpha}}^{x-tN^{-\alpha}} dy\left(2t - N^\alpha(x-y)\right)^{2H+1} \\
&= 2\frac{t^{2H+2}}{N^\alpha(2H+2)}\left(1 - \frac{4t}{N^\alpha}\right).
\end{aligned}
$$

By putting together the above computations, we obtain

$$I_{4,t,N} := \frac{1}{8(2H+1)}N^\alpha B_{H,N} = K_{4,1,t}(H) + K_{4,2,t}(H)\frac{1}{N^\alpha} \tag{4.20}$$

with

$$K_{4,1,t}(H) = \frac{t^{2H+2}}{8(H+1)(2H+1)} \text{ and } K_{4,2,t}(H) = -\frac{t^{2H+3}}{8(H+1)(2H+1)}\left(2 + \frac{2^{2H+3}}{2H+3}\right).$$

From (4.17), (4.18), (4.19) and (4.20) we obtain the conclusion. In particular, concerning the constant $K_{1,t}(H)$ which is needed in the sequel

$$K_{1,t}(H) = t^{2H+2}\left(\frac{c_H}{2H+2} - \frac{1}{2(2H+1)} + \frac{1}{2H+1} + \frac{1}{8(2H+1)(H+1)}\right) = \frac{1}{2H+2}t^{2H+2}$$

by using the expression of c_H in (2.6). $\qquad\square$

We also need to analyze $\mathbf{E}d(t, a_N, i)d(t, a_N, j)$ when $|i-j| = 1$. Only this correlation coefficient will be needed for the renormalization of the sequence (2.10).

Proposition 4.5. *Let $d(t, a, i)$ be given by (2.9) and assume (2.12) and (3.7). Then for every $t > 0, N \geq 1$*

$$\mathbf{E}d(t, a_N, i)d(t, a_N, i+1) = L_t(H)\frac{1}{N^\alpha}$$

with $L_t(H)$ from (4.28).

Proof. We have

$$\mathbf{E}d(t, a_N, i)d(t, a_N, j) = f_{H,N}(i-j)$$

where (recall $a_N = N^\alpha$)

$$
\begin{aligned}
f_{H,N}&(k)\\
&= a_N \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)\left[\frac{c_H}{2}a_N^{2H+1}|x-y+k|^{2H+1}\right.\\
&\quad \left. -\frac{t}{4}a_N^{2H}|x-y+k|^{2H} + \frac{t^{2H+1}}{2(2H+1)}\right]1_{\{|x-y+k|<\frac{t}{a_N}\}}\\
&\quad +a_N \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)\frac{(2t - a_N|x-y+k|)^{2H+1}}{8(2H+1)}1_{\{\frac{t}{a_N}\leq|x-y+k|<2\frac{t}{a_N}\}}.
\end{aligned}
\tag{4.21}
$$

Hence

$$\mathbf{E}d(t, a_N, i)d(t, a_N, i+1) = f_{H,N}(1).$$

We can write, *via* (4.2)

$$f_{H,N}(1) = J_{1,t,N} + J_{2,t,N} + J_{3,t,N} + J_{4,t,N}$$

with

$$J_{1,t,N} = \frac{c_H}{2}N^{\alpha(2H+2)}C_{H+\frac{1}{2},N}, \quad J_{2,t,N} = \frac{-t}{4}N^{\alpha(2H+1)}C_{H,N} \tag{4.22}$$

$$J_{3,t,N} = N^\alpha \frac{t^{2H+1}}{2(2H+1)}\int_0^1\int_0^1 \mathrm{d}x\mathrm{d}y\Psi(x)\Psi(y)1_{\{x-y+1<\frac{t}{N^\alpha}\}} \tag{4.23}$$

$$J_{4,t,N} = N^\alpha \int_0^1\int_0^1 \mathrm{d}x\mathrm{d}y\Psi(x)\Psi(y)\frac{(2t - N^\alpha|x-y+1|)^{2H+1}}{8(2H+1)}1_{\{\frac{t}{a_N}\leq|x-y+1|<2\frac{t}{a_N}\}} \tag{4.24}$$

where

$$C_{H,N} = \int_0^1 \int_0^1 \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)(x-y+1)^{2H} 1_{\{x-y+1 < \frac{t}{N^\alpha}\}}.$$

We have, if N is such that $\frac{t}{N^\alpha} < \frac{1}{2}$,

$$\begin{aligned}
C_{H,N} &= \int_0^1 \mathrm{d}x \int_{x+1-tN^{-\alpha}}^1 \mathrm{d}y \Psi(x)\Psi(y)(x-y+1)^{2H} \\
&= -\int_0^{\frac{t}{N^\alpha}} \int_{x+1-tN^{-\alpha}}^1 \mathrm{d}y (x-y+1)^{2H} \\
&= \left(\frac{1}{(2H+1)(2H+2)} - \frac{1}{2H+1}\right) \left(\frac{t}{N^\alpha}\right)^{2H+2}.
\end{aligned}$$

Therefore

$$J_{1,t,N} = K_{5,1,t}(H)\frac{1}{N^\alpha} \tag{4.25}$$

with

$$K_{5,1,t}(H) = \frac{c_H t^{2H+3}}{2}\left(\frac{1}{(2H+2)(2H+3)} - \frac{1}{2H+2}\right).$$

For the second term $J_{2,t,N}$ in (4.22), it is immediate to see that

$$J_{2,t,N} = \frac{-t}{4} N^{\alpha(2H+1)} C_{H,N} = K_{6,1,t}(H)\frac{1}{N^\alpha} \tag{4.26}$$

with

$$K_{6,1,t}(H) = \frac{-t^{2H+3}}{4}\left(\frac{1}{(2H+1)(2H+2)} - \frac{1}{2H+1}\right).$$

The third summand (4.23) gives

$$\begin{aligned}
J_{3,t,N} &= N^\alpha \frac{t^{2H+1}}{2(2H+1)} \int_0^1 \int_0^1 \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) 1_{\{x-y+1 < \frac{t}{N^\alpha}\}} \\
&= -N^\alpha \frac{t^{2H+1}}{2(2H+1)} \int_0^{\frac{t}{N^\alpha}} \mathrm{d}x \int_{x+1-tN^{-\alpha}}^1 \mathrm{d}y = \frac{t^{2H+3}}{4(2H+1)}\frac{1}{N^\alpha} =: K_{7,1,t}(H)\frac{1}{N^\alpha}.
\end{aligned}$$

Finally, concerning the summand $J_{4,t,N}$ in (4.24), if $2t/N^\alpha < \frac{1}{2}$,

$$\begin{aligned}
J_{4,t,N} &= N^\alpha \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y) \frac{(2t - N^\alpha|x-y+1|)^{2H+1}}{8(2H+1)} 1_{\{\frac{t}{a_N} \le |x-y+1| < 2\frac{t}{a_N}\}} \\
&= -\frac{1}{8(2H+1)} N^\alpha \int_0^{tN^{-\alpha}} \mathrm{d}x \int_{x+1-2tN^{-\alpha}}^{x+1-tN^{-\alpha}} \mathrm{d}y (2t - N^\alpha(x-y+1))^{2H+1}.
\end{aligned}$$

We obtain

$$J_{4,t,N} = K_{8,1,t}(H)\frac{1}{N^\alpha} \text{ with } K_{8,1,t}(H) = \frac{-t^{2H+3}}{8(2H+2)(2H+1)}. \tag{4.27}$$

Consequently,

$$f_{H,N}(1) = L_t(H)\frac{1}{N^\alpha} \text{ with } L_t(H) = K_{5,1,t}(H) + K_{6,1,t}(H) + K_{7,1,t}(H) + K_{8,1,t}(H). \tag{4.28}$$

$\square$

4.3. Renormalization of the wavelet variation

In order to analyze the asymptotic behavior of the wavelet variation (2.10), we will use the chaotic expression of $V_N(t,a)$. We will work with multiple stochastic integrals with respect to the fractional-white noise W^H.

Let $\mathcal{E}$ denote the space of all linear combinations of indicator functions $1_{[0,t]\times A}$ with $t \geq 0$ and $A \in \mathfrak{B}_b(\mathbb{R})$ (the bounded Borel subsets of $\mathbb{R}$). Let $\mathcal{H}$ be the completion of $\mathcal{E}$ with respect to the inner product

$$\langle 1_{[0,t]\times A}, 1_{[0,s]\times A}\rangle = \mathbf{E}(W_t^H(A)W_s^H(B)) = R^H(t,s)\lambda(A \cap B), \text{ for every } t, s \geq 0, A, B \in \mathfrak{B}_d(\mathbb{R}^n).$$

In particular (see [2])

$$\langle \varphi, \psi \rangle_{\mathcal{H}} = H(2H-1)\int_0^t \int_0^s dv_1 dv_2 |v_1 - v_2|^{2H-2} \int_{\mathbb{R}} dx \varphi(v_1, x)\psi(v_2, x)$$

for every $\varphi, \psi \in \mathcal{H}$ such that $\int_0^t \int_0^s dv_1 dv_2 |v_1 - v_2|^{2H-2} \int_{\mathbb{R}} dx |\varphi(v_1, x)\psi(v_2, x)| < \infty$.

Let I_q be the multiple stochastic integral of order q with respect to the isonormal process $(W(\varphi), \varphi \in \mathcal{H})$ (see the Appendix or [3]). Then

$$u(t,x) = \int_0^t \int_{\mathbb{R}} G_1(t-s, x-y)W^H(ds, dy) = I_1(g_{t,x})$$

where

$$g_{t,x}(s,y) = G_1(t-s, x-y)$$

and therefore the wavelet coefficient $d(t,a,i)$ given by (2.9) can be written as

$$d(t,a,i) = I_1(f_{t,a,i}) \text{ with } f_{t,a,i}(s,y) = \sqrt{a}\int_{\mathbb{R}} \psi(x)g_{t,a(x+i)}(s,y)dx \text{ for every } s > 0, y \in \mathbb{R}. \tag{4.29}$$

Then, by the product formula for multiple stochastic integrals (A.3), we have, for every $t > 0, a > 0$ and $N \geq 1$

$$V_N(t,a) = \frac{1}{N_a}\sum_{i=1}^{N_a}\left(\frac{I_2(f_{t,a,i}^{\otimes 2}) + \mathbf{E}d(t,a,i)^2}{\mathbf{E}d(t,a,i)^2} - 1\right)$$

$$= \frac{1}{N_a D(t,a)}\sum_{i=1}^{N_a} I_2(f_{t,a,i}^{\otimes 2}) \tag{4.30}$$

with $f_{t,a,i}$ given by (4.29).

Let us compute the L^2-norm of the random variable $V_N(t,a)$ given by (2.10). By using the isometry formula for multiple integrals (A.2),

$$
\begin{aligned}
\mathbf{E}V_N(t,a)^2 &= \frac{2}{D(t,a)^2 N_a^2} \sum_{i,j=1}^{N_a} \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}}^2 \\
&= \frac{2}{D(t,a)^2 N_a^2} \sum_{i,j=1}^{N_a} \left(\mathbf{E}d(t,a,i)d(t,a,j) \right)^2.
\end{aligned} \tag{4.31}
$$

Again we study the behavior of (4.31) as $N \to \infty$ when t varies with N and when t is fixed.

