Estimation of the linear fractional stable motion

STEPAN MAZUR[1], DMITRY OTRYAKHIN[2,*] and MARK PODOLSKIJ[2,**]

[1]*Unit of Statistics, Örebro University School of Business, Fakultetsgatan 1, 70281 Örebro, Sweden.*
E-mail: stepan.mazur@oru.se
[2]*Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus, Denmark.*
E-mail: *d.otryakhin@math.au.dk; **mpodolskij@math.au.dk

In this paper, we investigate the parametric inference for the linear fractional stable motion in high and low frequency setting. The symmetric linear fractional stable motion is a three-parameter family, which constitutes a natural non-Gaussian analogue of the scaled fractional Brownian motion. It is fully characterised by the scaling parameter $\sigma > 0$, the self-similarity parameter $H \in (0, 1)$ and the stability index $\alpha \in (0, 2)$ of the driving stable motion. The parametric estimation of the model is inspired by the limit theory for stationary increments Lévy moving average processes that has been recently studied in (*Ann. Probab.* **45** (2017) 4477–4528). More specifically, we combine (negative) power variation statistics and empirical characteristic functions to obtain consistent estimates of (σ, α, H). We present the law of large numbers and some fully feasible weak limit theorems.

Keywords: fractional processes; limit theorems; parametric estimation; stable motion

1. Introduction

Since the pioneering work by Mandelbrot and van Ness [18] fractional Brownian motion (fBm) became one of the most prominent Gaussian processes in the probabilistic and statistical literature. As a building block in stochastic models, it found various applications in natural and social sciences such as physics, biology or economics. Mathematically speaking, the scaled fBm is fully characterised by its scaling parameter $\sigma > 0$ and Hurst parameter $H \in (0, 1)$. More specifically, the scaled fBm $Z_t = \sigma B_t^H$ is a zero mean Gaussian process with covariance kernel determined by

$$\mathbb{E}\big[B_t^H B_s^H\big] = \frac{1}{2}\big(t^{2H} + s^{2H} - |t - s|^{2H}\big), \quad t, s \geq 0.$$

We recall that the (scaled) fBm with Hurst parameter $H \in (0, 1)$ is the unique Gaussian process with stationary increments and self-similarity index H, that is, it holds that $(a^H Z_t)_{t \geq 0} = (Z_{at})_{t \geq 0}$ in distribution for any $a > 0$. Over the last forty years there has been a lot of progress in limit theorems and statistical inference for fBm's. The estimation of the Hurst parameter H and/or the scaling parameter σ has been investigated in numerous papers both in low and high frequency framework. We refer to [13] for efficient estimation of the Hurst parameter H in the low frequency setting and to [9,12,16] for the estimation of (σ, H) in the high frequency setting, among many others. In the low frequency framework, the spectral density methods are usually

applied and the optimal convergence rate for the estimation of (σ, H) is known to be $\sqrt{n}$. In the high frequency setting the estimation of the pair (σ, H) typically relies upon power variations and related statistics, and the optimal convergence rate is known to be $(\sqrt{n}/\log(n), \sqrt{n})$. More recently, the class of multifractional Brownian motions, which accounts for time varying Hurst parameter, has been introduced in the literature (see, e.g., [2,19,30]). We refer to the work [4,17] for estimation techniques for the regularity of a multifractional Brownian motion.

If we drop the Gaussianity assumption, the class of stationary increments self-similar processes becomes much larger. This is a consequence of the work by Pipiras and Taqqu [20], which in turn applies the decomposition results from the seminal paper by Rosiński [25] (see also [26]). The crucial theorem proved in [25] shows that each stationary stable process can be uniquely decomposed (in distribution) into three independent parts: the mixed moving average process, the harmonizable process and the "third kind" process described by a conservative nonsingular flow. The most prominent example of a non-Gaussian stationary increments self-similar process is the linear fractional stable motion (an element of the first class), which has been introduced in [11]. It is defined as follows: On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}}, \mathbb{P})$, we introduce the process

$$X_t = \int_{\mathbb{R}} \left\{ (t-s)_+^{H-1/\alpha} - (-s)_+^{H-1/\alpha} \right\} dL_s, \quad x_+ := \max\{x, 0\}, \tag{1.1}$$

where L is a symmetric α-stable Lévy motion, $\alpha \in (0, 2)$, with scale parameter $\sigma > 0$ and $H \in (0, 1)$ (here we use the convention $x_+^a = 0$ for any $x \leq 0$ and $a \in \mathbb{R}$). In some sense the linear fractional stable motion is a non-Gaussian analogue of fBm. The process $(X_t)_{t \in \mathbb{R}}$ has symmetric α-stable marginals, stationary increments and it is self-similar with parameter H. Fractional stable motions are often used in natural sciences, e.g. in physics or internet traffic, where the process under consideration exhibits stationarity and self-similarity along with heavy tailed marginals (see, e.g., [15] for the context of turbulence modelling). The probabilistic properties of linear fractional stable motions, such as integration concepts, path and variational properties, have been intensively studied in several papers, see, for example, [6–8] among many others. However, from the statistical point of view, very little is known about the inference for the parameter $\theta = (\sigma, \alpha, H) \in \mathbb{R}_+ \times (0, 2) \times (0, 1)$ in high or low frequency setting. The few existing papers mostly concentrate on estimation of the self-similarity parameter H. The work [3,22] investigates the asymptotic theory for a wavelet-based estimator of H when $\alpha \in (1, 2)$. In [5,28] the authors suggest to use power variation statistics to obtain an estimator of H, but this method also requires the a priori knowledge of the lower bound for the stability parameter α. Recently, the work [14] suggested to use negative power variations to get a consistent estimator of H, which applies for any $\alpha \in (0, 2)$, but this article does not contain a central limit theorem for this method. Finally, in [5,15] the authors propose to use an empirical scale function to estimate the pair (α, H). However, this approach only provides a $\log(n)$-consistent estimator without any hope for a central limit theorem.

In this paper, we will propose a new estimation procedure for the parameter $\theta = (\sigma, \alpha, H)$ in high and low frequency framework. Our methodology is based upon the use of power variation statistics, with possibly negative powers, and the empirical characteristic function. The probabilistic techniques originate from the recent article [5], which has developed the asymptotic

theory for power variations of higher order differences of stationary increments Lévy moving averages (see also [21,22] for related asymptotic theory). However, we will prove a much stronger result, which in particular includes the complete asymptotic theory for negative power variations and empirical characteristic functions consisting of normal and stable limits (see Theorem 2.2). This novel result is key to obtaining a distributional theory for the estimator of the parameter $\theta \in \mathbb{R}_+ \times (0, 2) \times (0, 1)$. We will deduce a fully feasible asymptotic theory for our estimator with convergence rates $(\sqrt{n}, \sqrt{n}, \sqrt{n})$ in the low frequency setting and $(\sqrt{n}/\log(n), \sqrt{n}/\log(n), \sqrt{n})$ in the high frequency setting. We will distinguish two different scenarios $H - 1/\alpha > 0$ and the general case, where the first one corresponds to the continuous framework. In the continuous setting our estimation method is simpler than in the general case, which requires a two-step procedure. Finally, we present an extensive simulation study to uncover the finite sample properties of our estimators. We conclude that the estimators of the self-similarity parameter H are quite reliable in all considered scenarios, while estimators for the scaling parameter σ have the worst finite sample performance.

The paper is structured as follows. Section 2 presents the basic properties of the linear fractional stable motion, the review of the probabilistic results from [5] and a multivariate limit theorem, which plays a key role for the statistical estimation. Section 3 is devoted to the statistical inference in the continuous case $H - 1/\alpha > 0$. The general case is treated in Section 4. Finally, Section 5 demonstrates some simulation results. All proofs are collected in the supplementary material [27].

2. First properties and some asymptotic results

2.1. Distributional and path properties

In this section, we review some basic properties of the linear fractional stable motion. First of all, we recall that the symmetric α-stable process $(L_t)_{t \in \mathbb{R}}$ with scale parameter $\sigma > 0$ is uniquely determined by the characteristic function of L_1, which is given by

$$\mathbb{E}\big[\exp(it L_1)\big] = \exp\big(-\sigma^\alpha |t|^\alpha\big), \quad t \in \mathbb{R}. \tag{2.1}$$

Following the theory of integration with respect to infinitely divisible processes investigated in [23], we know that for any deterministic function $g : \mathbb{R} \to \mathbb{R}$

$$X = \int_{\mathbb{R}} g_s \, dL_s < \infty \quad \text{almost surely} \quad \Leftrightarrow \quad \|g\|_\alpha^\alpha := \int_{\mathbb{R}} |g_s|^\alpha \, ds < \infty.$$

Furthermore, if $\|g\|_\alpha < \infty$ then X has a symmetric α-stable distribution with scale parameter $\sigma \|g\|_\alpha$. In particular, setting

$$X_t = \int_{\mathbb{R}} g_t(s) \, dL_s, \quad g_t(s) := \big\{(t - s)_+^{H - 1/\alpha} - (-s)_+^{H - 1/\alpha}\big\}, \tag{2.2}$$

we see that $g_t \in L^\alpha(\mathbb{R})$ for any $t \in \mathbb{R}$, since $|g_t(s)| \leq C_t |s|^{H - 1 - 1/\alpha}$ when $s \to -\infty$ and $H \in (0, 1)$. Hence, X_t is well defined for any $t \in \mathbb{R}$ and all finite dimensional distributions of the

linear fractional stable motion $(X_t)_{t \in \mathbb{R}}$ are symmetric α-stable. It is easily seen that the linear fractional stable motion has stationary increments.

We recall that symmetric α-stable random variables with $\alpha \in (0, 2)$ do not exhibit finite second moments, and hence their dependence structure can't be measured via the classical covariance kernel. Instead it is often useful to consider the following measure of dependence. Let $X = \int_{\mathbb{R}} g_s \, dL_s$ and $Y = \int_{\mathbb{R}} h_s \, dL_s$ with $\|g\|_\alpha, \|h\|_\alpha < \infty$. Then we introduce the measure of dependence $U_{g,h} : \mathbb{R}^2 \to \mathbb{R}$ via

$$U_{g,h}(u, v) := \mathbb{E}\big[\exp\big(i(uX + vY)\big)\big] - \mathbb{E}\big[\exp(iuX)\big]\mathbb{E}\big[\exp(ivY)\big]$$
$$= \exp\big(-\sigma^\alpha \|ug + vh\|_\alpha^\alpha\big) - \exp\big(-\sigma^\alpha \big(\|ug\|_\alpha^\alpha + \|vh\|_\alpha^\alpha\big)\big). \tag{2.3}$$

The quantity $U_{g,h}$ is extremely useful when computing covariances $\mathrm{cov}(K_1(X), K_2(Y))$ for functions $K_1, K_2 \in L^1(\mathbb{R})$; see for instance [22]. Let $\mathfrak{F}$ denote the Fourier transform and let $\mathfrak{F}^{-1}$ be its inverse. Furthermore, let $p_{(X,Y)}$, p_X and p_Y denote the density of (X, Y), X and Y, respectively. We recall that these densities are not available in a closed form except in some special cases. Using the duality relationship we obtain the identity

$$\mathrm{cov}\big(K_1(X), K_2(Y)\big) = \int_{\mathbb{R}^2} K_1(x) K_2(y)\big(p_{(X,Y)}(x, y) - p_X(x) p_Y(y)\big) \, dx \, dy$$
$$= \int_{\mathbb{R}^2} K_1(x) K_2(y) \mathfrak{F}^{-1} U_{g,h}(x, y) \, dx \, dy$$
$$= \int_{\mathbb{R}^2} \big(\mathfrak{F}^{-1} K_1(x)\big)\big(\mathfrak{F}^{-1} K_2(y)\big) U_{g,h}(x, y) \, dx \, dy. \tag{2.4}$$

We remark that the latter provides an explicit formula for computation of covariances $\mathrm{cov}(K_1(X), K_2(Y))$.

Finally, we recall that the path properties of a linear fractional stable motion strongly depend on the interplay between the parameters H and α. When $H - 1/\alpha > 0$ the process $(X_t)_{t \in \mathbb{R}}$ is Hölder continuous on compact intervals of any order smaller than $H - 1/\alpha$; we refer to [6] for more details on this property. If $H - 1/\alpha < 0$ the linear fractional stable motion explodes at jump times of the driving Lévy process L; in particular, X has unbounded paths on compact intervals. We demonstrate some sample paths of the linear fractional stable motions in Figure 1. In the critical case $H - 1/\alpha = 0$, we obviously have the identity $X_t = L_t$. In this situation the parameter estimation has been investigated in [1].