4.4. The moving time case: Proof of Proposition 3.1

Assume (2.11) and let us prove the limit theorem (3.1). The formula (4.31) becomes

$$
\mathbf{E}V_N(t_N, a_N)^2 = \frac{2}{N_{a_N}^2} D(t_N, a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} \left[-\frac{t_N}{4} a_N^{2H+1} g_H(i-j) + \frac{c_H}{2} a_N^{2H+2} g_{H+\frac{1}{2}}(i-j) \right]^2
$$

with g_H given by (3.2). Thus, with $g_{N,H}$ defined by (4.6),

$$
\begin{aligned}
\mathbf{E}V_N(t_N, a_N)^2 = \frac{2}{N_{a_N}^2} D(t_N, a_N)^{-2} \\
\times \left[\frac{t_N^2}{16} a_N^{4H+2} g_{N,H}(a_N) + \frac{c_H^2}{4} a_N^{4H+4} g_{N,H+\frac{1}{2}}(a_N) \right. \\
\left. -\frac{t_N c_H}{4} a_N^{4H+3} \sum_{i,j=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \right].
\end{aligned}
$$

We will use the notation $f_N \sim g_N$ which in our work means that the sequences f_N and g_N have the same limit as $N \to \infty$.

Under assumption (2.11), we can estimate $\mathbf{E}V_N(t,a)^2$ as follows

$$
\begin{aligned}
\mathbf{E}V_N(t_N, a_N)^2 &\sim \frac{32}{K_{\Psi,H}^2} N^{2(\alpha-1)} D(t_N, a_N)^{-2} \left[\frac{1}{16} N^{2\beta+(4H+2)\alpha} g_{N,H}(a_N) \right. \\
&\left. + \frac{c_H^2}{4} N^{\alpha(4H+4)} g_{N,H+\frac{1}{2}}(a_N) - \frac{c_H}{4} N^{\beta+\alpha(4H+3)} \sum_{i,j=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \right] \\
&:= v_{1,N} + v_{2,N} + v_{3,N}.
\end{aligned} \tag{4.32}
$$

Let us estimate the three summands above. By the estimate of $D(t_N, a_N)^{-2}$ in Lemma 4.1 and by (4.7), we have

$$
\begin{aligned}
v_{1,N} &= \frac{2}{K_{\Psi,H}^2} N^{2(\alpha-1)} D(t_N, a_N)^{-2} N^{2\beta+(4H+2)\alpha} N^{1-\alpha} g_{N,H}(N^{\alpha}) \\
&\sim \frac{2}{K_{\Psi,H}^2} N^{2(\alpha-1)} N^{-2\beta-2\alpha(2H+1)} N^{2\beta+(4H+2)\alpha} N^{1-\alpha} g_{N,H}(N^{\alpha})
\end{aligned}
$$

$$\sim \frac{2}{K_{\Psi,H}^2} \sum_{k \in \mathbb{Z}} g_H(k)^2 N^{\alpha-1}$$

with $K_{\Psi,H}$ from (3.3). Consequently

$$N^{1-\alpha} v_{1,N} \to_{N\to\infty} \frac{2}{K_{\Psi,H}^2} \sum_{k \in \mathbb{Z}} g_H(k)^2. \tag{4.33}$$

Let us look at the term $v_{2,N}$. By (4.8),

$$v_{2,N} \le C_{\Psi,H} N^{\alpha-1} N^{-2\beta+2\alpha} \times \begin{cases} C_{\Psi,H,Q} & \text{if } Q \ge 2, \\ C_{\Psi,H,Q} N_a^{4H-1} & \text{if } Q = 1 \end{cases}$$

$$\le C_{\Psi,H} N^{\alpha-1} \begin{cases} C_{\Psi,H,Q} N^{-2\beta+2\alpha} & \text{if } Q \ge 2, \\ C_{\Psi,H,Q} N^{-2\beta+2\alpha} N^{(4H-1)(1-\alpha)} = C_{\Psi,H,Q} N^{\alpha-1} N^{\alpha(3-4H)-2\beta+4H-1} & \text{if } Q = 1, H < \frac{3}{4}. \end{cases}$$

Thus

$$N^{1-\alpha} v_{2,N} \le C_{\Psi,H,Q} \begin{cases} N^{-2\beta+2\alpha} & \text{if } Q \ge 2, \\ N^{\alpha(3-4H)-2\beta+4H-1} & \text{if } Q = 1, H < \frac{3}{4} \end{cases} \to_{N\to\infty} 0 \tag{4.34}$$

because $\alpha < 1 < \beta$, $\alpha(3-4H) < 0$ and $-2\beta + 4H - 1 < -2\beta + 2 < 0$. Finally we look at $v_{3,N}$. We can write

$$v_{3,N} \le C_{\Psi,H} N^{\alpha-1} N^{\alpha-\beta} \times \begin{cases} C_{\Psi,H,Q} & \text{if } Q \ge 2, \\ C_{\Psi,H,Q} N_{a_N}^{4H-2} & \text{if } Q = 1 \end{cases}$$

$$\le C_{\Psi,H} N^{\alpha-1} \begin{cases} C_{\Psi,H,Q} N^{\alpha-\beta} & \text{if } Q \ge 2, \\ C_{\Psi,H,Q} N^{\alpha-\beta} N^{(4H-2)(1-\alpha)} = C_{\Psi,H,Q} N^{\alpha-1} N^{\alpha(3-4H)-\beta+4H-2} & \text{if } Q = 1, H < \frac{3}{4}. \end{cases}$$

Thus

$$N^{1-\alpha} v_{3,N} \le C_{\Psi,H,Q} \begin{cases} N^{\alpha-\beta} & \text{if } Q \ge 2, \\ N^{\alpha(3-4H)-\beta+4H-2} & \text{if } Q = 1, H < \frac{3}{4} \end{cases} \to_{N\to\infty} 0 \tag{4.35}$$

since $\alpha < \beta$, $\alpha(3-4H) < 0$ and $-\beta + 4H - 1 < -\beta + 1 < 0$.

The bounds (4.33), (4.34), (4.35) lead to the desired conclusion. $\qquad \square$

4.4.1. The fixed time case: Proof of Proposition 3.4

If t is fixed, we can prove the approximation result (3.8). We have

$$\mathbf{E} V_N(t, a_N)^2 = \frac{2}{N_{a_N}^2} \sum_{i,j=1}^{N_a} \frac{(\mathbf{E} d(t, a_N, i) d(t, a_N, j))^2}{\mathbf{E} d(t, a, i)^2 \mathbf{E} d(t, a, j)^2}$$

$$= \frac{2}{N_{a_N}^2} D(t, a_N)^{-2} \sum_{i,j=1}^{N_a} (\mathbf{E} d(t, a_N, i) d(t, a_N, j))^2$$

$$= \frac{2}{N_{a_N}^2} D(t, a_N)^{-2} \sum_{i,j=1}^{N_a} (f_{H,N}(i-j))^2$$

with $f_{H,N}$ given by (4.21). Notice that $f_{H,N}(k) = f_{H,N}(-k)$ and

$$f_{H,N}(k) = 0 \text{ if } |k| \geq 2$$

by chosing N large enough. Since can be seen *via* (4.21), since the function Ψ has support included in $[0,1]$. Therefore

$$\mathbf{E}V_N(t, a_N)^2 = \frac{2}{N_{a_N}^2} D(t, a_N)^{-2} \left[N_a(f_{H,N}(0))^2 + 2(N_a - 1)(f_{H,N}(1))^2 \right]. \tag{4.36}$$

We have $f_{H,N}(0) = D(t, a_N)$ and $f_{H,N}(1)$ was computed before. Using (4.28), (4.36) can be written as follows

$$\mathbf{E}V_N(t, a_N)^2 = \frac{2}{N^{2(1-\alpha)}} D(t, a_N)^{-2} \left(N^{1-\alpha} D(t, a_N)^2 + \frac{2(N^{1-\alpha} - 1)L_t^2(H)}{D(t, a_N)^2 N^{2\alpha}} \right) \tag{4.37}$$

with $L_t(H)$ given by (4.28). Then

$$\mathbf{E}V_N(t, a_N)^2 \sim \left(\frac{2}{N^{1-\alpha}} + \frac{4L_t(H)^2}{K_{1,t}(H)^2} \frac{1}{N^{1+\alpha}} \right)$$

and the conclusion follows. $\square$

Remark 4.6. As already noticed in Remark 3.6, the renormalization of (2.10) is of the same order in both cases (fixed time or moving time) although the correlation structure of the wavelet coefficient is different. On the other hand, in the fixed time case, the diagonal term of $\mathbf{E}V_N(t, a_N)^2$ is dominant for the behavior of this quantity as $N \to \infty$ (here is only one non-diagonal term which does not contribute to the limit), while when t increases with N, all the diagonal and non-diagonal terms have contribution to the limit.

4.5. Central limit theorem and rate of convergence

We will show that, both in the moving time and fixed time cases, the renormalized wavelet variation satisfies a central limit theorem if $Q \geq 2$ or $Q = 1, H < \frac{3}{4}$.

Our main tool is the following result (see Thm. 5.2.6 and Cor. 5.2.10 in [16]). Recall that by d we denote the distance between distributions of random variable and below it can be each of the following distances: Kolmogorov, total variation, Wasserstein or Fortet-Mourier (see [16]).

Theorem 4.7. *Let $(F_N)_{N \geq 1}$ be a sequence of random variables in the qth Wiener chaos $(q \geq 1)$ with respect to an isonormal process indexed by the Hilbert space $\mathcal{H}$. Assume that $\mathbf{E}F_N^2 \to_{N \to \infty} \sigma^2 > 0$. Then the sequence $(F_N)_{N \geq 1}$ converges in law to the standard normal random variable Z if and only if $\|DF_N\|_{\mathcal{H}}^2$ converges in $L^2(\Omega)$ as $N \to \infty$ to $q\sigma^2$. In this case*

$$d(F_N, Z) \leq C \left(\sqrt{Var\left(\|DF_N\|_{\mathcal{H}}^2\right)} + \left|\mathbf{E}F_N^2 - \sigma^2\right| \right).$$

4.6. The moving time: Proof of Theorems 3.2 and 3.3

Consider the sequence $(F_N)_{N \geq 1}$ given by (3.4) and recall that from Proposition 3.1

$$\mathbf{E}F_N^2 \to_{N \to \infty} 1.$$

Also, by (4.30) we have the following chaos expansion of F_N, for every $N \geq 1$, with $f_{t,a,i}$ given by (4.29),

$$F_N = K_{0,\Psi,H}^{-\frac{1}{2}} N^{\frac{\alpha-1}{2}} I_2 \left(\sum_{i=1}^{N_{a_N}} \frac{f_{t_N,a_N,i}^{\otimes 2}}{\mathbf{E}(d(t_N,a_N,i)^2)} \right) = K_{0,\Psi,H}^{-\frac{1}{2}} N^{\frac{\alpha-1}{2}} D(t_N,a_N)^{-1} I_2 \left(\sum_{i=1}^{N_{a_N}} f_{t_N,a_N,i}^{\otimes 2} \right). \tag{4.38}$$

So, F_N is an element of the second Wiener chaos (with respect to the Gaussian noise W^H introduced in Section 4.3) for every $N \geq 1$ and we may apply Theorem 4.7 in order to check its asymptotic behavior in distribution.

Proof of Theorem 3.2. By taking the Malliavin derivative with respect to fractional-white noise W^H in (4.38) (see formula (A.4)),

$$DF_N = 2 K_{0,\Psi,H}^{-\frac{1}{2}} N^{\frac{\alpha-1}{2}} D(t_N,a_N)^{-1} \sum_{i=1}^{N_{a_N}} I_1(f_{t_N,a_N,i}) f_{t_N,a_N,i}$$

and, if $\mathcal{H}$ is the Hilbert space associated with the fractional-white Gaussian noise (see the beginning of Section 4.3),

$$\begin{aligned}
\|DF_N\|_{\mathcal{H}}^2 &= 4 K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N,a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_1(f_{t,a,i}) I_1(f_{t,a,j}) \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}} \\
&= 4 K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N,a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_2(f_{t,a,i} \otimes f_{t,a,j}) \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}} \\
&\quad + 4 K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N,a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}}^2 \\
&= 4 K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N,a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_2(f_{t,a,i} \otimes f_{t,a,j}) \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}} + \mathbf{E}\|DF_N\|_{\mathcal{H}}^2
\end{aligned}$$

with $f_{t,a,i}$ from (4.29). Notice that, since F_N belongs to the second Wiener chaos, we have $\mathbf{E}\|DF_N\|_{\mathcal{H}}^2 = 2\mathbf{E}F_N^2$. By Theorem 4.7,

$$d(F_N, Z) \leq c \left(\sqrt{Var\left(\|DF_N\|_{\mathcal{H}}^2\right)} + \mathbf{E}\|DF_N\|_{\mathcal{H}}^2 - 2 \right)$$

so

$$d(F_N, Z) \leq c(\sqrt{T_{1,N}} + T_{2,N})$$

where

$$T_{1,N} = \mathbf{E} \left(4 K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N,a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_2(f_{t,a,i} \otimes f_{t,a,j}) \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}} \right)^2 \tag{4.39}$$

and

$$T_{2,N} = \mathbf{E}\|DF_N\|_{\mathcal{H}}^2 - 2 = 4K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N, a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} \langle f_{t,a,i}, f_{t,a,j} \rangle_{\mathcal{H}}^2 - 2. \tag{4.40}$$

Let us first estimate $T_{2,N}$. Since

$$\langle f_{t_N,a_N,i} f_{t_N,a_N,j} \rangle_{\mathcal{H}} = \mathbf{E}d(t_N, a_N, i)d(t_N, a_N, j)$$

we can write, as in (4.32)

$$\begin{aligned}
T_{2,N} &= 4K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N, a_N)^{-2} \left[\frac{1}{16} N^{2\beta + (4H+2)\alpha} g_{N,H}(a_N) \right. \\
&\quad \left. + \frac{c_H^2}{4} N^{\alpha(4H+4)} g_{N,H+\frac{1}{2}}(a_N) - \frac{c_H}{4} N^{\beta+\alpha(4H+3)} \sum_{i,j=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \right] - 2 \\
&:= T_{2,1,N} + T_{2,2,N} + T_{2,3,N}.
\end{aligned}$$

First, we analyze the term $T_{2,1,N}$. We have

$$\begin{aligned}
T_{2,1,N} &= 4K_{0,\Psi,H}^{-1} N^{\alpha-1} D(t_N, a_N)^{-2} \frac{1}{16} N^{2\beta + (4H+2)\alpha} g_{N,H}(a_N) - 2 \\
&= \frac{1}{4} N^{2\beta + (4H+2)\alpha} K_{0,\Psi,H}^{-1} D(t_N, a_N)^{-2} \sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} \frac{N_{a_N} - |k|}{N_{a_N}} - 2 \\
&= \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta + (4H+2)\alpha} D(t_N, a_N)^{-2} \sum_{k\in\mathbb{Z}} g_H(k)^2 \\
&\quad + \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta + (4H+2)\alpha} D(t_N, a_N)^{-2} \left(\sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} \frac{N_{a_N} - |k|}{N_{a_N}} - \sum_{k\in\mathbb{Z}} g_H(k)^2 \right) - 2
\end{aligned}$$

and since by (4.4)

$$D(t_N, a_N) = \mathbf{E}d(t_N, a_N, i)^2 = -\frac{c_H}{2} K_{\Psi,H+\frac{1}{2}} N^{(2H+2)\alpha} + \frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha},$$

we obtain

$$\begin{aligned}
T_{2,1,N} &= \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta + (4H+2)\alpha} \left(\frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} \sum_{k\in\mathbb{Z}} g_H(k)^2 \\
&\quad + \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta + (4H+2)\alpha} \\
&\quad \times \left[\left(-\frac{c_H}{2} K_{\Psi,H+\frac{1}{2}} N^{(2H+2)\alpha} + \frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)H} \right)^{-2} - \left(\frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} \right] \sum_{k\in\mathbb{Z}} g_H(k)^2 \\
&\quad + \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta + (4H+2)\alpha} D(t_N, a_N)^{-2} \left(\sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} \frac{N_{a_N} - |k|}{N_{a_N}} - \sum_{k\in\mathbb{Z}} g_H(k)^2 \right) - 2.
\end{aligned}$$

The first term in the above expression vanishes with 2 so it remains

$$
\begin{aligned}
T_{2,1,N} = {}& \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta+(4H+2)\alpha} \\
& \times \left[\left(-\frac{c_H}{2} K_{\Psi,H+\frac{1}{2}} N^{(2H+2)\alpha} + \frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} - \left(\frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} \right] \sum_{k\in\mathbb{Z}} g_H(k)^2 \\
& + \frac{1}{4} K_{0,\Psi,H}^{-1} N^{2\beta+(4H+2)\alpha} D(t_N,a_N)^{-2} \left(\sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} \frac{N_{a_N}-|k|}{N_{a_N}} - \sum_{k\in\mathbb{Z}} g_H(k)^2 \right).
\end{aligned}
$$

We have the following bound for the first summand in $T_{2,1,N}$

$$
\left| \left(-\frac{c_H}{2} K_{\Psi,H+\frac{1}{2}} N^{(2H+2)\alpha} + \frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} - \left(\frac{1}{4} K_{\Psi,H} N^{\beta+(2H+1)\alpha} \right)^{-2} \right|
$$
$$
\le c N^{\alpha-\beta} N^{2\beta+(4H+2)\alpha}.
$$

To obtain a bound for the second term in the expression of $T_{2,1,N}$, we write $\sum_{k\in\mathbb{Z}} g_H(k)^2 = \sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} + \sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|>N_{a_N}\}}$ and using the fact that $|g_H|$ is bounded by $|k|^{4H-4Q}$ we get that

$$
\left| D(t_N,a_N)^{-2} \left(\sum_{k\in\mathbb{Z}} g_H(k)^2 1_{\{|k|\le N_{a_N}\}} \frac{N_{a_N}-|k|}{N_{a_N}} - \sum_{k\in\mathbb{Z}} g_H(k)^2 \right) \right|
$$
$$
\le c D(t_N,a_N)^{-2} N_{a_N}^{8H-8Q+2} = c N^{-2\beta-2\alpha(2H+1)+(1-\alpha)(8H-8Q+2)}.
$$

Therefore

$$
|T_{2,1,N}| \le c N^{\alpha-\beta}. \tag{4.41}
$$

For $T_{2,2,N}$ we have by (4.8) and Lemma 4.1

$$
\begin{aligned}
|T_{2,2,N}| &\le c D(t_N,a_N)^{-2} N^{\alpha-1} N^{\alpha(4H+4)} g_{H,\frac{1}{2}}(a_N) \\
&\le D(t_N,a_N)^{-2} N^{\alpha(4H+4)} \begin{cases} C, \text{ if } Q \ge 2 \\ N^{(1-\alpha)(4H-1)}, \text{ if } Q = 1 \end{cases} \\
&\le c \begin{cases} N^{2\alpha-2\beta} \text{ if } Q \ge 2 \\ N^{2\alpha-2\beta} N^{(1-\alpha)(4H-1)} \text{ if } Q = 1. \end{cases}
\end{aligned}
$$

Regarding $T_{2,3,N}$ we use (4.8) and Lemma 4.1 to get that

$$
\begin{aligned}
|T_{2,3,N}| &\le c D(t_N,a_N)^{-2} N^{\alpha-1} N^{\alpha(4H+3)} \sum_{i=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(i-j) \\
&\le D(t_N,a_N)^{-2} N^{\alpha(4H+3)} \begin{cases} C, \text{ if } Q \ge 2 \\ N^{(1-\alpha)(4H-2)}, \text{ if } Q = 1 \end{cases}
\end{aligned}
$$

$$\leq c \begin{cases} N^{\alpha-\beta} \text{ if } Q \geq 2 \\ N^{\alpha-\beta}N^{(1-\alpha)(4H-2)} \text{ if } Q = 1. \end{cases}$$

Combining (4.41), (4.42) and (4.42), we have the following bound for (4.40):

$$|T_{2,N}| \leq \begin{cases} CN^{\alpha-\beta} \text{ if } Q \geq 2 \\ N^{\alpha-\beta}N^{(1-\alpha)(4H-2)} \text{ if } Q = 1. \end{cases} \tag{4.42}$$

Concerning $T_{1,N}$, by the isometry of multiple integrals

$$T_{1,N} = Var(\|DF_N\|_{\mathcal{H}}^2) = \mathbf{E}\left(|\|DF_N\|_{\mathcal{H}}^2 - \mathbf{E}\|DF_N\|_{\mathcal{H}}^2\right)^2$$

$$= \mathbf{E}\left(4K_{0,\Psi,H}^{-1}N^{\alpha-1}D(t_N,a_N)^{-2}\sum_{i,j=1}^{N_{a_N}}I_2(f_{t_N,a_N,i}\otimes f_{t_N,a_N,j})\langle f_{t_N,a_N,i},f_{t_N,a_N,j}\rangle_{\mathcal{H}}\right)^2$$

$$= 16K_{0,\Psi,H}^{-2}N^{2(\alpha-1)}D(t_N,a_N)^{-4}$$

$$\mathbf{E}\left(\sum_{i,j,k,l=1}^{N_{a_N}}I_2(f_{t_N,a_N,i}\otimes f_{t_N,a_N,j})I_2(f_{t_N,a_N,k}\otimes f_{t_N,a_N,l})\langle f_{t_N,a_N,i};f_{t_N,a_N,j}\rangle_{\mathcal{H}}\langle f_{t_N,a_N,k},f_{t_N,a_N,l}\rangle_{\mathcal{H}}\right)$$

$$= 32K_{0\Psi,H}^{-2}N^{2(\alpha-1)}D(t_N,a_N)^{-4}\sum_{i,j,k,l=1}^{N_{a_N}}\langle f_{t,a,i};f_{t,a,j}\rangle_{\mathcal{H}}\langle f_{t,a,k};f_{t,a,l}\rangle_{\mathcal{H}}\langle f_{t_N,a_N,i},f_{t_N,a_N,k}\rangle_{\mathcal{H}}\langle f_{t,a,j};f_{t,a,l}\rangle_{\mathcal{H}}.$$