2.2. Review of the limit theory

In this section, we review some probabilistic results, which will be relevant for our estimation method. Due to stationarity of the increments and self-similarity of the process $(X_t)_{t \in \mathbb{R}}$, we can discuss the limit theory for the high and low frequency case simultaneously. We start by
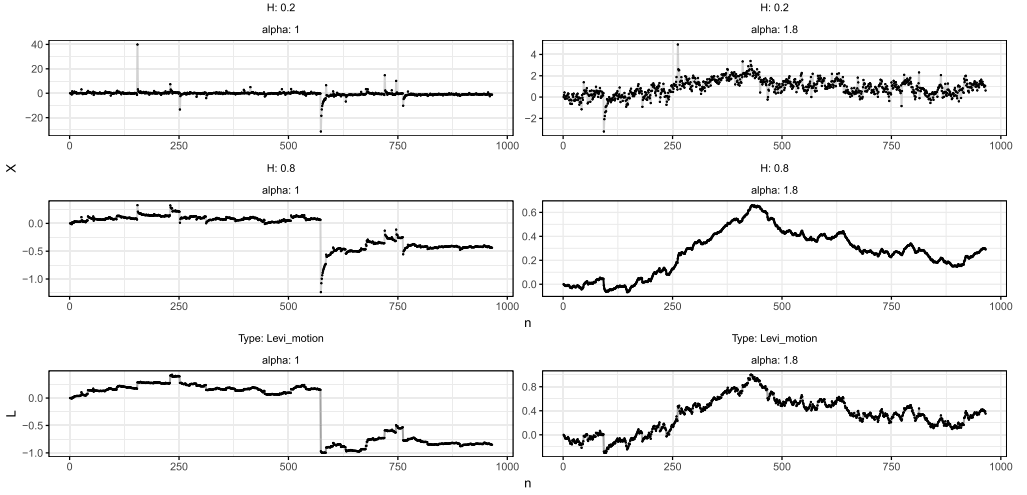
Figure 1. Left (from bottom to top): The driving symmetric stable Lévy process with $\alpha = 1$, linear fractional stable motions with parameters $\alpha = 1$, $H = 0.8$ and $\alpha = 1$, $H = 0.2$. Right (from bottom to top): The driving symmetric stable Lévy process with $\alpha = 1.8$, linear fractional stable motions with parameters $\alpha = 1.8$, $H = 0.8$ and $\alpha = 1.8$, $H = 0.2$.

introducing higher order increments of X. We denote by $\Delta_{i,k}^{n,r} X$ ($i, k, r, n \in \mathbb{N}$) the kth order increment of X at stage i/n and frequency r/n, that is,

$$\Delta_{i,k}^{n,r} X := \sum_{j=0}^{k} (-1)^j \binom{k}{j} X_{(i-rj)/n}, \quad i \geq rk. \tag{2.5}$$

Note that for $r = k = 1$ we obtain the usual increments $\Delta_{i,1}^{n,1} X = X_{i/n} - X_{(i-1)/n}$. For the ease of notation, we will often drop the index r (resp. k and n) in $\Delta_{i,k}^{n,r} X$ and other quantities when $r = 1$ (resp. $k = 1$ and $n = 1$). In particular, the low frequency kth order increments of X are denoted by

$$\Delta_{i,k}^{r} X := \sum_{j=0}^{k} (-1)^j \binom{k}{j} X_{i-rj}, \quad i \geq rk. \tag{2.6}$$

According to the self-similarity of the process $(X_t)_{t \in \mathbb{R}}$ we readily have that $(n^H \Delta_{i,k}^{n,r} X)_{i \geq rk} \stackrel{d}{=} (\Delta_{i,k}^{r} X)_{i \geq rk}$. Our main probabilistic tools will be statistics of the form

$$V_{\text{high}}(f; k, r)_n := \frac{1}{n} \sum_{i=rk}^{n} f\left(n^H \Delta_{i,k}^{n,r} X\right), \qquad V_{\text{low}}(f; k, r)_n := \frac{1}{n} \sum_{i=rk}^{n} f\left(\Delta_{i,k}^{r} X\right), \tag{2.7}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a measurable function. It is well known that the process $(X_t)_{t \in \mathbb{R}}$ is mixing, see, for example, [10]. Hence, Birkhoff's ergodic theorem implies the convergence $V_{\text{low}}(f; k, r)_n \to \mathbb{E}[f(\Delta^r_{rk,k} X)]$ almost surely whenever $\mathbb{E}[|f(\Delta^r_{rk,k} X)|] < \infty$. The same result holds in probability for the statistic $V_{\text{high}}(f; k, r)_n$ due to self-similarity of the process X. However, the weak limit theorems associated with the aforementioned law of large numbers and the framework of functions f with $\mathbb{E}[|f(\Delta^r_{rk,k} X)|] = \infty$ are not completely understood in the literature. To get an idea about possible limits that may appear, we briefly demonstrate some recent theoretical developments from the paper [5], where the case $f_p(x) = |x|^p$ ($p > 0$) has been investigated. We remark that their results are obtained for a wider class of processes, namely stationary increments Lévy moving average processes, and we adapt them to the setting of linear fractional stable motions.

We need to introduce some more notation to describe the various limits. For $p \in (-1, 1) \setminus \{0\}$ we define the constant

$$a_p := \begin{cases} \int_{\mathbb{R}} (1 - \cos(y)) |y|^{-1-p} \, dy : & p \in (0, 1) \\ \sqrt{2\pi} \Gamma(-p/2)/2^{p+1/2} \Gamma((p+1)/2) : & p \in (-1, 0), \end{cases} \tag{2.8}$$

where Γ denotes the Gamma function. It is easy to see that $a_p > 0$ is indeed finite in all relevant cases. For any functions $g, h \in L^\alpha(\mathbb{R})$, we introduce the notation

$$\theta(g, h)_p = a_p^{-2} \int_{\mathbb{R}^2} |xy|^{-1-p} U_{g,h}(x, y) \, dx \, dy, \tag{2.9}$$

where $U_{g,h}$ is defined in (2.3), whenever the above double integral is finite. Furthermore, for $k, r \in \mathbb{N}$, we define the function $h_{k,r} : \mathbb{R} \to \mathbb{R}$ by

$$h_{k,r}(x) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (x - rj)_+^{H-1/\alpha}, \quad x \in \mathbb{R}. \tag{2.10}$$

Below $(U_m)_{m \geq 1}$ is an i.i.d. $\mathcal{U}(0, 1)$-distributed sequence of random variables independent of L, $(T_m)_{m \geq 1}$ are jump times of L and $\Delta L_{T_m} := L_{T_m} - L_{T_{m-}}$ are jump sizes. The following result summarises the limit theory for the statistic $V_{\text{high}}(f_p; k)_n$ (i.e. $r = 1$) in the power variation setting.

Theorem 2.1 ([5], Theorems 1.1 and 1.2). *We consider the function $f_p(x) = |x|^p$ ($p > 0$) and assume that $H - 1/\alpha > 0$.*

(i) (*First order asymptotics*) *If $p > \alpha$ we obtain convergence in law*

$$n^{1-p/\alpha} V_{\text{high}}(f_p; k)_n \xrightarrow{d} \sum_{m: T_m \in [0,1]} |\Delta L_{T_m}|^p \left(\sum_{l=0}^{\infty} |h_k(l + U_m)|^p \right).$$

If $p < \alpha$, we deduce the law of large numbers

$$V_{\text{high}}(f_p; k)_n \xrightarrow{\mathbb{P}} m_{p,k} := \mathbb{E}[|\Delta_{k,k} X|^p].$$

(ii) (*Second order asymptotics*) *Assume that* $p < \alpha/2$. *If* $H < k - 1/\alpha$, *we obtain the central limit theorem*

$$\sqrt{n}\big(V_{\text{high}}(f_p; k)_n - m_{p,k}\big) \overset{d}{\longrightarrow} \mathcal{N}(0, \eta^2), \quad \eta^2 = \theta(h_k, h_k)_p + 2\sum_{j=1}^{\infty} \theta\big(h_k, h_k(\cdot + j)\big)_p,$$

where the quantity $\theta(g, h)$ *has been introduced at* (2.9). *If* $H > k - 1/\alpha$, *we deduce a non-central limit theorem*

$$n^{1-1/(1+\alpha(k-H))}\big(V_{\text{high}}(f_p; k)_n - m_{p,k}\big) \overset{d}{\longrightarrow} S,$$

where S *is a totally right skewed* $(1 + \alpha(k - H))$-*stable random variable with mean zero and scale parameter* $\widetilde{\sigma}$, *which is defined in* [5], *Theorem* 1.2.

We remark that the results of Theorem 2.1 remain valid for the low frequency statistic $V_{\text{low}}(f_p; k)_n$ due to self-similarity property of L. Apart from various critical cases Theorem 2.1 gives a rather complete understanding of the asymptotic behaviour of the power variation $V_{\text{high}}(f_p; k)_n$ in the setting $H - 1/\alpha > 0$. The strong law of large numbers in Theorem 2.1(i) will be useful for estimation of the parameter H. However, without an a priori knowledge about the stability parameter α, we can't insure that the condition $p < \alpha$ holds. Similarly, we would like to use the central limit theorem in Theorem 2.1(ii) whose convergence rate $\sqrt{n}$ is faster than the rate $n^{1-1/(1+\alpha(k-H))}$ in the non-central limit theorem. But the conditions of Theorem 2.1(ii) rely again on an a priori knowledge about α.

There are some related results in the literature. In [21], the authors have shown a central limit theorem for a standardised version of the statistic $\sum_{i=1}^{n} f(Y_i)$, where f is a *bounded* function and $(Y_t)_{t \in \mathbb{R}}$ is a stable moving average process. In a later work [22] the result has been extended to a certain class of unbounded functions f under the additional assumption that $\alpha \in (1, 2)$. Similarly to Theorem 2.1 the sufficient conditions for the validity of the central limit theorems in [21,22] depend on the interplay between the kernel function of the stable moving average process and the stability index α. We remark that extensions of these results in various directions will be necessary to obtain the full asymptotic theory for estimators of the parameter $\theta = (\sigma, \alpha, H)$.

2.3. A multivariate weak limit theorem

Although Theorem 2.1(ii) gives a rather complete picture of the weak limit theory in the power variation case, we will require a much stronger result for our statistical applications. We introduce the function $\psi_t : \mathbb{R} \to \mathbb{R}$ with $\psi_t(x) = \cos(tx)$ and define the statistics

$$\varphi_{\text{high}}(t; H, k)_n := V_{\text{high}}(\psi_t; k)_n \quad \text{and} \quad \varphi_{\text{low}}(t; k)_n := V_{\text{low}}(\psi_t; k)_n, \tag{2.11}$$

which correspond to $r = 1$. Notice that, in contrast to $\varphi_{\text{low}}(t; k)_n$, the high frequency statistic $\varphi_{\text{high}}(t; H, k)_n$ depends on the unknown self-similarity parameter H. In fact, this is the major difference between the high and low frequency settings, which will result in different rates of

convergence later on. Applying again the strong law of large numbers we readily obtain the strong consistency

$$\varphi_{\text{low}}(t;k)_n \xrightarrow{\text{a.s.}} \varphi(t;k) := \exp\big(-\big|\sigma\,\|h_k\|_\alpha t\big|^\alpha\big). \tag{2.12}$$

Clearly, the same result holds in probability for the high frequency statistic $\varphi_{\text{high}}(t;H,k)_n$. Next, we introduce various types of statistics, which will play a major role in estimation of the unknown parameter θ. More specifically, we will extend the definition of power variation to certain negative powers and prove a multivariate limit theorem for power variations and empirical characteristic functions. We fix $d \in \mathbb{N}$ and define the statistics for any $1 \le j \le d$, $r_j \in \{1,2\}$, $p \in (-1/2, 1/2) \setminus \{0\}$ and $t_j > 0$:

$$
\left.\begin{aligned}
W(n)_j^{(1)} &:= \sqrt{n}\big(V_{\text{low}}(f_p; k_j, r_j)_n - r_j^{pH} m_{p,k_j}\big) \\
W(n)_j^{(2)} &:= \sqrt{n}\big(V_{\text{low}}(\psi_{t_j}; k_j)_n - \varphi(t_j; k_j)\big)
\end{aligned}\right\} \quad \text{when } k_j > H + 1/\alpha
$$
$$
\left.\begin{aligned}
S(n)_j^{(1)} &:= n^{1-1/(1+\alpha(k-H))}\big(V_{\text{low}}(f_p; k, r_j)_n - r_j^{pH} m_{p,k}\big) \\
S(n)_j^{(2)} &:= n^{1-1/(1+\alpha(k-H))}\big(V_{\text{low}}(\psi_{t_j}; k)_n - \varphi(t_j; k)\big)
\end{aligned}\right\} \quad \text{when } k < H + 1/\alpha
\tag{2.13}
$$

Note the identity $\mathbb{E}[|\Delta_{rk,k}^r X|^p] = r^{pH} m_{p,k}$, which explains the centring of the statistics $W(n)^{(1)}$ and $S(n)^{(1)}$. We remark that the functionals $W(n)^{(1)}$ and $W(n)^{(2)}$ are in the domain of attraction of the normal distribution (under appropriate assumption on the powers p) while the functionals $S(n)^{(1)}$ and $S(n)^{(2)}$ are in the domain of attraction of the $(1+\alpha(k-H))$-stable distribution. The latter fact is rather surprising since the statistic $S(n)_j^{(2)}$ exhibits finite moments of any order.