Recall that for all integers p,q we have (see relation (4.5)) $\langle f_{t,a,p};f_{t,a,q}\rangle = \frac{c_H}{2}a_N^{2H+2}g_{H+\frac{1}{2}}(p-q) - \frac{t_N}{4}a_N^{2H+1}g_H(p-q)$. Hence,

$$T_{1,N} = 32K_{0,\Psi,H}^{-2}N^{2(\alpha-1)}D(t_N,a_N)^{-4}\sum_{i,j,k,l=1}^{N_{a_N}}\left[\frac{c_H}{2}a_N^{2H+2}g_{H+\frac{1}{2}}(i-j) - \frac{t_N}{4}a_N^{2H+1}g_H(i-j)\right]$$

$$\times\left[\frac{c_H}{2}a_N^{2H+2}g_{H+\frac{1}{2}}(k-l) - \frac{t_N}{4}a_N^{2H+1}g_H(k-l)\right]\left[\frac{c_H}{2}a_N^{2H+2}g_{H+\frac{1}{2}}(i-k) - \frac{t_N}{4}a_N^{2H+1}g_H(i-k)\right]$$

$$\times\left[\frac{c_H}{2}a_N^{2H+2}g_{H+\frac{1}{2}}(j-l) - \frac{t_N}{4}a_N^{2H+1}g_H(j-l)\right]$$

$$= \sum_{i=1}^{5}T_{1,i,N}.$$

Above we denoted by $T_{1,1,N}$ a product that depends only on functions g_H and $T_{1,2,N}$ is a product that depends only on one function g_H (see (3.2)) and three functions $g_{H+\frac{1}{2}}$. $T_{1,3,N}$ is the product of two functions g_H and two $g_{H+\frac{1}{2}}$. As well, $T_{1,4,N}$ contains three g_H and one $g_{H+\frac{1}{2}}$ while $T_{1,5,N}$ contains the four functions $g_{H+\frac{1}{2}}$. To study each term we will use the same technique of proof of the Theorem 7.3.1 in [16] or Lemma 3 in [12]. Let $g_{H,N}(k) := |g_H(k)|1_{\{|k|\leq N_a\}}$ (this is not the same as $g_{N,H}$ in (4.6)) and $g_{H+\frac{1}{2},N}(k) := |g_{H+\frac{1}{2}}(k)|1_{\{|k|\leq N_a\}}$.

Let $u,v : \mathbb{Z} \to \mathbb{R}$ be two sequences, we define their convolution by

$$(u*v)(k) = \sum_{n\in\mathbb{Z}}u(n)v(k-n).$$

We will need Young's inequality, which can be written, for $s, p, q \geq 1$ such that $\frac{1}{s} + 1 = \frac{1}{p} + \frac{1}{q}$ as

$$\|u * v\|_{l^s(\mathbb{Z})} \leq \|u\|_{l^p(\mathbb{Z})} \|v\|_{l^q(\mathbb{Z})}.$$

Now we can estimate each term $T_{1,i,N}$, for $i = 1, \ldots, 5$. First

$$|T_{1,1,N}| = \left| 32 K_0^{-2} N^{2(\alpha-1)} \frac{t_N^4}{4^4} a_N^{4(2H+1)} D(t_N, a_N)^{-4} \sum_{i,j,k,l=1}^{N_{a_N}} g_H(i-j) g_H(k-l) g_H(i-k) g_H(j-l) \right|$$

$$\leq C N^{\alpha(8H+6)-2+4\beta} D(t_N, a_N)^{-4} \sum_{j,k=1}^{N_{a_N}} \sum_{i,l \in \mathbb{Z}} g_{H,N}(i-j) g_{H,N}(k-l) g_{H,N}(i-k) g_{H,N}(j-l)$$

$$= C N^{\alpha(8H+6)-2+4\beta} D(t_N, a_N)^{-4} \sum_{j,k=1}^{N_{a_N}} (g_{H,N} * g_{H,N})(k-j)^2$$

$$= C N^{\alpha(8H+6)-2+4\beta} N^{1-\alpha} D(t_N, a_N)^{-4} \|g_{H,N} * g_{H,N}\|_{l^2(\mathbb{Z})}^2$$

$$\leq C N^{\alpha(8H+5)-1+4\beta} D(t_N, a_N)^{-4} \|g_{H,N}\|_{l^{\frac{4}{3}}(\mathbb{Z})}^4.$$

where the last inequality follows from Young inequality

$$|T_{1,1,N}| \leq C N^{\alpha(8H+5)-1+4\beta} D(t_N, a_N)^{-4} \left(\sum_{k=-N}^{N} |g_H(k)|^{4/3} \right)^3.$$

The series $\sum_{k \in \mathbb{Z}} |g_H(k)|^{4/3}$ converges because $|g_H(k)|$ is bounded, for k large, by $C|k|^{4H-4Q}$, see Lemma 4.3. Finally,

$$|T_{1,1,N}| \leq C N^{\alpha-1}. \tag{4.43}$$

Next, we have

$$T_{1,2,N} = -C N^{2(\alpha-1)} D(t_N, a_N)^{-4} N^{\beta+\alpha(8H+7)} \sum_{i,j,k,l=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(k-l) g_{H+\frac{1}{2}}(i-k) g_{H+\frac{1}{2}}(j-l).$$

We apply the same idea as above

$$|T_{1,2,N}| \leq C N^{2(\alpha-1)} D(t_N, a_N)^{-4} N^{\beta} N^{\alpha(8H+7)} \sum_{i,j,k,l=1}^{N_{a_N}} g_H(i-j) g_{H+\frac{1}{2}}(k-l) g_{H+\frac{1}{2}}(i-k) g_{H+\frac{1}{2}}(j-l)$$

$$\leq C N^{\beta-2+\alpha(8H+9)} D(t_N, a_N)^{-4} \sum_{j,k=0}^{N_{a_N}} \sum_{i,l \in \mathbb{Z}} g_{H,N}(i-j) g_{H+\frac{1}{2},N}(k-l) g_{H+\frac{1}{2},N}(i-k) g_{H+\frac{1}{2},N}(j-l)$$

$$\leq C N^{\beta-1+\alpha(8H+8)} D(t_N, a_N)^{-4} \sum_{k \in \mathbb{Z}} (g_{H,N} * g_{H+\frac{1}{2},N})(k)(g_{H+\frac{1}{2},N} * g_{H+\frac{1}{2},N})(k)$$

$$\leq C N^{\beta-1+\alpha(8H+8)} D(t_N, a_N)^{-4} \|g_{H,N} * g_{H+\frac{1}{2},N}\|_{l^2(\mathbb{Z})} \|g_{H+\frac{1}{2},N} * g_{H+\frac{1}{2},N}\|_{l^2(\mathbb{Z})}$$

$$\leq C N^{\beta-1+\alpha(8H+8)} D(t_N, a_N)^{-4} \|g_{H,N}\|_{l^{\frac{4}{3}}(\mathbb{Z})} \|g_{H+\frac{1}{2},N}\|_{l^{\frac{4}{3}}(\mathbb{Z})}^3$$

$$= CN^{\beta-1+\alpha(8H+8)}D(t_N,a_N)^{-4}\left(\sum_{k=-N}^{N}|g_{H,N}|^{\frac{4}{3}}\right)^{\frac{4}{3}}\left(\sum_{k=-N}^{N}|g_{H+\frac{1}{2},N}|^{\frac{4}{3}}\right)^{\frac{9}{4}}$$

$$\leq CN^{4\alpha-3\beta-1}\left(\sum_{k=-N}^{N}|g_{H,N}|^{\frac{4}{3}}\right)^{\frac{3}{4}}\left(\sum_{k=-N}^{N}|g_{H+\frac{1}{2},N}|^{\frac{4}{3}}\right)^{\frac{9}{4}}.$$

Now, by Lemma 4.3, $g_{H+\frac{1}{2}}(k)$ is bounded for k large by $C|k|^{4H+2-4Q}$ and this yields to

$$\sum_{k=-N}^{N}|g_{H+\frac{1}{2},N}|^{\frac{4}{3}} \leq \begin{cases} O(1) & \text{if } Q \geq 2 \text{ and } Q = 1, H < \frac{5}{16} \\ O(\log(N)) & \text{if } Q = 1, H = \frac{5}{16} \\ O(N^{\frac{16H-5}{3}}) & \text{if } Q = 1, H > \frac{5}{16}. \end{cases}$$

Recall that $\frac{1}{2} < H < 1$, hence

$$|T_{1,2,N}| \leq C \begin{cases} N^{4\alpha-3\beta-1} & \text{if } Q \geq 2 \\ N^{\frac{3(16H-5)}{4}+4\alpha-3\beta-1} & \text{if } Q = 1, H \in (\frac{1}{2}, \frac{3}{4}). \end{cases} \qquad (4.44)$$

In the same way,

$$|T_{1,3,N}| = C\left|N^{2(\alpha-1)}N^{\alpha(8H+6)+2\beta}D(t_N,a_N)^{-4}\sum_{i,j,k,l=1}^{N_{a_N}}g_H(i-j)g_H(k-l)g_{H+\frac{1}{2}}(i-k)g_{H+\frac{1}{2}}(j-l)\right|$$

$$\leq CN^{3\alpha-2\beta-1}\left(\sum_{k=-N}^{N}|g_{H,N}|^{\frac{4}{3}}\right)^{\frac{3}{2}}\left(\sum_{k=-N}^{N}|g_{H+\frac{1}{2},N}|^{\frac{4}{3}}\right)^{\frac{3}{2}}.$$

Using the previous case we have

$$T_{1,3,N} \leq C \begin{cases} N^{3\alpha-2\beta-1} & \text{if } Q \geq 2 \\ N^{\frac{16H-5}{2}+3\alpha-2\beta-1} & \text{if } Q = 1, H \in (\frac{1}{2}, \frac{3}{4}) \end{cases} \qquad (4.45)$$

and as well,

$$|T_{1,4,N}| = C\left|N^{2(\alpha-1)}N^{\alpha(8H+5)+3\beta}D(t_N,a_N)^{-4}\sum_{i,j,k,l=1}^{N_{a_N}}g_H(i-j)g_H(k-l)g_H(i-k)g_{H+\frac{1}{2}}(j-l)\right|$$

$$\leq CN^{2\alpha-\beta-1}\left(\sum_{k=-N}^{N}|g_{H,N}|^{\frac{4}{3}}\right)^{\frac{9}{4}}\left(\sum_{k=-N}^{N}|g_{H+\frac{1}{2},N}|^{\frac{4}{3}}\right)^{\frac{3}{4}}.$$

which implies that

$$|T_{1,4,N}| \leq C \begin{cases} N^{2\alpha-\beta-1} & \text{if } Q \geq 2 \\ N^{\frac{16H-5}{4}+\alpha-\beta-1} & \text{if } Q = 1, H \in (\frac{1}{2}, \frac{3}{4}) \end{cases}. \qquad (4.46)$$

It remains to look at $T_{1,5,N}$,

$$|T_{1,5,N}| = C \left| N^{2(\alpha-1)} N^{4\alpha(2H+2)} D(t_N, a_N)^{-4} \sum_{i,j,k,l=1}^{N_{a_N}} g_{H+\frac{1}{2}}(i-j) g_{H+\frac{1}{2}}(k-l) g_{H+\frac{1}{2}}(i-k) g_{H+\frac{1}{2}}(j-l) \right|$$

$$\leq C N^{5\alpha-4\beta-1} \left(\sum_{k=-N}^{N} |g_{H+\frac{1}{2},N}|^{\frac{4}{3}} \right)^3$$

and this gives us

$$|T_{1,5,N}| \leq \begin{cases} N^{5\alpha-4\beta-1} & \text{if } Q \geq 2 \\ N^{16H+5\alpha-4\beta-6} & \text{if } Q = 1, H \in (\frac{1}{2}, \frac{3}{4}). \end{cases}$$

Combining $(34)-(38)$ we see that $T_{1,1,N}$ is the biggest term and finally we get a simple estimate for $T_{1,N}$

$$|T_{1,N}| \leq C N^{\alpha-1}. \tag{4.47}$$

By (4.42) and (4.47), we obtain the conclusion. $\qquad\square$

Proof of Theorem 3.3. By Theorem 3.2 (applied to the scale La_N, $L = 1, \ldots, d$), we know that the each component of the vector $\left(N^{1-\alpha} V_n(t_N, La_N) \right)_{L=1,\ldots,d}$ converges in distribution, as $N \to \infty$, to a centered Gaussian random variable. By the main result in [19], it suffices to show that for every $L_1, L_2 = 1, \ldots, d$

$$N^{1-\alpha} \mathbf{E} V_N(t_N, L_1 a_N) V_N(t_N, L_2 a_N)$$

converges as $N \to \infty$ to Γ_{L_1, L_2}. We have

$$N^{1-\alpha} \mathbf{E} V_N(t_N, L_1 a_N) V_N(t_N, L_2 a_N)$$
$$= \frac{2}{D(t_N, L_1 a_N) D(t_N, L_2 a_N)} N^{1-\alpha} \sum_{i=1}^{N_{L_1 a_A}} \sum_{j=1}^{N_{L_2 a_A}} \left(\mathbf{E} d(t_N, L_1 a_N, i) d(t_N, L_2 a_N, j) \right)^2$$

where

$$\mathbf{E} d(t_N, L_1 a_N, i) d(t_N, L_2 a_N, j) = \frac{c_H}{2} a_N^{2H+2} g_{L_1, L_2, H+\frac{1}{2}}(i, j) - \frac{t_N}{4} a_N^{2H+1} g_{L_1, L_2, H}(i, j)$$

and

$$g_{L_1, L_2, H}(i, j) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x \mathrm{d}y \Psi(x) \Psi(y) |L_1 x - L_2 y + L_1 - L_2 j|^{2H}.$$

We know (see Sect. 3.2 in [7] or Prop. 2.3 in [5]), that for every i, j large and for $H \in \left(\frac{1}{2}, \frac{3}{2} \right)$,

$$|g_{L_1, L_2, H}(i, j)| \leq C(1 + |L_1 i - L_2 j|^{2H-2Q}) \tag{4.48}$$

and if $H \in \left(\frac{1}{2}, 1\right)$,

$$N^{1-\alpha} \sum_{i=1}^{N_{L_1 a_N}} \sum_{j=1}^{N_{L_2 a_N}} (g_{L_1, L_2, H}(i, j))^2 \to_{N \to \infty} C(L_1, L_2, H) \tag{4.49}$$

with $C(L_1, L_2, H)$ being explicit constant. From Lemma 4.1, (4.49) and (4.48) we obtain the conclusion by following the proof of Proposition 3.1. $\qquad\square$

4.6.1. The case of fixed time: Proof of Theorem 3.5

Now consider the random sequence $(G_N)_{N \geq 1}$ defined by (3.9). It satisfies $\mathbf{E}G_N^2 \to 1$ as $N \to \infty$ and it admits the following chaos expansion

$$G_N =: \frac{1}{\sqrt{2}} N^{\frac{1-\alpha}{2}} V(t, a_N) = \frac{1}{\sqrt{2}} N^{\frac{1-\alpha}{2}} D(t, a_N)^{-1} I_2 \left(\sum_{i=1}^{N} f_{t, a_N, i}^{\otimes 2} \right). \tag{4.50}$$

From (4.50) and (A.4), the Malliavin derivative of G_N writes as

$$DG_N = \frac{1}{\sqrt{2}} N^{\frac{1-\alpha}{2}} d(t, a_N)^{-2} I_1 \left(\sum_{i=1}^{N} f_{t, a_N, i} \right) f_{t, a_N, i}$$

so

$$\|DG_N\|_{\mathcal{H}}^2$$
$$= 2N^{\alpha-1} D(t, a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_2(f_{t, a_N, i} \otimes f_{t, a_N, i}) f_{t, a_N, j} + 2N^{\alpha-1} D(t, a_N)^{-2} \sum_{i,j=1}^{N} \langle (f_{t, a_N, i}, f_{t, a_N, j}) \rangle^2$$
$$= 2N^{\alpha-1} D(t, a_N)^{-2} \sum_{i,j=1}^{N_{a_N}} I_2(f_{t, a_N, i} \otimes f_{t, a_N, i}) f_{t, a_N, j} + \mathbf{E}\|DG_N\|^2.$$

From Theorem 4.7,

$$d(G_N, Z) \leq c \left(\sqrt{Var\left(\|DG_N\|_{\mathcal{H}}^2\right)} + \mathbf{E}\|DG_N\|_{\mathcal{H}}^2 - 2 \right)$$
$$=: c(\sqrt{T_{1,N}} + T_{2,N}).$$

We analyze first $T_{1,N}$,

$$|T_{1,N}| = \left| 8N^{2(\alpha-1)} D(t, a_N)^{-4} \sum_{i,j,k,l=1}^{N_{a_N}} f_{H,N}(i-j) f_{H,N}(k-l) f_{H,N}(i-k) f_{H,N}(j-l) \right|$$
$$\leq CN^{\alpha-1} D(t, a_N)^{-4} \|\tilde{f}_{H,N}\|_{l^{4/3}(\mathbb{Z})}^4$$
$$\leq CN^{\alpha-1} D(t, a_N)^{-4} \left(\sum_{k=-N_{a_N}}^{N_{a_N}} |f_{H,N}(k)|^{4/3} \right)^3$$

$$= CN^{\alpha-1}D(t,a_N)^{-4}\left(2\sum_{k=1}^{N_{a_N}}|f_{H,N}(k)|^{4/3}+|f_{H,N}(0)|^{4/3}\right)^3$$

$$= CN^{\alpha-1}D(t,a_N)^{-4}\left(2|f_{H,N}(1)|^{4/3}+|f_{H,N}(0)|^{4/3}\right)^3$$

$$= CN^{\alpha-1}D(t,a_N)^{-4}\left(2\frac{|K_{3,t}|^{4/3}}{N^{\frac{4\alpha}{3}}}+|D(t,a_N)|^{4/3}\right)^3$$

$$\leq CN^{\alpha-1}D(t,a_N)^{-4}\left(8\frac{|K_{3,t}|^4}{N^{4\alpha}}+4|D(t,a_n)|^4\right)\leq C\frac{1}{N^{1-\alpha}}.$$

For $T_{2,N}$ we can write

$$|T_{2,N}| = \left|2N^{\alpha-1}D(t,a_N)^{-2}\sum_{i,j=1}^{N_{a_N}}\langle(f_{t,a_N,i};f_{t,a_N,j})\rangle^2-2\right|$$

$$= \left|2N^{\alpha-1}D(t,a_N)^{-2}\left(N_{a_N}f_{H,N}(0)^2+2(N_{a_N}-1)f_{H,N}(1)^2\right)-2\right|$$

$$= \left|4N^{\alpha-1}D(t,a_N)^{-2}(N_{a_N}-1)\frac{K_3^2}{N^{2\alpha}}\right|\leq\frac{C}{N^{2\alpha}}.$$

Therefore we obtain

$$d(G_N,Z)\leq c\left(\sqrt{Var\left(\|DF_N\|_{\mathcal{H}}^2\right)}+\mathbf{E}\|DF_N\|_{\mathcal{H}}^2-2\right)$$

$$= c\left(\frac{1}{N^{\frac{1-\alpha}{2}}}+\frac{1}{N^{2\alpha}}\right).$$

$\square$

5. Estimation of the Hurst parameter

We will apply our theoretical results in Section 3 in order to construct an estimator for the Hurst parameter of the solution to the stochastic wave equation (2.1). The estimator will be constructed by using the wavelet variation (2.10). We will assume that the solution is observed at discrete points in space $x_i = i, i = 1, \ldots, N$ and at a certain time t (fixed or depending on N). Different estimators (but all of them constructed *via* the wavelet variation) are obtained in these two situations treated in our work (moving or fixed time). While for fixed time the logarithm of the variance of the wavelet coefficient depends linearly on H (Lem. 4.2), a linear log–log regression will give the explicit form of the estimator. For fixed time, this variance of the wavelet coefficient have a more complex dependence on the Hurst parameter (Prop. 4.4) and a different argument will be employed.

5.1. The moving time case

First we introduce a discrete version of the wavelet variation (2.10). Then we define an estimator in terms of the discrete wavelet variation and we prove its asymptotic properties.