Before we proceed with the main result of this section, we need to introduce some more notation. In the first step, for any $x \in \mathbb{R}$, we define the functions

$$
\begin{aligned}
\Phi_j^{(1)}(x) &= \mathbb{E}\big[f_p\big(\Delta_{r_j k,k}^{r_j} X + x\big)\big] - \mathbb{E}\big[f_p\big(\Delta_{r_j k,k}^{r_j} X\big)\big], \\
\Phi_j^{(2)}(x) &= \mathbb{E}\big[\psi_{t_j}(\Delta_{k,k} X + x)\big] - \mathbb{E}\big[\psi_{t_j}(\Delta_{k,k} X)\big].
\end{aligned}
\tag{2.14}
$$

Since the functions f_p and ψ_t are even we readily obtain that $\Phi_j^{(l)}(0) = \nabla \Phi_j^{(l)}(0) = 0$ for all l, j. Thus, using Lemma 0.5, we deduce the growth estimates

$$
\big|\Phi_j^{(1)}(x)\big| \le C\big(x^2 \wedge |x|^{\max\{p,0\}}\big), \qquad \big|\Phi_j^{(2)}(x)\big| \le C\big(x^2 \wedge 1\big),
\tag{2.15}
$$

for some positive constant C. Next, we introduce the functions

$$
\overline{\Phi}_j^{(1)}(x) = \sum_{i=1}^{\infty} \Phi_j^{(1)}\big(h_{k,r_j}(i)x\big), \qquad \overline{\Phi}_j^{(2)}(x) = \sum_{i=1}^{\infty} \Phi_j^{(2)}\big(h_k(i)x\big).
\tag{2.16}
$$

Note that these functions are indeed finite due to (2.15) and the estimate $|h_{k,r}(x)| \le C|x|^{H-1/\alpha-k}$ for large x. Finally, we set $\overline{\Phi} = (\overline{\Phi}^{(1)}, \overline{\Phi}^{(2)}) = (\overline{\Phi}_1^{(1)}, \dots, \overline{\Phi}_d^{(1)}, \overline{\Phi}_1^{(2)}, \dots, \overline{\Phi}_d^{(2)})$. The main probabilistic result of this paper is the following theorem.

Theorem 2.2. *Assume that either $p \in (-1/2, 0)$ or $p \in (0, 1/2)$ and $p < \alpha/2$. Set $W(n)^{(i)} = (W(n)_1^{(i)}, \ldots, W(n)_d^{(i)})$ and $S(n)^{(i)} = (S(n)_1^{(i)}, \ldots, S(n)_d^{(i)})$ for $i = 1, 2$. Then we obtain weak convergence in law on $\mathbb{R}^{4d}$:*

$$\left(W(n)^{(1)}, W(n)^{(2)}, S(n)^{(1)}, S(n)^{(2)}\right) \xrightarrow{d} \left(W^{(1)}, W^{(2)}, S^{(1)}, S^{(2)}\right), \tag{2.17}$$

where $W = (W^{(1)}, W^{(2)})$ and $S = (S^{(1)}, S^{(2)})$ are independent, W is a centred $2d$-dimensional normal distribution with covariance matrix determined by

$$\mathrm{cov}\left(W_j^{(i)}, W_{j'}^{(i')}\right) = \lim_{n \to \infty} \mathrm{cov}\left(W(n)_j^{(i)}, W(n)_{j'}^{(i')}\right) \quad 1 \le j, j' \le d, i, i' = 1, 2,$$

and $S^{(1)}$, $S^{(2)}$ are independent d-dimensional $(1 + \alpha(k - H))$-stable random variables. The law of $S^{(1)}$ (resp. $S^{(2)}$) is determined by the Lévy measure ν_1 (resp. ν_2) whose support is the cone $(\mathbb{R}_+)^d$ (resp. $(\mathbb{R}_-)^d$). More specifically, for any Borel sets $A_1 \in (\mathbb{R}_+)^d$, $A_2 \in (\mathbb{R}_-)^d$ bounded away from 0 the quantities $\nu_1(A_1)$, $\nu_2(A_2)$ are determined by the identity

$$\nu_l(A_l) = \lim_{n \to \infty} n \mathbb{P}\left(n^{-1/(1+\alpha(k-H))} \overline{\Phi}^{(l)}(L_1) \in A_l\right), \quad l = 1, 2. \tag{2.18}$$

The probabilistic result of Theorem 2.2 is new in the literature; neither the negative power variations nor the (real part of) empirical characteristic function have been studied from the distributional perspective. We remark that the statistics $W(n)^{(1)}$ and $S(n)^{(1)}$ use the same powers p while the quantities $S(n)^{(1)}$ and $S(n)^{(2)}$ are based on the same order of increments k. The result of Theorem 2.2 does not really use these particular restrictions, but its statement is sufficient for the statistical application under investigation.

There exists an explicit expression for the covariance matrix of the limit W. We obtain the following representations:

$$\mathrm{cov}\left(W_j^{(1)}, W_{j'}^{(1)}\right) = \sum_{l \in \mathbb{Z}} \theta\left(h_{k_j, r_j}, h_{k_{j'}, r_{j'}}(\cdot + l)\right)_p,$$

$$\mathrm{cov}\left(W_j^{(2)}, W_{j'}^{(2)}\right) = \frac{1}{2} \sum_{l \in \mathbb{Z}} \left(U_{h_{k_j}, h_{k_{j'}}(\cdot + l)}(t_j, t_{j'}) + U_{h_{k_j}, -h_{k_{j'}}(\cdot + l)}(t_j, t_{j'})\right), \tag{2.19}$$

$$\mathrm{cov}\left(W_j^{(1)}, W_{j'}^{(2)}\right) = \sum_{l \in \mathbb{Z}} \overline{\theta}(l)_{jj'},$$

with

$$\overline{\theta}(l)_{jj'} = -a_p^{-1} \int_{\mathbb{R}} |y|^{-1-p} U_{h_{k_j, r_j}, h_{k_{j'}}(\cdot + l)}(y, t_{j'}) \, dy.$$

We will prove that $\mathrm{cov}(W) < \infty$ in all relevant cases and the mapping $(\sigma, \alpha, H) \mapsto \mathrm{cov}(W)$ is continuous (see Section 0.1). In principle, the latter allows us to estimate the covariance matrix $\mathrm{cov}(W) < \infty$ and thus to obtain a feasible version of the central limit theorem in Theorem 2.2, although we will use a different approach in the simulation study.

Similarly, the Lévy measures ν_l ($l = 1, 2$) can be determined explicitly. First of all, the representation (0.2) from Section 0.1 implies the identities

$$\Phi_j^{(1)}(x) = a_p^{-1} \int_{\mathbb{R}} \left(1 - \cos(ux)\right) \exp\left(-|\sigma\|h_{k,r_j}\|_\alpha u|^\alpha\right) |u|^{-1-p} \, du,$$

$$\Phi_j^{(2)}(x) = \left(\cos(t_j x) - 1\right) \exp\left(-|\sigma\|h_k\|_\alpha t_j|^\alpha\right).$$

In particular, it holds that $\Phi_j^{(1)}(x) \geq 0$ and $\Phi_j^{(2)}(x) \leq 0$. In the next step we need to determine the asymptotic behaviour of $\overline{\Phi}_j^{(1)}(x)$ (resp. $\overline{\Phi}_j^{(2)}(x)$) as $x \to \infty$ (resp. as $x \to -\infty$). By the substitution $u = (x/z)^{1/(k+1/\alpha - H)}$ we have that

$$x^{1/(H-k-1/\alpha)} \overline{\Phi}_j^{(1)}(x)$$

$$= x^{1/(H-k-1/\alpha)} \int_0^\infty \Phi_j^{(1)}\left(h_{k,r_j}\left(\lfloor u \rfloor + 1\right)x\right) du$$

$$= (k + 1/\alpha - H)^{-1} \int_0^\infty \Phi_j^{(1)}\left(h_{k,r_j}\left(\lfloor (x/z)^{1/(k+1/\alpha - H)} \rfloor + 1\right)x\right) z^{-1+1/(H-k-1/\alpha)} \, dz$$

$$\to c_j^{(1)} := (k + 1/\alpha - H)^{-1} \int_0^\infty \Phi_j^{(1)}\left(r_j^k \prod_{i=0}^{k-1}(H - 1/\alpha - i) \cdot z\right)$$

$$\times z^{-1+1/(H-k-1/\alpha)} \, dz \tag{2.20}$$

as $x \to \infty$. The convergence at (2.20) follows from the asymptotic behaviour $h_{k,r_j}(x) \sim r_j^k \prod_{i=0}^{k-1}(H - 1/\alpha - i) \cdot x^{H-1/\alpha - k}$ as $x \to \infty$. Applying the same technique, we deduce that

$$|x|^{1/(H-k-1/\alpha)} \overline{\Phi}_j^{(2)}(x)$$

$$\to c_j^{(2)} := (k + 1/\alpha - H)^{-1} \int_0^\infty \Phi_j^{(2)}\left(\prod_{i=0}^{k-1}(H - 1/\alpha - i) \cdot z\right)$$

$$\times z^{-1+1/(H-k-1/\alpha)} \, dz \tag{2.21}$$

as $x \to -\infty$. Now, both measures ν_1 and ν_2 from Theorem 2.2 can be related to the Lévy measure ν of L. We introduce the mappings $\tau_1 : \mathbb{R} \to (\mathbb{R}_+)^d$ and $\tau_2 : \mathbb{R} \to (\mathbb{R}_-)^d$ via

$$\tau_1(x) = x^{1/(k+1/\alpha - H)}\left(c_1^{(1)}, \ldots, c_d^{(1)}\right), \qquad \tau_2(x) = |x|^{1/(k+1/\alpha - H)}\left(c_1^{(2)}, \ldots, c_d^{(2)}\right).$$

Then, for Borel sets A_1, A_2 as defined in Theorem 2.2, we deduce the identity

$$\nu_l(A_l) = \lim_{n \to \infty} n\mathbb{P}\left(\tau_l\left(n^{-1/\alpha} L_1\right) \in A_l\right) = \nu\left(\tau_l^{-1}(A_l)\right), \quad l = 1, 2. \tag{2.22}$$

3. Statistical inference in the continuous case $H - 1/\alpha > 0$

We start with the continuous case $H - 1/\alpha > 0$, which turns out to be somewhat easier to treat compared to the general setting. Since $H \in (0, 1)$ and $\alpha \in (0, 2)$, condition $H - 1/\alpha > 0$ implies the restrictions

$$\alpha \in (1, 2) \quad \text{and} \quad H \in (1/2, 1).$$

It is the lower bound $\alpha > 1$ that enables us to use the law of large numbers in Theorem 2.1(i) whenever $p < 1$, and the central limit theorem in Theorem 2.1(ii) whenever $p < 1/2$ and $H < k - 1/\alpha$. The latter condition $H < k - 1/\alpha$ never holds for $k = 1$ since $0 < H - 1/\alpha < 1 - 2/\alpha < 0$ gives a contradiction, but it is always satisfied for any $k \geq 2$ since

$$H < 1 < k - 1/\alpha \quad \text{for any } k \geq 2,$$

because $\alpha > 1$.