5.2. Discretization of the wavelet variation

We will use an estimator constructed by using the wavelet variation (2.10) or more precisely, by using its discretized version defined below. Notice that the wavelet coefficient $d(t,a,i)$ is defined as a continuous integral

(see (2.9)) and it cannot be observed directly when the process u is observed. Therefore, by approximating the integral in (2.9) by Riemann sums, we define the discrete wavelet coeffcient, for $a > 0, t > 0$,

$$e_N(t, a, i) = \frac{1}{\sqrt{a}} \sum_{k=1}^{N} \Psi\left(\frac{k}{a} - i\right) u(t, k). \quad N \geq 1.$$

Since Ψ has its support contained in the interval $[0, 1]$, the above coefficient can also expressed as

$$e_N(t, a, i) = \frac{1}{\sqrt{a}} \sum_{k=0}^{[a]} \Psi\left(\frac{k}{a}\right) u(t, k + ai). \tag{5.1}$$

Let us also define the discrete version of the wavelet variation by setting

$$\widehat{V}_N(t, a) = \frac{1}{N_a} \sum_{i=1}^{N_a} \left(\tilde{e}(t, a, i)^2 - 1\right) \tag{5.2}$$

with, see notation (4.1),

$$\tilde{e}_N(t, a, i) = \frac{e_N(t, a, i)}{(\mathbf{E}(d(a, t, i))^2)^{\frac{1}{2}}} = \frac{1}{\sqrt{D(t, a)}} e_N(t, a, i).$$

In a first step, we will show that the sequence $\widehat{V}_N(t_N, a_N)$ has the same limit behavior in distribution as $V_N(t_N, a_N)$ when N goes to infinity. We need to assume some differentiability of the mother wavelet (several examples satisfy this assumption, among others the Daubechies wavelet or the mexican hat wavelet, see [5] or [7]).

Proposition 5.1. *Suppose that* $\Psi \in C^m(\mathbb{R})$ *with* $m > \frac{H\beta}{\alpha}$. *Assume (2.11) with* $\alpha \in \left(\frac{1}{2}, 1\right)$ *and let* $V_N(t_N, a_N), \widehat{V}_N(t_N, a_N)$ *be given by (2.10) and (5.2) respectively. Then*

$$\mathbf{E}\left| N^{\frac{1-\alpha}{2}}\left(V_N(t_N, a_N) - \widehat{V}_N(t_N, a_N)\right)\right| \to_{N\to\infty} 0.$$

Proof. We start by estimating the difference between the coefficient $d(t_N, a_N, i)$ and its discrete counterpart $e_N(t_N, a_N, i)$ with $i = 1, \ldots, N_{a_N}$ and with t_N, a_N as in (2.11). Let us compute the $L^2(\Omega)$-norm of this difference. We write

$$\mathbf{E}\left(d(t_n, a_N, i) - e_N(t_N, a_N, i)\right)^2 = \mathbf{E}d(t_N, a_N, i)^2 - 2\mathbf{E}d(t_N, a_N, i)e_N(t_N, a_N, i) + \mathbf{E}e_N(t_N, a_N, i)^2.$$

The first summand $\mathbf{E}d(t_N, a_N, i)^2$ has already been computed in (4.3). Let us compute the other two terms. For $N \geq 1$ and $i = 1, \ldots, N_{a_N}$ we have from the covariance formula (2.6)

$$\mathbf{E}e_N(t_N, a_N, i)^2 = \frac{1}{a_N} \sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{l}{a_N}\right) \mathbf{E}u(t_N, k + a_N i)u(t_N, l + a_N i) \tag{5.3}$$

$$= \frac{1}{a_N} \sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{k}{a_N}\right) \left[\frac{c_H}{2}|k - l|^{2H+1} - \frac{t_N}{4}|k - l|^{2H} + \frac{t_N^{2H+1}}{2(2H+1)}\right].$$

We used the fact that $|k - l| \leq a_N = N^\alpha < t_N = N^\beta$ under (2.11), so the last summand in (2.6) vanishes. We also have from (2.6), (2.9) and (5.1)

$$\mathbf{E}d(t_N, a_N, i)e_N(t_N, a_N, i) \tag{5.4}$$

$$= \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \int_{\mathbb{R}} \mathrm{d}x \Psi(x) \mathbf{E}u(t_N, k + a_N i)u(t_N, a_N(x + i))$$

$$= \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \int_{\mathbb{R}} \mathrm{d}x \Psi(x) \left[\frac{c_H}{2}|k - a_N x|^{2H+1} - \frac{t_N}{4}|k - a_N x|^{2H} + \frac{t_N^{2H+1}}{2(2H+1)}\right].$$

Via (4.3), (5.3) and (5.4),

$$\mathbf{E}\left(d(t_n, a_N, i) - e_N(t_N, a_N, i)\right)^2$$

$$= \frac{c_H}{2} a_N^{2H+2} \left(\frac{1}{a_N^2} \sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{l}{a_N}\right) \left|\frac{k-l}{a_N}\right|^{2H+1} + \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y|^{2H+1}\right.$$

$$\left. -2\frac{1}{a_N} \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \int_{\mathbb{R}} \mathrm{d}x \Psi(x) \left|k - \frac{a_N}{x}\right|^{2H+1}\right)$$

$$+ \frac{c_H}{2} a_N^{2H+1} \left(\sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{l}{a_N}\right) \left|\frac{k-l}{a_N}\right|^{2H} + \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y|^{2H}\right.$$

$$\left. -2\frac{1}{a_N} \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \int_{\mathbb{R}} \mathrm{d}x \Psi(x) \left|k - \frac{a_N}{x}\right|^{2H}\right)$$

$$+ \frac{t_N^{2H+1}}{2(2H+1)} \frac{1}{a_N} \sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{l}{a_N}\right). \tag{5.5}$$

Now we use the following bounds (we refer to [5] for their proofs, see also [7]) for N large

$$\left|\frac{1}{a_N^2} \sum_{k,l=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \Psi\left(\frac{l}{a_N}\right) \left|\frac{k-l}{a_N}\right|^{2H} - \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y|^{2H}\right| \leq C\frac{1}{a_N}, \tag{5.6}$$

and

$$\left|\frac{1}{a_N} \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right) \int_{\mathbb{R}} \mathrm{d}x \Psi(x) \left|k - \frac{a_N}{x}\right|^{2H} - \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{d}x\mathrm{d}y \Psi(x)\Psi(y)|x - y|^{2H}\right| \leq C\frac{1}{a_N} \tag{5.7}$$

and, for Ψ of class $C^m(\mathbb{R})$,

$$\left|\frac{1}{a_N} \sum_{k=0}^{[a_N]} \Psi\left(\frac{k}{a_N}\right)\right| \leq C\frac{1}{a_N^m}. \tag{5.8}$$

with $C > 0$ not depending on N. By using the inequalities (5.6), (5.7) and (5.8) in (5.5), we obtain

$$
\begin{aligned}
\mathbf{E} \left(d(t_n, a_N, i) - e_N(t_N, a_N, i) \right)^2 &\leq C \left[a_N^{2H+1} + t_N a_N^{2H} + t_N^{2H+1} a_N^{1-2m} \right] \\
&\leq C \left[N^{(2H+1)\alpha} + N^{\beta + 2H\alpha} + N^{(2H+1)\beta + \alpha(1-2m)} \right].
\end{aligned}
\tag{5.9}
$$

For the renormalized coefficients, we have the estimate

$$
\begin{aligned}
\mathbf{E} \left(\tilde{d}(t_n, a_N, i) - \tilde{e}_N(t_N, a_N, i) \right)^2 &= D(t_N, a_N)^{-1} \mathbf{E} \left(d(t_n, a_N, i) - e_N(t_N, a_N, i) \right)^2 \\
&\leq C N^{-\beta + (2H+1)\alpha} \left[N^{(2H+1)\alpha} + N^{\beta + 2H\alpha} + N^{(2H+1)\beta + \alpha(1-2m)} \right] \\
&= C \left[N^{-\beta} + N^{-\alpha} + N^{2H\beta - \alpha(2m+2H)} \right].
\end{aligned}
$$

If $m > \frac{H\beta}{\alpha}$, then for all $i = 1, \ldots, N_{a_N}$,

$$
\mathbf{E} \left(\tilde{d}(t_n, a_N, i) - \tilde{e}_N(t_N, a_N, i) \right)^2 \leq C N^{-\alpha}.
\tag{5.10}
$$

Finally, we regard the $L^1(\Omega)$-norm of the difference $V_N(t_N, a_N) - \widehat{V}_N(t_N, a_N)$. By using Cauchy-Scwarz inequality as proceeding as in the the proof of Lemma 1 in [7], we can write, with $C_1, C_2 > 0$,

$$
\begin{aligned}
&\mathbf{E} \left| V_N(t_N, a_N) - \widehat{V}_N(t_N, a_N) \right| \\
&\leq C_1 \left(\frac{1}{N_{a_N}} \sum_{i=1}^{N_{a_N}} \mathbf{E} \left(\tilde{d}(t_N, a_N, i) - \tilde{e}_N(t_N, a_N, i) \right)^2 \right)^{\frac{1}{2}} \\
&\quad \times \left(C_2 + 2 \frac{1}{N_{a_N}} \sum_{i=1}^{N_{a_N}} \mathbf{E} \left(\tilde{d}(t_N, a_N, i) - \tilde{e}_N(t_N, a_N, i) \right)^2 \right)^{\frac{1}{2}}
\end{aligned}
$$

and by (5.10),

$$
\mathbf{E} \left| V_N(t_N, a_N) - \widehat{V}_N(t_N, a_N) \right| \leq C N^{-\frac{\alpha}{2}} \left(C + \frac{1}{N^\alpha} \right)^{\frac{1}{2}} \leq C N^{-\frac{\alpha}{2}}.
$$

Consequently,

$$
\mathbf{E} \left| N^{\frac{1-\alpha}{2}} \left(V_N(t_N, a_N) - \widehat{V}_N(t_N, a_N) \right) \right| \leq C N^{\frac{1-\alpha}{2}} N^{-\frac{\alpha}{2}} = C N^{\frac{1}{2} - \alpha}
$$

and the conclusion is obtained since $\alpha > \frac{1}{2}$. $\square$

As a consequence of the above result, the discrete wavelet variation $\widehat{V}_N(t_N, a_N)$ has the same limit in law as $V_N(t_N, a_N)$.

Corollary 5.2. *Let the assumptions in Proposition 5.1 prevail and let $\widehat{V}_N$ be given by (5.2). Then the d-dimensional random vector $\left(N^{\frac{1-\alpha}{2}} \widehat{V}_N(t_N, L a_N) \right)_{L=1,\ldots,d}$ converges in distribution, as $N \to \infty$, to a centered*

d-dimensional Gaussian vector with covariance matrix $\Gamma = (\Gamma_{L_1,L_2})_{L_1,L_2=1,\ldots,d}$ *where the matrix* Γ *is given by* (3.5).