Now, we introduce an estimator for the parameter $\theta = (\sigma, \alpha, H)$ in high and low frequency setting. We start with the statistical inference for the self-similarity parameter H, which is based upon a ratio statistic that compares power variations at two different frequencies. More specifically, we define the quantities

$$R_{\text{high}}(p, k)_n := \frac{\sum_{i=2k}^{n} |\Delta_{i,k}^{n,2} X|^p}{\sum_{i=k}^{n} |\Delta_{i,k}^{n,1} X|^p}, \qquad R_{\text{low}}(p, k)_n := \frac{\sum_{i=2k}^{n} |\Delta_{i,k}^{2} X|^p}{\sum_{i=k}^{n} |\Delta_{i,k}^{1} X|^p}, \qquad (3.1)$$

where the increments $\Delta_{i,k}^{r} X$ have been defined at (2.6). We obtain the convergence

$$R_{\text{high}}(p, k)_n \xrightarrow{\mathbb{P}} 2^{pH}, \qquad R_{\text{low}}(p, k)_n \xrightarrow{\text{a.s.}} 2^{pH}$$

for any $p \in (0, 1)$ as an immediate consequence of Theorem 2.1(i). Consequently, defining the statistics

$$\widehat{H}_{\text{high}}(p, k)_n := \frac{1}{p} \log_2 R_{\text{high}}(p, k)_n, \qquad \widehat{H}_{\text{low}}(p, k)_n := \frac{1}{p} \log_2 R_{\text{low}}(p, k)_n, \qquad (3.2)$$

we deduce the consistency $\widehat{H}_{\text{high}}(p, k)_n \xrightarrow{\mathbb{P}} H$, $\widehat{H}_{\text{low}}(p, k)_n \xrightarrow{\text{a.s.}} H$ as $n \to \infty$ for any $k \geq 1$ and any $p \in (0, 1)$. We remark that this type of ratio statistics is commonly used in the framework of fBm's when estimating the Hurst parameter H (see e.g. [16] among many others). In the Gaussian setting, which corresponds to $\alpha = 2$, the central limit theorem for the quantity $\sqrt{n}(\widehat{H}_{\text{high}}(p, k)_n - H)$ holds for all $k \geq 2$ and also for $k = 1$ if further $H \in (0, 3/4)$. As we indicated above, in the framework of pure jump α-stable driving motion L the central limit theorem never holds if $k = 1$. Hence, there is no smooth transition between the non-Gaussian and Gaussian setting when $\alpha \to 2$.

The estimation strategy for the parameter $\theta = (\sigma, \alpha, H)$ based on high frequency observations is now straightforward: Infer the self-similarity parameter H by (3.2) and use the plug-in estimator $\varphi_{\text{high}}(t; \widehat{H}_{\text{high}}(p, k), k)_n$ for two different values of t to infer the scale parameter σ and the

stability index α. For the latter step, we consider $t_2 > t_1 > 0$ and observe the identities

$$\sigma = \left(-\log \varphi(t_1; k)\right)^{1/\alpha} / t_1 \|h_k\|_\alpha, \qquad \alpha = \frac{\log |\log \varphi(t_2; k)| - \log |\log \varphi(t_1; k)|}{\log t_2 - \log t_1}.$$

Recalling that h_k depends on α and H, we readily obtain a function G such that

$$(\sigma, \alpha) = G\big(\varphi(t_1; k), \varphi(t_2; k), H\big), \tag{3.3}$$

where we applied the above identities. Next, we present the estimator of the pair (σ, α) in high and low frequency setting, recalling that the estimators of the self-similarity parameter H have been defined at (3.2). We introduce the following estimators:

$$\big(\widehat{\sigma}_{\text{high}}(k, t_1, t_2)_n, \widehat{\alpha}_{\text{high}}(k, t_1, t_2)_n\big)$$
$$= G\big(\varphi_{\text{high}}\big(t_1; \widehat{H}_{\text{high}}(p, k)_n, k\big)_n, \varphi_{\text{high}}\big(t_2; \widehat{H}_{\text{high}}(p, k)_n, k\big)_n, \widehat{H}_{\text{high}}(p, k)_n\big), \tag{3.4}$$
$$\big(\widehat{\sigma}_{\text{low}}(k, t_1, t_2)_n, \widehat{\alpha}_{\text{low}}(k, t_1, t_2)_n\big) = G\big(\varphi_{\text{low}}(t_1; k)_n, \varphi_{\text{low}}(t_2; k)_n, \widehat{H}_{\text{low}}(p, k)_n\big).$$

Before we present the main result of this section we need to introduce more notation. We define the functions $v_p : \mathbb{R}^2_+ \to \mathbb{R}$ and $F : \mathbb{R}^2_+ \times \mathbb{R}^2 \to \mathbb{R}^3$ by

$$v_p(x, y) = p^{-1}(\log_2 y - \log_2 x), \qquad F(x, y, u, w) = \big(G\big(u, w, v_p(x, y)\big), v_p(x, y)\big), \tag{3.5}$$

and let JF denotes the Jacobian of F. For any matrix A we write $A^\star$ for its transpose. The asymptotic normality in the low and high frequency setting is summarised in the following theorem.

Theorem 3.1. *Consider the linear fractional stable motion* $(X_t)_{t \in \mathbb{R}}$ *introduced at* (1.1). *Let* $k \geq 2$ *and* $t_2 > t_1 > 0$.

(i) (*Low frequency case*) *Let* $W = (W^{(1)}, W^{(2)})$ *be the 4-dimensional normal limit defined in Theorem* 2.2 *associated with* $d = 2$, $p \in (0, 1/2)$, $k_1 = k_2 = k$ *and* $r_j = j$. *Then we obtain the central limit theorem*

$$\sqrt{n} \begin{pmatrix} \widehat{\sigma}_{\text{low}}(k, t_1, t_2)_n - \sigma \\ \widehat{\alpha}_{\text{low}}(k, t_1, t_2)_n - \alpha \\ \widehat{H}_{\text{low}}(p, k)_n - H \end{pmatrix} \xrightarrow{d} B^{\text{nor}}_{\text{low}}(p, k) = JF\big(m_{p,k}, 2^H m_{p,k}, \varphi(t_1; k), \varphi(t_2; k)\big) W^\star.$$

(ii) (*High frequency case*) *We obtain the central limit theorem*

$$\begin{pmatrix} \sqrt{n}(\log n)^{-1}\big(\widehat{\sigma}_{\text{high}}(k, t_1, t_2)_n - \sigma\big) \\ \sqrt{n}(\log n)^{-1}\big(\widehat{\alpha}_{\text{high}}(k, t_1, t_2)_n - \alpha\big) \\ \sqrt{n}\big(\widehat{H}_{\text{high}}(p, k)_n - H\big) \end{pmatrix}$$
$$\xrightarrow{d} B^{\text{nor}}_{\text{high}}(p, k) = \nabla v\big(m_{p,k}, 2^H m_{p,k}\big)\big(W^{(1)}\big)^\star$$

$$\times \begin{pmatrix} \nabla G_1\big(\varphi(t_1; k), \varphi(t_2; k), H\big)\big(t_1\varphi'(t_1; k), t_2\varphi'(t_2; k), 0\big)^\star \\ \nabla G_2\big(\varphi(t_1; k), \varphi(t_2; k), H\big)\big(t_1\varphi'(t_1; k), t_2\varphi'(t_2; k), 0\big)^\star \\ 1 \end{pmatrix}.$$

We remark that the central limit theorem of Theorem 3.1(i) is a simple consequence of Theorem 2.2 and the delta method. In contrast to the low frequency case Theorem 3.1(ii) is degenerate in the sense that the limit distribution is solely driven by the asymptotics of the term $\sqrt{n}(\widehat{H}_{\mathrm{high}}(p, k)_n - H)$. Since the parameter H enters the quantity $\varphi_{\mathrm{high}}(t; H, k)_n$ via n^H the additional term $(\log n)^{-1}$ appears in the convergence rate.

For a later use, we need to extend the definition of the random variables $B_{\mathrm{high}}^{\mathrm{nor}}(p, k)$ and $B_{\mathrm{low}}^{\mathrm{nor}}(p, k)$ to various directions. First of all, we will allow for negative powers $-p$ with $p \in (0, 1/2)$. Secondly, we would like to define the same limiting variables but associated with the stable limit $S = (S^{(1)}, S^{(2)})$ from Theorem 2.2 rather than W. Thus, for $d = 2$, $p \in (-1/2, 1/2) \setminus \{0\}$, $k_1 = k_2 = k$ and $r_j = j$, we set

$$B_{\mathrm{low}}^{\mathrm{sta}}(p, k) = J F\big(m_{p,k}, 2^H m_{p,k}, \varphi(t_1; k), \varphi(t_2; k)\big)S^\star,$$

$$B_{\mathrm{high}}^{\mathrm{sta}}(p, k) = \nabla v\big(m_{p,k}, 2^H m_{p,k}\big)\big(S^{(1)}\big)^\star$$

$$\times \begin{pmatrix} \nabla G_1\big(\varphi(t_1; k), \varphi(t_2; k), H\big)\big(t_1\varphi'(t_1; k), t_2\varphi'(t_2; k), 0\big)^\star \\ \nabla G_2\big(\varphi(t_1; k), \varphi(t_2; k), H\big)\big(t_1\varphi'(t_1; k), t_2\varphi'(t_2; k), 0\big)^\star \\ 1 \end{pmatrix}.$$

Remark 3.1. In Theorem 3.1, we use two values $t_1, t_2 \in \mathbb{R}_+^2$ and an estimator $\widehat{H}$ to infer the parameters (σ, α). Applying basic statistical principles it is more natural to use all $t \in \mathbb{R}_+$ for the estimation procedure. For example, when considering the low frequency framework, we may estimate the parameters (σ, α) via a minimal contrast approach. Given a positive weight function $w \in L^1(\mathbb{R}_+)$ we obtain an estimator $(\widetilde{\sigma}_n, \widetilde{\alpha}_n)$ of (σ, α) by

$$(\widetilde{\sigma}_n, \widetilde{\alpha}_n) \in \operatorname*{argmin}_{\theta \in \mathbb{R}_+ \times (0,2)} \int_0^\infty \big(\varphi_{\mathrm{low}}(t; k)_n - \widehat{\varphi}(t; k)\big)^2 w(t)\, dt,$$

where $\widehat{\varphi}(t; k) = \varphi(\widehat{H}_{\mathrm{low}}(p, k)_n, t; k)$. In this setting, we are likely to require tightness or a similar property of the stochastic process $\varphi_{\mathrm{low}}(\cdot; k)_n$ to prove asymptotic normality of $(\widetilde{\sigma}_n, \widetilde{\alpha}_n)$. However, this seems to be a non-trivial problem, at least when using standard tightness criteria for the space $(C(\mathbb{R}_+), \|\cdot\|_\infty)$. We leave it for future research.

Remark 3.2. The described statistical methodology can be applied to more general processes than the mere linear fractional stable motion. In the paper, [5] the authors investigated limit theorems for stochastic processes of the form

$$Y_t = \int_{\mathbb{R}} \big\{g(t - s) - g_0(-s)\big\} dL_s,$$

where g, g_0 are deterministic functions vanishing on $\mathbb{R}_-$ with $g(x) = x^{H - 1/\alpha} f(x)$ and $f(0) \neq 0$, and L is a symmetric α-stable Lévy motion. In the high frequency setting, the process Y exhibits

the *tangent process* $f(0)X$, that is, we have that

$$\Delta_{i,k}^{n,r} Y \approx f(0)\Delta_{i,k}^{n,r} X.$$

In particular, we believe that under certain assumptions on f (cf. [5]) the central limit theorem part of Theorem 2.2 should hold for the more general class of processes Y (however, it would require a non-trivial modification of the proofs). Hence, in this semi-parametric model it should be possible to estimate the parameter $(|f(0)|\sigma, \alpha, H)$ via the same approach as presented in Theorem 3.1(ii). We remark that the function f can't be inferred from high frequency observations on a fixed time interval.

4. Statistical inference in the general case

In this section, we treat the case of a general linear fractional stable motion as it has been introduced at (1.1). We recall that in the continuous setting the restriction $H - 1/\alpha > 0$ has led to the lower bound $\alpha > 1$, which is essential for obtaining the asymptotic results of Theorem 3.1. Without having an explicit lower bound for the stability parameter α statistical inference turns out to be more complex. As a consequence, we will require a different estimation method for the self-similarity parameter H and a two-step procedure to choose the right order of increments k. Furthermore, in order to obtain fast rates of convergence we need different treatments for the low and high frequency frameworks.

4.1. Low frequency setting

We note that the basic idea behind the ratio statistic $R_{\text{low}}(p, k)_n$ introduced in (3.1) is the homogeneity of the function $f_p(x) = |x|^p$ and the fact that $m_{p,k} < \infty$ which is a consequence of $p < \alpha$ (for the associated central limit theorem we need the stronger condition $p < \alpha/2$). In order to keep both properties, we may instead consider the negative power variation, which corresponds to the function $f_{-p}(x) = |x|^{-p}$, and we assume throughout this section that $p \in (0, 1/2)$. This approach has been originally proposed in [14], although central limit theorems have not been investigated in this setting. Note that the function f_{-p} is still homogenous and $m_{-2p,k} < \infty$, which is due to the fact that for any random variable Y with bounded density near 0 it holds that $\mathbb{E}[|Y|^a] < \infty$ for all $a \in (-1, 0)$. Thus, $\widehat{H}_{\text{low}}(-p, k)_n$ is a strongly consistent estimator of the parameter H for any $p \in (0, 1/2)$.

In the next step, we need to ensure that we end up in the domain of attraction of the central limit theorem in Theorem 2.1(ii), which requires that $k > H + 1/\alpha$. To guarantee this, we need a preliminary estimator of the parameter α. They are obtained as in (3.4) using the function f_{-p} and $k = 1$:

$$\widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n = G_2\big(\varphi_{\text{low}}(t_1)_n, \varphi_{\text{low}}(t_2)_n, \widehat{H}_{\text{low}}(-p)_n\big), \tag{4.1}$$

where $G = (G_1, G_2)$. Notice that this estimator is consistent, but we do not know if it is in the domain of attraction of a normal distribution or not. Now, we define

$$\widehat{k}_{\text{low}}(t_1, t_2)_n = 2 + \big\lfloor \widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n^{-1} \big\rfloor. \tag{4.2}$$

For the sake of brevity we write $\widehat{k}_{\text{low}} = \widehat{k}_{\text{low}}(t_1, t_2)_n$. In the second step we estimate the parameter $\theta = (\sigma, \alpha, H)$ using $\widehat{k}_{\text{low}}$. The self-similarity parameter H is thus estimated by $\widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n$. Next, similarly to definitions at (3.4), we introduce the estimators

$$\big(\widetilde{\sigma}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n, \widetilde{\alpha}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n\big)$$
$$= G\big(\varphi_{\text{low}}(t_1; \widehat{k}_{\text{low}})_n, \varphi_{\text{low}}(t_2; \widehat{k}_{\text{low}})_n, \widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n\big). \tag{4.3}$$

In order to determine, the asymptotic distribution of the proposed estimators we will need the full force of Theorem 2.2. Due to definition (4.2), we also require a separate treatment of the cases $\alpha^{-1} \notin \mathbb{N}$ and $\alpha^{-1} \in \mathbb{N}$. In the first case $\widehat{k}_{\text{low}} \xrightarrow{\text{a.s.}} 2 + \lfloor \alpha^{-1} \rfloor$ while in the second case we will have

$$\mathbb{P}\big(\widehat{k}_{\text{low}} = 2 + \alpha^{-1}\big) \to \lambda \quad \text{and} \quad \mathbb{P}\big(\widehat{k}_{\text{low}} = 1 + \alpha^{-1}\big) \to 1 - \lambda$$

for a certain constant $\lambda \in (0, 1)$. In the first setting, which is easier to treat, we obtain the following result.

Theorem 4.1. *Let X be the linear fractional stable motion defined at (1.1). Assume that $p \in (0, 1/2)$ and $\alpha^{-1} \notin \mathbb{N}$. We obtain the central limit theorem*

$$\sqrt{n} \begin{pmatrix} \widetilde{\sigma}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \sigma \\ \widetilde{\alpha}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \alpha \\ \widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n - H \end{pmatrix} \xrightarrow{d} B_{\text{low}}^{\text{nor}}\big(-p, 2 + \lfloor \alpha^{-1} \rfloor\big).$$

In the framework $\alpha^{-1} \in \mathbb{N}$ we distinguish two further cases, that determine the asymptotic behaviour of the preliminary estimate $\widehat{\alpha}_{\text{low}}^0$, which is constructed using $k = 1$. According to Theorem 2.2 we are in the domain of the validity of a central limit theorem when $H < 1 - 1/\alpha$ while a non-central limit theorem holds if $H > 1 - 1/\alpha$.

Proposition 4.2. *Let X be the linear fractional stable motion defined at (1.1). Assume that $p \in (0, 1/2)$.*

(i) *(Normal case) Assume that $H < 1 - 1/\alpha$. Then we obtain the central limit theorem*

$$\sqrt{n}\big(\widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n - \alpha\big) \xrightarrow{d} B_{\text{low}}^{\text{nor}}(-p, 1)_2.$$

(ii) *(Stable case) Assume that $H > 1 - 1/\alpha$. Then we obtain the weak limit theorem*

$$n^{1-1/(1+\alpha(1-H))}\big(\widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n - \alpha\big) \xrightarrow{d} B_{\text{low}}^{\text{sta}}(-p, 1)_2.$$

We note that the result of Proposition 4.2(ii) is essentially the same as in the asymptotically normal regime except that the convergence rate is now $n^{1-1/(1+\alpha(1-H))}$ and the normal limit W is replaced by S.

The next theorem presents the statistical behaviour of the estimator $(\widetilde{\sigma}_{\text{low}}, \widetilde{\alpha}_{\text{low}}, \widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n)$ in the case $\alpha^{-1} \in \mathbb{N}$.

Theorem 4.3. *Let X be the linear fractional stable motion defined at (1.1). Assume that $p \in (0, 1/2)$ and $\alpha^{-1} \in \mathbb{N}$.*

(i) *(Case $H < 1 - 1/\alpha$) Assume that $H < 1 - 1/\alpha$. Then we obtain*

$$\sqrt{n} \begin{pmatrix} \widetilde{\sigma}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \sigma \\ \widetilde{\alpha}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \alpha \\ \widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n - H \end{pmatrix} \xrightarrow{d} D_{\text{low}}^{\text{nor}},$$

where the probability distribution $D_{\text{low}}^{\text{nor}}$ on $\mathbb{R}^3$ is given by

$$D_{\text{low}}^{\text{nor}}(\cdot) = \mathbb{P}\big(\{B_{\text{low}}^{\text{nor}}(-p, 2 + \alpha^{-1}) \in \cdot\} \cap \{B_{\text{low}}^{\text{nor}}(-p, 1)_2 < 0\}\big)$$
$$+ \mathbb{P}\big(\{B_{\text{low}}^{\text{nor}}(-p, 1 + \alpha^{-1}) \in \cdot\} \cap \{B_{\text{low}}^{\text{nor}}(-p, 1)_2 > 0\}\big).$$

(ii) *(Case $H > 1 - 1/\alpha$) Assume that $H > 1 - 1/\alpha$. Then we obtain*

$$\sqrt{n} \begin{pmatrix} \widetilde{\sigma}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \sigma \\ \widetilde{\alpha}_{\text{low}}(\widehat{k}_{\text{low}}, t_1, t_2)_n - \alpha \\ \widehat{H}_{\text{low}}(-p, \widehat{k}_{\text{low}})_n - H \end{pmatrix} \xrightarrow{d} D_{\text{low}}^{\text{sta}},$$

where the probability distribution $D_{\text{low}}^{\text{sta}}$ on $\mathbb{R}^3$ is given by

$$D_{\text{low}}^{\text{sta}}(\cdot) = \mathbb{P}\big(B_{\text{low}}^{\text{sta}}(-p, 1)_2 < 0\big)\mathbb{P}\big(B_{\text{low}}^{\text{nor}}(-p, 2 + \alpha^{-1}) \in \cdot\big)$$
$$+ \mathbb{P}\big(B_{\text{low}}^{\text{sta}}(-p, 1)_2 > 0\big)\mathbb{P}\big(B_{\text{low}}^{\text{nor}}(-p, 1 + \alpha^{-1}) \in \cdot\big).$$

According to Theorem 2.2 the statistic $(B_{\text{low}}^{\text{nor}}(-p, k), B_{\text{low}}^{\text{nor}}(-p, 1))$ is jointly normal for $k \in \{1 + \alpha^{-1}, 2 + \alpha^{-1}\}$. Thus, the probability distribution $D_{\text{low}}^{\text{nor}}$ can be easily computed using conditioning rules for normal distribution.

Note however that it is problematic to use Theorem 4.3 for constructing confidence regions since we do not know a priori whether part (i) or part (ii) applies. We now introduce a decision rule that helps us to solve this problem. Let $t_4 > t_3 > t_2 > t_1 > 0$ be given real numbers and let $\widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n$, $\widehat{\alpha}_{\text{low}}^0(t_3, t_4)_n$ be two estimators of parameter $\alpha \in (0, 2)$ defined at (4.1). Then, similarly to Proposition 4.2, we deduce that

$$a_n\big(\widehat{\alpha}_{\text{low}}^0(t_3, t_4)_n - \widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n\big) \quad \text{converges in law,}$$

where $a_n = \sqrt{n}$ if $H < 1 - 1/\alpha$ and $a_n = n^{1 - 1/(1 + \alpha(1 - H))}$ if $H > 1 - 1/\alpha$. Hence, we immediately conclude the convergence

$$d_n := -\frac{\log|\widehat{\alpha}_{\text{low}}^0(t_3, t_4)_n - \widehat{\alpha}_{\text{low}}^0(t_1, t_2)_n|}{\log(n)} \xrightarrow{\mathbb{P}} \begin{cases} 1/2: & \text{if } H < 1 - 1/\alpha \\ 1 - 1/(1 + \alpha(1 - H)): & \text{if } H > 1 - 1/\alpha \end{cases}$$

In other word, the statistic d_n helps us to identify the rate of convergence, but it has a bias of order $1/\log(n)$. Our decision rule is now as follows: Use Theorem 4.3(i) to perform statistical inference if

$$d_n > 1/2 - \big(\log(n)\big)^{-1+\epsilon}$$

for some small chosen $\epsilon > 0$; otherwise use Theorem 4.3(ii).

Remark 4.1. While we can obtain fully feasible asymptotic theory if we know whether $\alpha^{-1} \in \mathbb{N}$ or not, we are not yet able to deduce a complete statistical method without this a priori knowledge. Possibly subsampling procedures are required to obtain empirical confidence regions that automatically adapt to a given setting.

4.2. High frequency setting

In the framework of high frequency observations the application of the empirical characteristic function might lead to suboptimal convergence rates for the estimator of (σ, α). This comes from the following observation. Assume that $\alpha < 1$. Using the inequality $|\cos(x) - \cos(y)| \le |x - y|^{\alpha'}$ for any $\alpha' < \alpha$ we obtain the upper bound

$$\big|\varphi_{\text{high}}\big(t; \widehat{H}_{\text{high}}(p, k)_n, k\big)_n - \varphi_{\text{high}}(t; H, k)_n\big|$$
$$\le \frac{t^{\alpha'} (n^{\widehat{H}_{\text{high}}(p,k)_n - H} - 1)^{\alpha'}}{n} \sum_{i=k}^{n} \big|n^H \Delta_{i,k}^n X\big|^{\alpha'} = O_{\mathbb{P}}\big((n^{-1/2} \log n)^{-\alpha'/2}\big),$$

where the last statement follows from $\mathbb{E}[|\Delta_{k,k} X|^{\alpha'}] < \infty$ and the ergodic theorem. Since the above expression is predominant in the asymptotic theory and it seems hard to improve it, we obtain slow rates of convergence for the parameters σ and α if we apply the same estimation procedure as in the previous section. For this reason we require a different approach in the high frequency setting.