Proof. The proof immediately follows from Theorem 3.3 and Proposition 5.1. □

5.3. The definition of the estimator

Let us denote, for $t > 0, a > 0$

$$S_N(t,a) := \frac{1}{N_a} \sum_{i=1}^{N_a} d(t,a,i)^2. \tag{5.11}$$

Notice that the sequence $S_N(t,a)$ is related to the wavelet variation $V_N(t,a)$ in (2.10) as follows

$$S_N(t,a) = D(t,a)\left(V_N(t,a) + 1\right) \tag{5.12}$$

with $D(t,a)$ defined by (4.1). Let $d \geq 1$ and assume (2.11). By taking the expectation in (5.11) (note that $\mathbf{E}V_N(t_N,a_N) = 0$), we have, for every $L = 1,\ldots,d$, by using Lemma 4.1,

$$N^{-\beta-(2H+1)\alpha}\mathbf{E}S_N(t_N,La_N) = N^{-\beta-(2H+1)\alpha}D(t_N,La_N) \underset{N\to+\infty}{\to} \frac{1}{4}K_{\Psi,H}L^{2H+1} > 0$$

with $K_{\Psi,H}$ from Lemma 4.1. We write the above relation for $t_N = t_{La_N}$ (which means that we replace $t_N = N^\beta$ by $L^\beta N^{\alpha\beta}$ in (4.3). To do this, we will assume that in the sequel $\alpha\beta > 1$ and with this assumption all our theoretical results (such as Thm. 3.3) can be applied. So

$$N^{-\alpha\beta-(2H+1)\alpha}\mathbf{E}S_N(t_{La_N},La_N) = N^{-\alpha\beta-(2H+1)\alpha}D(t_{La_N},La_N) \underset{N\to+\infty}{\to} \frac{1}{4}K_{\Psi,H}L^{\beta+2H+1} > 0. \tag{5.13}$$

The above relation (5.13) implies, for $N \geq 1$,

$$\log\mathbf{E}(S_N(t_{La_N},La_N)) = \log(D(t_{La_N},La_N)$$
$$= (\beta+2H+1)\alpha\log N + (\beta+2H+1)\log L + \log\left(\frac{1}{4}K_{\Psi,H}\right) + \log(1+\varepsilon_N)$$
$$= (\beta+2H+1)(\alpha\log N + \log L) + \log\left(\frac{1}{4}K_{\Psi,H}\right) + \log(1+\varepsilon_N)$$
$$= (\beta+2H+1)\log(La_N) + \log\left(\frac{1}{4}K_{\Psi,H}\right) + \log(1+\varepsilon_N) \tag{5.14}$$

where $(\varepsilon_N)_{N\geq 1}$ is a a deterministic sequence defined by, for every $N \geq 1$,

$$\varepsilon_N = \left(\frac{1}{4}K_{\Psi,H}L^{2H+1}\right)^{-1} N^{-\beta-(2H+1)\alpha}\mathbf{E}S_N(t_N,La_N) - 1$$
$$= \left(\frac{1}{4}K_{\Psi,H}L^{2H+1}\right)^{-1} N^{-\alpha\beta-(2H+1)\alpha}D(t_{La_N},La_N) - 1 \to_{N\to\infty} 0.$$

Let us also introduce the discretized counterpart of $S_N(t, a)$, *i.e.*

$$\widehat{S}_N(t, a) := \frac{1}{N_a} \sum_{i=1}^{N_a} e_N(t, a, i)^2 = D(t_{a_N}, a_N)(1 + \widehat{V}_N(t_{a_N}, a_N)) \tag{5.15}$$

with e_N from (5.1).

The above relation (5.14) inspires the following definition of the estimator for the Hurst parameter

$$2\widehat{H}_N + \beta + 1 = (X^T X)^{-1} X^T Y \tag{5.16}$$

where ^{T} denotes the transpose and with the notation

$$X = \log(La_N)_{L=1,\dots,d} = (\alpha \log N + logL)_{L=1,\dots,d} \text{ and } Y = \left(\log \widehat{S}_N(t_{La_N}, La_N)\right)_{L=1,\dots,d}. \tag{5.17}$$

Equivalently, we have

$$\widehat{H}_N = \frac{1}{2}(X^T X)^{-1} X^T Y - \frac{1}{2} = \frac{1}{2} \sum_{L=1}^{d} \frac{\log \widehat{S}_N(t_{La_N}, La_N) \log(La_N)}{\sum_{L=1}^{d} (\log(La_N))^2} - \frac{1}{2} - \frac{\beta}{2} \tag{5.18}$$

with X, Y from (5.17).

Remark 5.3. 1. The following heuristics leads to the expression (5.16):
- We approximate $\mathbf{E}(S_N(t_{La_N}, La_N)$ by $S_N(t_{La_N}, La_N)$ and then $\log S_N(t_{La_N}, a_N)$ by $\log \widehat{S}_N(t_{La_N}, a_N)$.
- In the expression

$$\log(\widehat{S}_N(t_{La_N}, La_N)) \sim (\beta + 2H + 1) \log(La_N) + \log\left(\frac{1}{4} K_{\Psi, H}\right) + \log(1 + \varepsilon_N)$$

we use a log–log regression of $(\log(S_N(t_{La_N}, La_N))_{L=1,\dots,d}$ on $(\log La_N)_{L=1,\dots,d} = (\alpha + \log L)_{L=1,\dots,d}$ *i.e.* we minimize with respect to $H \in \left(\frac{1}{2}, 1\right)$ the function

$$f(H) = \sum_{L=1}^{d} (\log S_N(t_{La_N}, La_N) - (2H + 1 + \beta) \log(La_N))^2.$$

2. Notice that the estimator (5.18) is expressed in terms of the sequence (5.15) which depends on the discrete wavelet coefficients e_N. Therefore, the estimator can be computed from the data, that is, from the observations $u(t, k), k = 1, 2, \dots, N$ with $t = t_{a_N} = N^{\alpha\beta}$ with $\alpha\beta > 1$. So, if we have at our disposal N observations in space, one needs to be able to observe them at time $N^{\alpha\beta}$. Recall that in practice our wave equation describes the vertical displacement of a vibrating string under a random force. This means that the observation time of the vibrating time should be sufficiently long and it is related to the number of spatial observations.

Using Theorem 3.3, we can deduce the limit behavior of the estimator $\widehat{H}_N$.

Theorem 5.4. *Consider the estimator $\widehat{H}_N$ given by (5.18). Let the assumptions in Proposition 5.1 prevail. Assume also $\alpha\beta > 1$. Then the estimator (5.18) is stronly consistent and*

$$2\alpha(\log N)N^{\frac{1-\alpha}{2}}\left(\widehat{H}_N - H\right) \to^{(d)} N\left(0, \Gamma(1, \dots, 1)^T\right) \tag{5.19}$$

with the matrix Γ *defined by* (3.5).

Proof. From (5.15) and (5.14), we have for large enough N,

$$
\begin{aligned}
\log \widehat{S}_N(t_{La_N}, La_N) &= \log D(t_{La_N}, La_N) + \log\left(1 + \widehat{V}_N(t_{La_N}, La_N)\right) \\
&= (\beta + 2H + 1)\log(La_N) \\
&\quad + \log\left(\frac{1}{4}K_{\Psi,H}\right) + \log\left(1 + \widehat{V}_N(t_{La_N}, La_N)\right) + \log(1 + \varepsilon_N)
\end{aligned}
$$

with ε_N defined in Section 5.3. By plugging the above relation into (5.18), we obtain, for N large,

$$
\begin{aligned}
\beta + 2\widehat{H}_N + 1 &= \beta + 2H + 1 + \frac{\log\left(\frac{1}{4}K_{\Psi,H}\right)\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2} + \log(1 + \widehat{V}_N(t_{La_N}, La_N))\frac{\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2} \\
&\quad + \log(1 + \varepsilon_N)\frac{\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2}
\end{aligned}
$$

and so

$$
\begin{aligned}
2(\widehat{H}_N - H) &= \frac{\log\left(\frac{1}{4}K_{\Psi,H}\right)\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2} + \log(1 + \widehat{V}_N(t_{La_N}, La_N))\frac{\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2} \\
&\quad + \log(1 + \varepsilon_N)\frac{\sum_{L=1}^{d}\log(La_N)}{\sum_{L=1}^{d}(\log(La_N))^2}.
\end{aligned}
$$

Note that $\widehat{V}_N(t_{a_N}, a_N)$ converges to zero almost surely as $N \to \infty$, this is a consequence of Proposition 3.1 and of a standard Borel-Cantelli argument, see *e.g.* [24]. Therefore, as ε_N tends to zero we get that $\widehat{H}_N \to_{N\to\infty} H$ almost surely and by using Theorem 3.3 we obtain the convergence (5.19). □

5.4. Estimation when the time is fixed

Assume now that the time t is fixed, as in (2.12). We would like to estimate the parameter H of the mild solution (2.4) based on the observation of the solution at a fied time and at discrete points in space. The result in Proposition 4.4 shows that the behavior of the wavelet coefficient is not a power-function with exponent depending on H and their relationship is more complex. Actually

$$
\mathbf{E}d(t, a_N, i)^2 = \frac{1}{N_{a_N}}\left[K_{1,t}(H) + K_{2,t}(H)N^{-\alpha}\right] \to_{N\to\infty} K_{1,t}(H) \tag{5.20}
$$

with $K_{1,t}(H), K_{2,t}(H)$ from Proposition 4.4. Therefore the log–log regression argument employed above cannot work when the time is fixed. We proposed and alternative method *via* the analysis of the constant $K_{1,t}(H)$.

Consider the sequence S_N given by (5.11) and assume now (2.12). By Proposition 4.4,

$$
\mathbf{E}S_N(t, a_N) = D(t, a_N) = K_{1,t}(H) + K_{2,t}(H)N^{-\alpha} \to_{N\to\infty} K_{1,t}(H) \tag{5.21}
$$

with $K_{1,t}(H) = \frac{1}{2(H+1)}t^{2H+2}$, see (4.10). By approximating, as usual, $\mathbf{E}S_N(t, a_N)$ by $S_N(t, a_N)$, we can say that for N large enough, $S_N(t, a_N)$ is close to $K_{1,t}(H)$.

Lemma 5.5. *For some $t > 0$ sufficiently large (not depending on N), the equation*

$$S_N(t, a_N) - K_{1,t}(x) = 0 \tag{5.22}$$

has a unique solution in the interval $\left[\frac{1}{2}, 1\right]$.

Proof. Consider the function

$$f_{N,t}(H) = S_N(t, a_N) - K_{1,t}(H)$$

with $H \in \left[\frac{1}{2}, 1\right]$. We have

$$f'_{N,t}(x) = -\frac{\partial}{\partial H} K_{1,t}(H) = -f_1(H) t^{2H+2} \log t + f_2(H) t^{2H+2}$$

with

$$f_1(H) = \frac{1}{2(H+1)} > 0$$

and $f_2(H) = \frac{1}{2(H+1)^2}$ for $H \in \left[\frac{1}{2}, 1\right]$. When $t \to \infty$, this derivative behaves as $f_1(H) t^{2H+2} \log t$ so it is positive by choosing a suitable time t large enough. Consequently, the function $f_{N,t}$ is invertible on $\left[\frac{1}{2}, 1\right]$ and the conclusion follows. □

We will assume in the sequel that t is large enough in order to ensures the existence and uniqueness of the solution to (5.22).

Definition 5.6. We define $\widehat{H}_N$ to be the unique solution of the equation (5.22).

We derive the asymptotic properties of the estimator $\widehat{H}_N$.