First of all, we give an explicit formula for the constant $m_{-p,k} = \mathbb{E}[|\Delta_{k,k} X|^{-p}]$, $p \in (0, 1/2)$, which has been introduced in Theorem 2.1. We recall that the random variable $\Delta_{k,k} X$ is symmetric α-stable with scale parameter $\sigma \|h_k\|_\alpha$. Consequently, applying the identity [14], Eq. (18), we conclude that

$$m_{-p,k} = \frac{(\sigma \|h_k\|_\alpha)^{-p}}{a_{-p}} \int_{\mathbb{R}} \exp\big(-|y|^\alpha\big)|y|^{-1+p} \, dy = \frac{2(\sigma \|h_k\|_\alpha)^{-p}}{\alpha a_{-p}} \Gamma(p/\alpha),$$

where the last equality follows by substitution $z = y^\alpha$ for $y > 0$. Now, we use the idea that has been originally proposed in [14] to identify the parameter α via power variation statistics. We consider $p, p' \in (0, 1/2)$, $p \ne p'$, and observe that

$$\frac{m_{-p',k}^p}{m_{-p,k}^{p'}} = \frac{(2/\alpha)^{p-p'} a_{-p}^{p'} \Gamma(p'/\alpha)^p}{a_{-p'}^p \Gamma(p/\alpha)^{p'}} =: \phi_{p,p'}(\alpha). \tag{4.4}$$

It has been shown in [14] that the mapping $\alpha \mapsto \phi_{p,p'}(\alpha)$ is invertible for any $p \neq p'$. Hence, we have $\alpha = \phi_{p,p'}^{-1}(m_{-p',k}^p / m_{-p,k}^{p'})$. Now, assuming that we know α and H (recall that the norm $\|h_k\|_\alpha$ depends on these parameters), we can recover the scale parameter σ via

$$\sigma = \left(\frac{\alpha a_{-p} m_{-p,k}}{2\Gamma(p/\alpha)} \right)^{-\frac{1}{p}} / \|h_k\|_\alpha.$$

Summarising the above identities we obtain the function $\overline{G} : (\mathbb{R}_+)^2 \times (0,1) \to \mathbb{R}^2$ such that

$$(\sigma, \alpha) = \overline{G}(m_{-p,k}, m_{-p',k}, H). \tag{4.5}$$

Next, we follow the same two-stage routine as in the previous section. We first compute $\widehat{H}_{\text{high}}(-p)_n = \widehat{H}_{\text{high}}(-p, 1)_n$ with $p \in (0, 1/2)$ and define the preliminary estimator of α by

$$\widehat{\alpha}_{\text{high}}^0(-p, -p')_n = \overline{G}_2 \big(V_{\text{high}}\big(f_{-p}, \widehat{H}_{\text{high}}(-p)_n\big)_n,$$
$$V_{\text{high}}\big(f_{-p'}, \widehat{H}_{\text{high}}(-p)_n\big)_n, \widehat{H}_{\text{high}}(-p)_n \big), \tag{4.6}$$

where the statistic $V_{\text{high}}(f_{-p}, \widehat{H}_{\text{high}}(-p)_n)_n$ refers to power variation introduced in (2.7) with $k = 1$ and with H replaced by $\widehat{H}_{\text{high}}(-p)_n$. Now, we define

$$\widehat{k}_{\text{high}} = \widehat{k}_{\text{high}}(-p, -p')_n = 2 + \lfloor \widehat{\alpha}_{\text{high}}^0(-p, -p')_n^{-1} \rfloor \tag{4.7}$$

and introduce the estimator

$$\big(\widetilde{\sigma}_{\text{high}}(\widehat{k}_{\text{high}}, -p, -p')_n, \widetilde{\alpha}_{\text{high}}(\widehat{k}_{\text{high}}, -p, -p')_n \big)$$
$$= \overline{G}\big(V_{\text{high}}(f_{-p}, \widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n; \widehat{k}_{\text{high}})_n,$$
$$V_{\text{high}}(f_{-p'}, \widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n; \widehat{k}_{\text{high}})_n, \widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n \big).$$

We again require a separate treatment of the cases $\alpha^{-1} \notin \mathbb{N}$ and $\alpha^{-1} \in \mathbb{N}$. We start with the first setting. When $H < k - 1/\alpha$ we consider the statistic $W(n)^{(1)} = (W(n)_1^{(1)}, W(n)_2^{(1)})$ associated with the power $-p$ and

$$k_1 = \widehat{k}_{\text{high}}, r_1 = 1 \quad \text{and} \quad k_2 = \widehat{k}_{\text{high}}, r_2 = 2.$$

Recall that $W(n)^{(1)} \xrightarrow{d} W^{(1)}$ according to Theorem 2.1. Now, similarly to Theorem 3.1, we define

$$\overline{B}_{\text{high}}^{\text{nor}}(-p, -p', k) := \nabla v_p\big(m_{-p,k}, 2^H m_{-p,k}\big)\big(W^{(1)}\big)^\star$$
$$\times \begin{pmatrix} \nabla \overline{G}_1(m_{-p,k}, m_{-p',k}, H)\big(-pm_{-p,k}, -p'm_{-p',k}, H\big)^\star \\ \nabla \overline{G}_2(m_{-p,k}, m_{-p',k}, H)\big(-pm_{-p,k}, -p'm_{-p',k}, H\big)^\star \\ 1 \end{pmatrix}, \tag{4.8}$$

where the function v_p has been introduced at (3.5). Our first result is the following theorem.

Theorem 4.4. *Let X be the linear fractional stable motion defined at (1.1). Assume that $p, p' \in (0, 1/2)$ and $\alpha^{-1} \notin \mathbb{N}$. Then we obtain the central limit theorem*

$$\begin{pmatrix} \sqrt{n}(\log n)^{-1}\big(\widetilde{\sigma}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \sigma\big) \\ \sqrt{n}(\log n)^{-1}\big(\widetilde{\alpha}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \alpha\big) \\ \sqrt{n}\big(\widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n - H\big) \end{pmatrix} \xrightarrow{d} \overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 2 + \lfloor \alpha^{-1} \rfloor\big).$$

Next, we treat the case $\alpha^{-1} \in \mathbb{N}$. For this purpose, whenever $H > k - 1/\alpha$, we introduce the notation $\overline{B}^{\text{sta}}_{\text{high}}(-p, -p', k)$ to denote the random variable at (4.8) where $W^{(1)}$ is replaced by $S^{(1)}$. We deduce the following result, which is the analogue of Theorem 4.3.

Theorem 4.5. *Let X be the linear fractional stable motion defined at (1.1). Assume that $p, p' \in (0, 1/2)$ and $\alpha^{-1} \in \mathbb{N}$.*

(i) *(Case $H < 1 - 1/\alpha$) Assume that $H < 1 - 1/\alpha$. Then we obtain*

$$\begin{pmatrix} \sqrt{n}(\log n)^{-1}\big(\widetilde{\sigma}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \sigma\big) \\ \sqrt{n}(\log n)^{-1}\big(\widetilde{\alpha}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \alpha\big) \\ \sqrt{n}\big(\widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n - H\big) \end{pmatrix} \xrightarrow{d} D^{\text{nor}}_{\text{high}},$$

where the probability distribution $D^{\text{nor}}_{\text{high}}$ on $\mathbb{R}^3$ is given by

$$D^{\text{nor}}_{\text{high}}(\cdot) = \mathbb{P}\big(\big\{\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 2 + \alpha^{-1}\big) \in \cdot\big\} \cap \big\{\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 1\big)_2 < 0\big\}\big)$$
$$+ \mathbb{P}\big(\big\{\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 1 + \alpha^{-1}\big) \in \cdot\big\} \cap \big\{\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 1\big)_2 > 0\big\}\big).$$

(ii) *(Case $H > 1 - 1/\alpha$) Assume that $H > 1 - 1/\alpha$. Then we obtain*

$$\begin{pmatrix} \sqrt{n}(\log n)^{-1}\big(\widetilde{\sigma}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \sigma\big) \\ \sqrt{n}(\log n)^{-1}\big(\widetilde{\alpha}_{\text{high}}\big(\widehat{k}_{\text{high}}, -p, -p'\big)_n - \alpha\big) \\ \sqrt{n}\big(\widehat{H}_{\text{high}}(-p, \widehat{k}_{\text{high}})_n - H\big) \end{pmatrix} \xrightarrow{d} D^{\text{sta}}_{\text{high}},$$

where the probability distribution $D^{\text{sta}}_{\text{high}}$ on $\mathbb{R}^3$ is given by

$$D^{\text{sta}}_{\text{high}}(\cdot) = \mathbb{P}\big(\overline{B}^{\text{sta}}_{\text{high}}\big(-p, -p', 1\big)_2 < 0\big)\mathbb{P}\big(\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 2 + \alpha^{-1}\big) \in \cdot\big)$$
$$+ \mathbb{P}\big(\overline{B}^{\text{sta}}_{\text{high}}\big(-p, -p', 1\big)_2 > 0\big)\mathbb{P}\big(\overline{B}^{\text{nor}}_{\text{high}}\big(-p, -p', 1 + \alpha^{-1}\big) \in \cdot\big).$$

Remark 4.2. We may use a similar decision rule as proposed in Section 4.1 to figure out whether part (i) or (ii) of Theorem 4.5 is applicable. Let $p_1, \ldots, p_4 \in (0, 1/2)$ be distinct real numbers. As in the previous subsection we have that

$$\overline{d}_n := -\frac{\log |\widehat{\alpha}^0_{\text{high}}(-p_1, -p_2)_n - \widehat{\alpha}^0_{\text{high}}(-p_3, -p_4)_n|}{\log(n)}$$

$$\xrightarrow{\mathbb{P}} \begin{cases} 1/2: & \text{if } H < 1 - 1/\alpha \\ 1 - 1/\big(1 + \alpha(1 - H)\big): & \text{if } H > 1 - 1/\alpha \end{cases}$$

We thus use Theorem 4.5(i) to perform statistical inference when

$$\overline{d}_n > 1/2 - \big(\log(n)\big)^{-1+\epsilon}.$$

5. A simulation study

In this section, we demonstrate the finite sample performance of our estimators based upon the theoretical results of Theorems 3.1, 4.1 and 4.4, where the latter two correspond to the setting $\alpha^{-1} \notin \mathbb{N}$ (we dispense with the numerical analysis associated with Theorems 4.3 and 4.5). We simulate high and low frequency observations of the linear fractional stable motion defined at (1.1) for $n = 100$, 1.000 and 10.000. To generate the paths of the linear fractional stable motion, we discretise the integral in (1.1) and use the fast Fourier transform as proposed in [29]. Whenever we use the statistics $V_{\text{high}}(f; k, r)_n$ and $V_{\text{low}}(f; k, r)_n$ introduced in (2.7), we multiply them by $(n - rk + 1)/n$ to account for the actual number of summands. Throughout the section, we set $t_1 = 1$ and $t_2 = 2$. We use 5000 repetitions to uncover the finite sample properties of our estimators. The asymptotic variances appearing in central limit theorems are rather hard to compute numerically due to slow rate of convergence of the involved infinite sums/integrals. Instead we perform Monte Carlo simulations to estimate them. More specifically, after computing an estimator $(\widehat{\sigma}, \widehat{\alpha}, \widehat{H})$, we generate sample paths of the process X with these parameters and calculate the associated empirical means/variances for all relevant statistics. These are used to construct empirical density functions.

We begin with the discussion of Theorem 3.1. Table 1 reports the bias and the standard deviation of the estimator of $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$ in high and low frequency settings, where we use the power $p = 0.4$ and the order $k = 2$. We observe that our estimators exhibit a rather convincing finite sample performance in both settings. As expected from the theoretical statements of Theorem 3.1, the estimators of the self-similarity parameter H exhibit similar finite sample properties in high and low frequency settings, while the performance of the low frequency estimators for the parameters σ and α is better than in the high frequency case. This is obviously a consequence of a slightly slower convergence rate in the high frequency setting. Figure 2 plots the empirical densities of the standardised estimators from Theorem 3.1 in comparison to the density of the standard normal distribution. As mentioned earlier, we use Monte Carlo simulations

Table 1. Bias/standard deviation of the estimators $(\widehat{\sigma}_{\text{low}}, \widehat{\alpha}_{\text{low}}, \widehat{H}_{\text{low}})$ and $(\widehat{\sigma}_{\text{high}}, \widehat{\alpha}_{\text{high}}, \widehat{H}_{\text{high}})$. We use $p = 0.4$ and $k = 2$, and the true parameter is $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$

n	$\widehat{\sigma}_{\text{low}}$	$\widehat{\alpha}_{\text{low}}$	$\widehat{H}_{\text{low}}$	$\widehat{\sigma}_{\text{high}}$	$\widehat{\alpha}_{\text{high}}$	$\widehat{H}_{\text{high}}$
100	$-0.024/0.06$	$-0.038/0.18$	$-0.05/0.12$	$0.06/0.18$	$-0.07/0.2$	$0.02/0.10$
1000	$-0.0008/0.02$	$0.012/0.068$	$-0.012/0.05$	$-0.001/0.12$	$0.015/0.07$	$-0.009/0.05$
10,000	$0.00014/0.006$	$0.0005/0.022$	$-0.005/0.016$	$-0.010/0.05$	$0.001/0.022$	$-0.005/0.016$
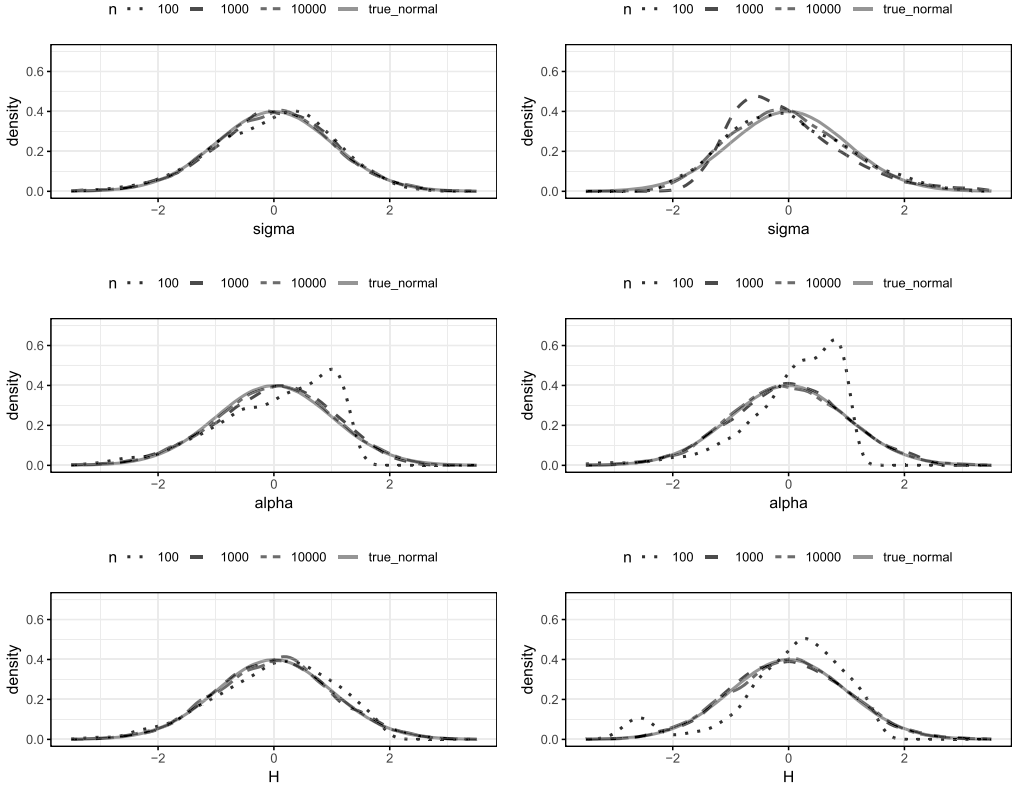
Figure 2. Empirical pdfs of $(\widehat{\sigma}, \widehat{\alpha}, \widehat{H})$ in high and low frequency settings. The right column corresponds to the high frequency case and the left one to the low frequency case. The true parameter is $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$, $k = 2$, $p = 0.4$.

to estimate the theoretical variances. We again observe a very good performance of estimators of the parameter H, while the numerical results for the estimators of σ and α are better in the low frequency case.

Another approach to estimation of the self-similarity parameter H is the log–log regression, which is a generalisation of our approach. The key idea is the observation that the convergence $V_{\text{low}}(f_p; k, r) \xrightarrow{\text{a.s.}} r^{pH} m_{p,k}$ for $p \in (0, \alpha)$ or $p \in (-1, 0)$ leads to the approximative identity

$$\log\big(V_{\text{low}}(f_p; k, r)_n\big) \approx \log(m_{p,k}) + pH \log(r), \quad r = 1, 2, \ldots, \overline{r}.$$

Note that the latter is a linear regression and the slope identifies the parameter H. Indeed, H can be estimated from low frequency data via

$$\widehat{H}_{\text{low}}^{\log} = \frac{\sum_{r=1}^{\overline{r}} (x_r - \overline{x})(y_r - \overline{y})}{p \sum_{r=1}^{\overline{r}} (x_r - \overline{x})^2},$$

Table 2. Bias/standard deviation of the regression-based estimator for H. Low frequency case. Here $p = -0.4$, $k = 2$ and $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$

n	$\overline{r}$	$\widehat{H}_{\mathrm{low}}^{\log}$
100	4	$-5.8\mathrm{e}{-}3/0.13$
1000	6	$-2.9\mathrm{e}{-}4/0.04$
10,000	9	$-1.5\mathrm{e}{-}4/0.013$

where $x_r = \log(r)$, $y_r = \log(V_{\mathrm{low}}(f_p; k, r)_n)$ and $\overline{x}$ (resp. $\overline{y}$) denotes the empirical mean of x_r's (resp. y_r's). Obviously, the asymptotic theory for the estimator $\widehat{H}_{\mathrm{low}}^{\log}$ can be directly deduced from Theorem 2.2; we leave the details to the reader. Instead we restrict our attention to the empirical performance of $\widehat{H}_{\mathrm{low}}^{\log}$. The next table demonstrates the finite sample bias/standard deviation of the estimator $\widehat{H}_{\mathrm{low}}^{\log}$ in the setting of Theorem 3.1 with $p = 0.4$, $k = 2$ and $\overline{r} = \lfloor \log(n) \rfloor$.

Comparing Tables 1 and 2, we observe that the standard deviations of $\widehat{H}_{\mathrm{low}}^{\log}$ and $\widehat{H}_{\mathrm{low}}$ are quite similar in all scenarios, but $\widehat{H}_{\mathrm{low}}^{\log}$ has a much lower bias.

Now, we turn our attention to the low frequency estimation discussed in Theorem 4.1. We use the power $p = -0.4$ and consider the true parameter $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$ and $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$. Observe that the first case corresponds to the setting of Theorem 3.1 and the second parameter corresponds to the discontinuous setting. The estimated order $\widehat{k}_{\mathrm{low}}$ is computed via (4.2). We also compute the log–log regression estimator $\widehat{H}_{\mathrm{low}}^{\log}$ based upon $\widehat{k}_{\mathrm{low}}$ and $\overline{r} = \lfloor \log(n) \rfloor$ in both scenarios. Table 3 displays the bias and standard deviation in the case $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$, while Table 4 demonstrates the numerical results in the case $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$. Finally, Table 5 shows the finite sample performance of $\widehat{H}_{\mathrm{low}}^{\log}$.

Comparing the simulation results of Theorems 3.1 and 4.1, we see that the finite sample performance of estimators σ and H in Theorem 4.1 is inferior; on the other hand $\widehat{H}_{\mathrm{low}}^{\log}$ outperforms $\widehat{H}_{\mathrm{low}}$ in terms of bias and and also in terms of variance for $n = 1000, 10,000$. This is not really surprising, since the methodology of Theorem 4.1 requires preliminary estimation of α and k, and hence leads to an accumulation of errors. On the other hand, the estimator of α is not as sensitive to preliminary estimation. Furthermore, in the setting of a fractional Brownian motion it is well known that low values of the parameter k give more efficient estimators. We conjecture that

Table 3. Bias/standard deviation of the estimator $(\widetilde{\sigma}_{\mathrm{low}}, \widetilde{\alpha}_{\mathrm{low}}, \widetilde{H}_{\mathrm{low}})$. Here $p = -0.4$, $\widehat{k}_{\mathrm{low}}$ is computed from (4.2) and $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$

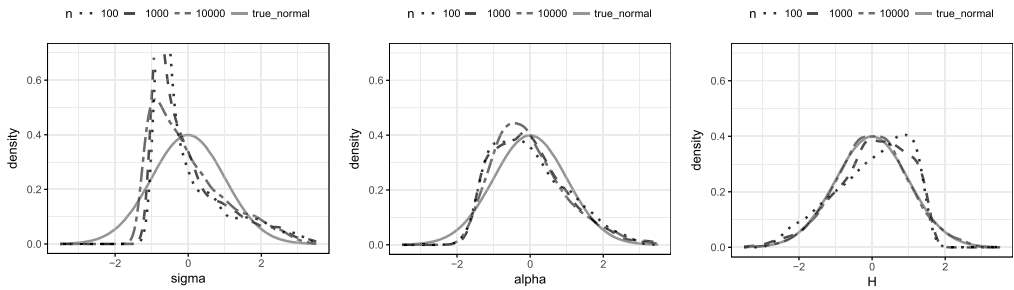| n	$\widetilde{\sigma}_{\mathrm{low}}$	$\widetilde{\alpha}_{\mathrm{low}}$	$\widetilde{H}_{\mathrm{low}}$
100	$-0.05/0.09$	$-0.031/0.18$	$-0.12/0.23$
1000	$-0.004/0.04$	$0.01/0.068$	$-0.018/0.12$
10,000	$0.0003/0.015$	$0.001/0.022$	$-0.003/0.05$

Table 4. Bias/standard deviation of the estimator $(\widetilde{\sigma}_{\mathrm{low}}, \widetilde{\alpha}_{\mathrm{low}}, \widetilde{H}_{\mathrm{low}})$. Here $p = -0.4$, $\widehat{k}_{\mathrm{low}}$ is computed from (4.2) and $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$

n	$\widetilde{\sigma}_{\mathrm{low}}$	$\widetilde{\alpha}_{\mathrm{low}}$	$\widetilde{H}_{\mathrm{low}}$
100	$-0.06/0.31$	$-0.003/0.41$	$-0.15/0.24$
1000	$-0.05/0.27$	$-0.08/0.31$	$0.003/0.13$
10,000	$0.03/0.26$	$0.008/0.27$	$0.04/0.05$

Table 5. Bias/standard deviation of the regression-based estimator for H with preliminary estimation of k. Low frequency case. Here $p = -0.4$, $\widehat{k}_{\mathrm{low}}$ is computed from (4.2). The left column corresponds to the triplet $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$, and the right one – to $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$

n	$\overline{r}$	$\widehat{H}_{\mathrm{low}}^{\log}$	$\widehat{H}_{\mathrm{low}}^{\log}$
100	4	$-4.9\mathrm{e}{-}03/0.22$	$0.018/0.3$
1000	6	$-1.1\mathrm{e}{-}03/0.06$	$0.01/0.09$
10,000	9	$-2.9\mathrm{e}{-}04/0.017$	$8\mathrm{e}{-}3/0.03$

a similar effect appears for linear fractional stable motions. This would explain the superiority of the results in Table 3 compared to those in Table 4, since $\lfloor \alpha^{-1} \rfloor + 2 = 2$ in the first setting while $\lfloor \alpha^{-1} \rfloor + 2 = 3$ in the second setting. Figures 3 and 4 show the empirical density functions, where the theoretical variances have been estimated via a Monte Carlo simulations. They confirm the better performance of the estimators in the continuous setting $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$. We also observe that the estimator of the parameter σ exhibits the worst finite sample properties in the setting $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$.
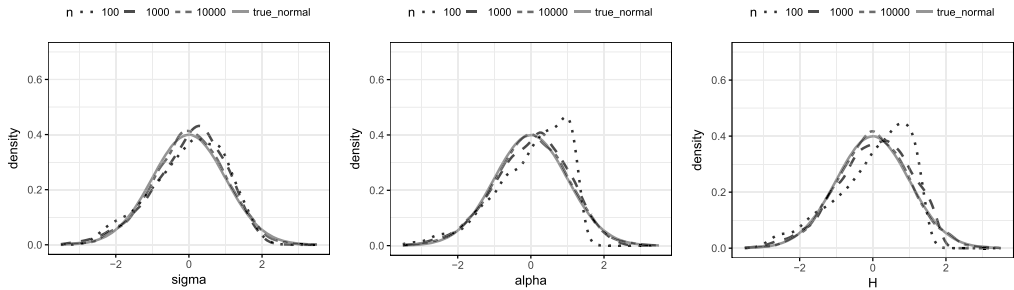
Figure 3. Empirical pdfs of $(\widetilde{\sigma}_{\mathrm{low}}, \widetilde{\alpha}_{\mathrm{low}}, \widetilde{H}_{\mathrm{low}})$. Here $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$ and $p = -0.4$.

Figure 4. Empirical pdfs of $(\widetilde{\sigma}_{\mathrm{low}}, \widetilde{\alpha}_{\mathrm{low}}, \widetilde{H}_{\mathrm{low}})$. Here $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$ and $p = -0.4$.