Proposition 5.7. *The estimator $\widehat{H}_N$ from Definition 5.6 is strongly consistent. Moreover, it satisfies the following limit behavior in distribution*

$$N^{\frac{1-\alpha}{2}} (\widehat{H}_N - H) \to_{N\to\infty}^{(d)} N\left(0, 2K_{1,t}(H)^2 \left(\frac{\partial}{\partial H} K_{1,t}^{-1}(H)\right)^2\right). \tag{5.23}$$

Proof. By Lemma 5.5, for some $t > 0$

$$S_N(t, a_N) = K_{1,t}(\widehat{H}_N)$$

and from (5.12),

$$S_N(t, a_N) = D(t, a_N)(1 + V_N(t, a_N)).$$

Thus

$$K_{1,t}(\widehat{H}_N) = D(t, a_N)(1 + V_N(t, a_N)).$$

We let $N \to \infty$ above. Since $V_N(t, a_N)$ tends to zero almost surely and $D(t, a_N)$ converges to $K_{1,t}(H)$ as $N \to \infty$, we get

$$\lim_{N \to \infty} K_{1,t}(\widehat{H}_N) = K_{1,t}(H) \text{ almost surely.} \tag{5.24}$$

By the proof of Lemma 5.5, we deduce that $K_{1,H}$ is invertible on $\left[\frac{1}{2}, 1\right]$ and its inverse is continuously differentiable on this interval. By applying $K_{1,t}^{-1}$ to (5.24) we deduce that $\widehat{H}_N \to_{N \to \infty} H$ almost surely.

Let us show that the estimator is asymptotically normal. Indeed, from (5.12)

$$\frac{1}{D(t, a_N)} \left(S_N(t, a_N) - D(t, a_N)\right) = V_N(t, a_N)$$

and consequently, in distribution, by Theorem 3.5

$$\frac{1}{\sqrt{2}} N^{\frac{1-\alpha}{2}} \frac{1}{D(t, a_N)} \left(S_N(t, a_N) - D(t, a_N)\right) \to_{N \to \infty} N(0, 1).$$

Given the asymptotic behavior of $D(t, a_N)$ (see 5.21)), we can write

$$N^{\frac{1-\alpha}{2}} \left(K_{1,t}(\widehat{H}_N) - K_{1,t}(H)\right) \to_{N \to \infty}^{(d)} N(0, 2K_{1,t}(H)^2).$$

By using the delta-method with the continuously differentiable function $K_{1,t}^{-1}$ on $\left[\frac{1}{2}, 1\right]$, we obtain (5.23). □

Let us end this statistical inference part with some comments:

Remark 5.8. 1. There exist alternative ways to estimate the Hurst parameter for t fixed by exploiting the relation (5.20). But using this result for different fixed times $t_1, \ldots, t_d > 0$, and by applying a nonlinear least-squares regression of $S_N(t_i, a_N)$ on t_i with $i = 1, \ldots, d$ we can obtain another estimator for H (which should be, in principle, consistent and asymptotically normal). Being related to nonlinear regression, it analysis could be more complex but probabably we can avoid the restriction of t large, assumed in Proposition 5.7.
2. The estimator from Definition 5.6 is based on the variation $S_N(t, a, i)$ which is written in terms of the continuous wavelet transforms $d(t, a, i)$ and the of the inverse function $K_{1,t}^{-1}$. While the inverse can be (at least numerically) computed, the wavelet coefficients are not directly computed from the observations $u(t, k), k = 1, \ldots, N$. An approach to compute them is use Riemann sums approximations, as in (5.1). Another possibility is to use a pyramidal multiresolution algorithm (see for example the survey [6]).
3. Other methods (not based on wavelets) to estimate the Hurst parameter based on observations of (2.4) at fixed time are obtained *via* the generalized spatial variations of the solution (see *e.g.* [12] or [22]).

Appendix A.

The basic tools from the analysis on Wiener space are presented in this section. We will focus on some elementary facts about multiple stochastic integrals. We refer to [17] for a complete review on the topic.

Consider $\mathcal{H}$ a real separable infinite-dimensional Hilbert space with its associated inner product $\langle ., . \rangle_{\mathcal{H}}$, and $(B(\varphi), \varphi \in \mathcal{H})$ an isonormal Gaussian process on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, which is a centered Gaussian family of random variables such that $\mathbf{E}(B(\varphi)B(\psi)) = \langle \varphi, \psi \rangle_{\mathcal{H}}$ for every $\varphi, \psi \in \mathcal{H}$. Denote by I_q the qth multiple stochastic integral with respect to B, which is an isometry between the Hilbert space $\mathcal{H}^{\odot q}$ (symmetric tensor product) equipped with the scaled norm $\frac{1}{\sqrt{q!}} \| \cdot \|_{\mathcal{H}^{\otimes q}}$ and the Wiener chaos of order q, which is defined as the

closed linear span of the random variables $H_q(B(\varphi))$ where $\varphi \in \mathcal{H}$, $\|\varphi\|_{\mathcal{H}} = 1$ and H_q is the Hermite polynomial of degree $q \geq 1$ defined by:

$$H_q(x) = (-1)^q \exp\left(\frac{x^2}{2}\right) \frac{\mathrm{d}^q}{\mathrm{d}x^q}\left(\exp\left(-\frac{x^2}{2}\right)\right), \ x \in \mathbb{R}. \tag{A.1}$$

The isometry of multiple integrals can be written as follows: for p, $q \geq 1$, $f \in \mathcal{H}^{\otimes p}$ and $g \in \mathcal{H}^{\otimes q}$

$$\mathbf{E}\left(I_p(f)I_q(g)\right) = \begin{cases} q!\langle \tilde{f}, \tilde{g}\rangle_{\mathcal{H}^{\otimes q}} & \text{if } p = q, \\ 0 & \text{otherwise.} \end{cases} \tag{A.2}$$

We have the following product formula: if $f \in \mathcal{H}^{\odot p}$ and $g \in \mathcal{H}^{\odot q}$, then

$$I_p(f)I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r}\binom{q}{r} I_{p+q-2r}\left(f \otimes_r g\right) \tag{A.3}$$

where $f \otimes_r g$ denotes the contraction of order $r = 0, 1, \ldots, p \wedge q$.

We denote by D the Malliavin derivative operator that acts on cylindrical random variables of the form $F = g(B(\varphi_1), \ldots, B(\varphi_n))$, where $n \geq 1$, $g : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with compact support and $\varphi_i \in \mathcal{H}$ in the following way

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(B(\varphi_1), \ldots, B(\varphi_n))\varphi_i.$$

The operator D is closable and it can be extended to the closure of the set of cylindrical random variables (denotes $\mathbb{D}^{1,2}$) with respect to the norm

$$\|F\|_{1,2}^2 := \mathbf{E}|F|^2 + \mathbf{E}\|DF\|_{\mathcal{H}}^2.$$

If $F = I_p(f)$ with $f \in \mathcal{H}^{\odot p}$ and $p \geq 1$, then

$$DF = pI_{p-1}\left(f(\cdot, *)\right) \tag{A.4}$$

where "$*$" stands for $p - 1$ variables.

References

[1] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch, Self-similarity and long range dependence through the wavelet lens. Long-range dependence: Theory and Applications, edited by P. Doukhan, G. Oppenheim, M.S. Taqqu. Birhäuser (2003).

[2] R. Balan and C.A. Tudor , The stochastic heat equation with fractional-colored noise: existence of the solution. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** (2008) 57–87.

[3] R. Balan and C.A. Tudor, Stochastic heat equation with multiplicative fractional-colored noise. *J. Theor. Probab.* **23** (2010) 834–870.

[4] R.M. Balan and C.A. Tudor, The stochastic wave equation with fractional colored noise: a random field approach. *Stoch. Process. Appl.* **120** (2010) 2468–2494.

[5] J.-M. Bardet, Statistical study of the wavelet analysis of fractional Brownian motion. *IEEE Trans. Inform. Theory* **48** (2002) 991–999.

[6] J.-M. Bardet, G. Lang, G. Oppenheim, Georges, A. Philippe, S. Stoev and M.S. Taqqu, Murad, Semi-parametric estimation of the long-range dependence parameter: a survey. Theory and applications of long-range dependence. Birkhäuser Boston, Boston, MA (2003) 557–577.

[7] J.-M. Bardet and C.A. Tudor, A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter. *Stoch. Process. Appl.* **120** (2010) 2331–2362.

[8] I. Cialenco, Statistical inference for SPDEs: an overview. *Stat. Inference Stoch. Process.* **21** (2018) 309–329.

[9] J. Clarke De la Cerda, C.A. Tudor, Hitting probabilities for the stochastic wave equation with fractional colored noise. *Rev. Mat. Iberoam.* **30** (2014) 685–709.

[10] R.C. Dalang and M. Sanz-Solé, Hölder-Sobolev regularity of the solution to the stochastic wave equation in dimension three. *Memoirs Am. Math. Soc.* **199** (2009) 1–70.

[11] P. Flandrin, Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory* **38** (1992) 910–917.

[12] M. Khalil and C.A. Tudor, Correlation structure, quadratic variations and parameter estimation for the solution to the wave equation with fractional noise. *Electr. J. Stat.* **12** (2018) 3639–3672.

[13] M. Kleptsyna and A. Le Breton, Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat. Inference Stoch. Processes* **5** (2002) 229–241.

[14] Yu.A. Kutoyants, Statistical Inference for Ergodic Diffusion Processes. *Springer Series in Statistics.* Springer (2004).

[15] Y. Meyer, F. Sellan and M.S. Taqqu, Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *J. Fourier Anal. Appl.* **5** (2008) 465–494.

[16] I. Nourdin and G. Peccati, Normal Approximations with Malliavin Calculus From Stein 's Method to Universality. Cambridge University Press (2012).

[17] D. Nualart, Malliavin Calculus and Related Topics, 2nd ed. Springer (2006).

[18] D. Nualart and G. Peccati, Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (2005) 177–193.

[19] G. Peccati and C.A. Tudor, Gaussian limits for vector-valued multiple stochastic integrals. *Séminaire de Probabilités* **XXXIV** (2004) 247–262.

[20] B.L.S. Praksa Rao, Semimartingales and their Statistical Inference. Chapman and Hall (1999).

[21] B.L.S. Prakasa Rao, Statistical inference for fractional diffusion processes. *Wiley Series in Probability and Statistics.* Chichester, John Wiley & Sons (2010).

[22] R. Shevchenko, M. Slaoui and C.A. Tudor, Generalized k-variations and Hurst parameter estimation for the fractional wave equation *via* Malliavin calculus. *J. Statist. Plann. Inference* **207** (2020) 155–180.

[23] S. Torres, C.A. Tudor and F. Viens, Quadratic variations for the fractional-colored stochastic heat equation. *Electr. J. Probab.* **19** (2014) 76.

[24] C.A. Tudor, Analysis of variations for self-similar processes. A stochastic calculus approach. *Probability and its Applications* (New York). Springer, Cham (2013).

[25] C.A. Tudor and F.G. Viens, Statistical aspects of the fractional stochastic calculus. *Ann. Statist.* **35** (2007) 1183–1212.