Finally, let us discuss the finite sample performance of the high frequency estimators from Theorem 4.4. We again consider two parameter settings $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$ and $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$, and we use $p = -0.4$ and $p' = -0.2$. The estimated order $\widehat{k}_{\mathrm{high}}$ is computed via (4.7). Tables 6 and 7 display the biases and standard deviations in both parameter settings. We observe that the estimators of the parameter σ have the worst performance and we only obtain reasonable results for $n = 10.000$. Similar conclusions can be drawn from Figures 5 and 6 that plot the empirical density functions. The bad performance of the estimator of σ in Theorem 4.4 is explained by the fact that we not only require a preliminary estimation step for our procedure, but we also need to estimate the parameters H and α first to obtain an estimator of σ. This leads to accumulation of finite sample errors, which results in large bias and variance for small n. To

Table 6. Bias/standard deviation of the estimator $(\widetilde{\sigma}_{\mathrm{high}}, \widetilde{\alpha}_{\mathrm{high}}, \widetilde{H}_{\mathrm{high}})$. Here $p = -0.4$, $p' = -0.2$ and $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$

n	$\widetilde{\sigma}_{\mathrm{high}}$	$\widetilde{\alpha}_{\mathrm{high}}$	$\widetilde{H}_{\mathrm{high}}$
100	60/1443	$-0.02/0.77$	0.23/0.33
1000	0.18/0.82	0.19/0.67	0.02/0.13
10,000	$-0.003/0.17$	0.052/0.26	$-0.003/0.05$

Table 7. Bias/standard deviation of the estimator $(\widetilde{\sigma}_{\mathrm{high}}, \widetilde{\alpha}_{\mathrm{high}}, \widetilde{H}_{\mathrm{high}})$. Here $p = -0.4$, $p' = -0.2$ and $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$

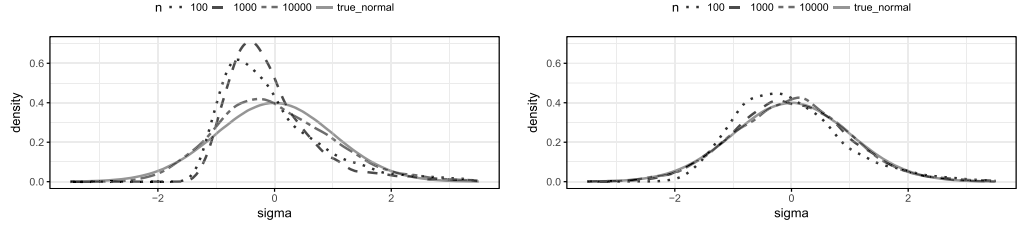| n	$\widetilde{\sigma}_{\mathrm{high}}$	$\widetilde{\alpha}_{\mathrm{high}}$	$\widetilde{H}_{\mathrm{high}}$
100	16/341	0.19/0.37	0.13/0.4
1000	0.103/1	0.02/0.09	0.06/0.16
10,000	$-0.11/0.12$	0.003/0.04	0.04/0.06
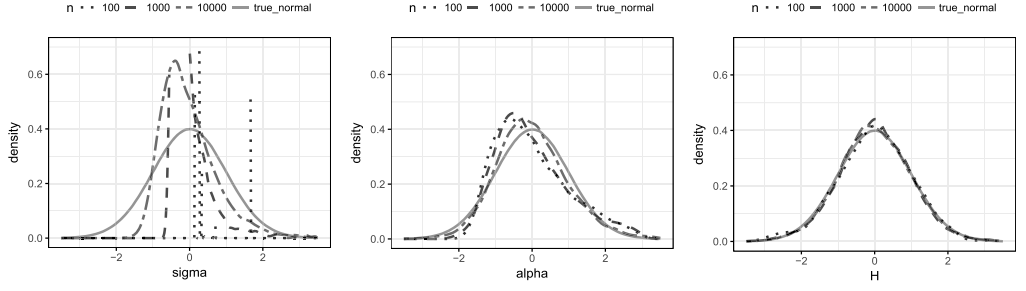
Figure 5. Empirical pdfs of $(\widetilde{\sigma}_{\text{high}}, \widetilde{\alpha}_{\text{high}}, \widetilde{H}_{\text{high}})$. Here $p = -0.4$, $p' = -0.2$ and $(\sigma, \alpha, H) = (0.3, 1.8, 0.8)$.
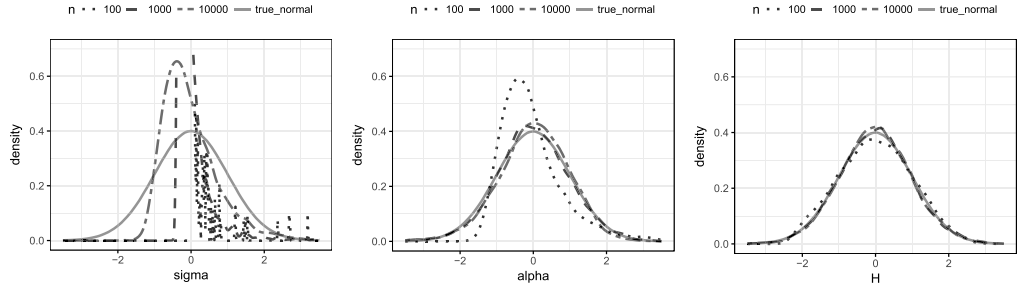
Figure 6. Empirical pdfs of $(\widetilde{\sigma}_{\text{high}}, \widetilde{\alpha}_{\text{high}}, \widetilde{H}_{\text{high}})$. Here $p = -0.4$, $p' = -0.2$ and $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$.

Figure 7. Empirical pdfs for σ (Left=Theorem 4.1, Right=Theorem 4.4) when the parameter $(\alpha, H) = (0.8, 0.8)$ is known. Here $\sigma = 0.3$, $p = -0.4$, $p' = -0.2$ and $k = 3$.

further highlight this issue, we have plotted the empirical densities for the estimators of σ from Theorems 4.1 and 4.4 in Figure 7 in the setting $(\sigma, \alpha, H) = (0.3, 0.8, 0.8)$ where the parameter (α, H) is assumed to be known. We observe a much better finite sample performance, which confirms that the bad finite sample properties of the estimator of σ are largely due to preliminary estimation of (α, H).

Acknowledgements

Supplementary Material

Supplement to "Estimation of the linear fractional stable motion" (DOI: 10.3150/19-BEJ1124SUPP; .pdf). Supplementary information.

References

[1] Aït-Sahalia, Y. and Jacob, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist*. **35** 355–392. MR2332279 https://doi.org/10.1214/009053606000001190

[2] Ayache, A., Cohen, S. and Lévy Véhel, J. (2000). The covariance structure of multifractional Brownian motion, with application to long range dependence (extended version). *ICASSP*, Refereed Conference Contribution.

[3] Ayache, A. and Hamonier, J. (2012). Linear fractional stable motion: A wavelet estimator of the α parameter. *Statist. Probab. Lett*. **82** 1569–1575. MR2930661 https://doi.org/10.1016/j.spl.2012.04.005

[4] Bardet, J.-M. and Surgailis, D. (2013). Nonparametric estimation of the local Hurst function of multifractional Gaussian processes. *Stochastic Process. Appl*. **123** 1004–1045. MR3005013 https://doi.org/10.1016/j.spa.2012.11.009

[5] Basse-O'Connor, A., Lachièze-Rey, R. and Podolskij, M. (2017). Power variation for a class of stationary increments Lévy driven moving averages. *Ann. Probab*. **45** 4477–4528. MR3737916 https://doi.org/10.1214/16-AOP1170

[6] Benassi, A., Cohen, S. and Istas, J. (2004). On roughness indices for fractional fields. *Bernoulli* **10** 357–373. MR2046778 https://doi.org/10.3150/bj/1082380223

[7] Bender, C., Lindner, A. and Schicks, M. (2012). Finite variation of fractional Lévy processes. *J. Theoret. Probab*. **25** 594–612. MR2914443 https://doi.org/10.1007/s10959-010-0339-y

[8] Bender, C. and Marquardt, T. (2008). Stochastic calculus for convoluted Lévy processes. *Bernoulli* **14** 499–518. MR2544099 https://doi.org/10.3150/07-BEJ115

[9] Brouste, A. and Fukasawa, M. (2018). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist*. **46** 2045–2061. MR3845010 https://doi.org/10.1214/17-AOS1611

[10] Cambanis, S., Hardin, C.D. Jr. and Weron, A. (1987). Ergodic properties of stationary stable processes. *Stochastic Process. Appl*. **24** 1–18. MR0883599 https://doi.org/10.1016/0304-4149(87)90024-X

[11] Cambanis, S. and Maejima, M. (1989). Two classes of self-similar stable processes with stationary increments. *Stochastic Process*. *Appl*. **32** 305–329. MR1014456 https://doi.org/10.1016/0304-4149(89)90082-3

[12] Coeurjolly, J.-F. and Istas, J. (2001). Cramèr–Rao bounds for fractional Brownian motions. *Statist*. *Probab. Lett*. **53** 435–447. MR1856169 https://doi.org/10.1016/S0167-7152(00)00197-8

[13] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann*. *Statist*. **17** 1749–1766. MR1026311 https://doi.org/10.1214/aos/1176347393

[14] Dang, T.T.N. and Istas, J. (2017). Estimation of the Hurst and the stability indices of a H-self-similar stable process. *Electron*. *J*. *Stat*. **11** 4103–4150. MR3715823 https://doi.org/10.1214/17-EJS1357

[15] Grahovac, D., Leonenko, N.N. and Taqqu, M.S. (2015). Scaling properties of the empirical structure function of linear fractional stable motion and estimation of its parameters. *J. Stat*. *Phys*. **158** 105–119. MR3296276 https://doi.org/10.1007/s10955-014-1126-4

[16] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **33** 407–436. MR1465796 https://doi.org/10.1016/S0246-0203(97)80099-4

[17] Lebovits, J. and Podolskij, M. (2017). Estimation of the global regularity of a multifractional Brownian motion. *Electron*. *J*. *Stat*. **11** 78–98. MR3597564 https://doi.org/10.1214/16-EJS1221

[18] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev*. **10** 422–437. MR0242239 https://doi.org/10.1137/1010093

[19] Peltier, R. and Lévy Véhel, J. (1995). Multifractional Brownian motion: Definition and preliminary results. Rapport de recherche de l'INRIA, number 2645.

[20] Pipiras, V. and Taqqu, M.S. (2002). The structure of self-similar stable mixed moving averages. *Ann*. *Probab*. **30** 898–932. MR1905860 https://doi.org/10.1214/aop/1023481011

[21] Pipiras, V. and Taqqu, M.S. (2003). Central limit theorems for partial sums of bounded functionals of infinite-variance moving averages. *Bernoulli* **9** 833–855. MR2047688 https://doi.org/10.3150/bj/1066418880

[22] Pipiras, V., Taqqu, M.S. and Abry, P. (2007). Bounds for the covariance of functions of infinite variance stable random variables with applications to central limit theorems and wavelet-based estimation. *Bernoulli* **13** 1091–1123. MR2364228 https://doi.org/10.3150/07-BEJ6143

[23] Rajput, B.S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probab*. *Theory Related Fields* **82** 451–487. MR1001524 https://doi.org/10.1007/BF00339998

[24] Resnick, S. and Greenwood, P. (1979). A bivariate stable characterization and domains of attraction. *J. Multivariate Anal*. **9** 206–221. MR0538402 https://doi.org/10.1016/0047-259X(79)90079-4

[25] Rosiński, J. (1995). On the structure of stationary stable processes. *Ann*. *Probab.* **23** 1163–1187. MR1349166

[26] Samorodnitsky, G. (2005). Null flows, positive flows and the structure of stationary symmetric stable processes. *Ann*. *Probab*. **33** 1782–1803. MR2165579 https://doi.org/10.1214/009117905000000305

[27] Mazur, S., Otryakhin, D. and Podolskij, M. (2020). Supplement to "Estimation of the linear fractional stable motion." https://doi.org/10.3150/19-BEJ1124SUPP.

[28] Stoev, S., Pipiras, V. and Taqqu, M.S. (2002). Estimation of the self-similarity parameter in linear fractional stable motion. *Signal Process*. **82** 1873–1901.

[29] Stoev, S. and Taqqu, M.S. (2004). Simulation methods for linear fractional stable motion and FARIMA using the fast Fourier transform. *Fractals* **12** 95–121. MR2052735 https://doi.org/10.1142/S0218348X04002379

[30] Stoev, S.A. and Taqqu, M.S. (2006). How rich is the class of multifractional Brownian motions? *Stochastic Process*. *Appl*. **116** 200–221. MR2197974 https://doi.org/10.1016/j.spa.2005.09.007