Asymptotically efficient estimation of a scale parameter in Gaussian time series and closed-form expressions for the Fisher information

T. Sabel* and J. Schmidt-Hieber†‡

Abstract

Mimicking the maximum likelihood estimator, we construct first order Cramer-Rao efficient and explicitly computable estimators for the scale parameter σ^2 in the model $Z_{i,n} = \sigma n^{-\beta} X_i + Y_i,\ i = 1, \ldots, n,\ \beta > 0$ with independent, stationary Gaussian processes $(X_i)_{i\in\mathbb{N}}$, $(Y_i)_{i\in\mathbb{N}}$, and $(X_i)_{i\in\mathbb{N}}$ exhibits possibly long-range dependence. In a second part, closed-form expressions for the asymptotic behavior of the corresponding Fisher information are derived. Our main finding is that depending on the behavior of the spectral densities at zero, the Fisher information has asymptotically two different scaling regimes, which are separated by a sharp phase transition. The most prominent example included in our analysis is the Fisher information for the scaling factor of a high-frequency sample of fractional Brownian motion under additive noise.

AMS 2010 Subject Classification: Primary 62F12; secondary 60G15, 60G22, 60G10.

Keywords: efficient estimation; fractional Brownian motion; Fisher information; general monotone sequence; regular variation; slowly varying functions; spectral density.

*Institut für Mathematische Stochastik, Universität Göttingen, Goldschmidtstr. 7, 37077 Göttingen, Germany.

†ENSAE, Paris, 3 Avenue Pierre Larousse, 92240 Malakoff, France.

‡For correspondence `Johannes.Schmidt.Hieber@ensae.fr`

1 Introduction

Let $(X_i)_{i\in\mathbb{N}}$ and $(Y_i)_{i\in\mathbb{N}}$ be independent Gaussian processes with known distribution. Suppose that we observe $\mathbf{Z} := \mathbf{Z}_n := (Z_{1,n}, \ldots, Z_{n,n})$ with

$$Z_{i,n} = \sigma n^{-\beta} X_i + Y_i, \quad i = 1, \ldots, n, \quad \text{and} \quad \beta, \sigma > 0. \tag{1.1}$$

In our framework, the parameter β is assumed to be known. We are interested in the case where $(X_i)_{i\in\mathbb{N}}$ is stationary and $(Y_i)_{i\in\mathbb{N}}$ is a noise process. Our theory includes white noise and increments of white noise as special cases for $(Y_i)_{i\in\mathbb{N}}$ (cf. Assumption 2). The problem, which we address in this work, is asymptotically optimal estimation of the scale parameter σ^2. In order to understand its asymptotic properties the key ingredient is knowledge of the Fisher information, for which closed-form expressions will be derived as well.

Our study is motivated by estimation of the variance σ^2 of a fractional Brownian motion (fBM) $(B_t^H)_{t\geq 0}$ at time points i/n, $i = 1, \ldots, n$ under additive Gaussian white noise (WN), i.e.

$$V_{i,n} := \sigma B_{i/n}^H + \tau \epsilon_i, \quad i = 1, \ldots, n. \tag{1.2}$$

Here, H refers to the Hurst index (or self-similarity parameter) and $(\epsilon_i)_i$ is a sequence of i.i.d. standard normal random variables. This model has attracted a lot of attention, recently (cf. Gloter and Hoffmann [14, 15] and for the special case $H = 1/2$, cf. Stein [26], Gloter and Jacod [16, 17], as well as Cai et *al.* [7]). Let us call it the fBM+WN model and note that the increment vector is of type (1.1) with $\beta = H$. This shows that models (1.1) and (1.2) coincide, if X_i and Y_i are chosen as the increments of $n^H(B_{i/n}^H - B_{(i-1)/n}^H)$ and $\tau(\epsilon_{i,n} - \epsilon_{i-1,n})$, with $\epsilon_{0,n} := 0$, respectively.

Estimation of σ^2 (and H) was discussed in slightly more general settings than the fBM+WN model by Gloter and Hoffmann [14, 15]. In these papers it was proven that for $H > \frac{1}{2}$ the optimal rate of convergence for σ^2 is $n^{-1/(4H+2)}$. More extensively studied and of particular interest is the case $H = \frac{1}{2}$, due to its applications to high-frequency modeling of stock returns. For this case, the asymptotic Fisher information is known to be $n^{\frac{1}{2}}(8\tau\sigma^3)^{-1}(1 + o(1))$ (cf. Gloter and Jacod [16, 17], and Cai et *al.* [7]). This result had a big impact as a benchmark for estimation of the integrated volatility (cf. Barndorff-Nielsen et *al.* [1], Podolskij and Vetter [21], Jacod et *al.* [19], and Zhang [28]) as well as for the asymptotic equivalence theorem by Reiß [23]. The fact that the multiplicative inverse of the asymptotic Fisher information is linear in τ and proportional to the cube of σ is surprising and requires further understanding.

The main contribution of our work to the existing literature is that for $0 < H < 1$, the Fisher information $I^n_{\sigma^2}$ for estimation of σ^2 in the fBM+WN model is given by

$$I^n_{\sigma^2} = n^{\frac{1}{2H+1}} \sigma^{-\frac{8H+2}{2H+1}} \tau^{-\frac{2}{2H+1}} c_H + o(n^{\frac{1}{2H+1}}), \tag{1.3}$$

where c_H is a constant only depending on H (for an explicit expression of c_H, cf. Corollary 1).

In general, we focus on the situation, where the Fisher information converges to infinity for $n \to \infty$, which corresponds to consistent estimation of σ^2. Note that in view of $n^{-\beta} X_i = O_p(n^{-\beta})$ and $Y_i = O_p(1)$ it is not clear at all that there are such situations. In fact, the rate at which the Fisher information tends to infinity can be rather unexpected. In a first place, one might guess that the optimal rate of convergence for estimation of the "parameter" $\sigma^2 n^{-2\beta}$ is $n^{-\frac{1}{2}}$ and hence the Fisher information of σ^2 should be of the order $n^{1-4\beta}$ (corresponding to the rate of convergence $n^{2\beta-1/2}$). However, this heuristic reasoning is in general not true and *better rates* can be obtained, as for instance in (1.3). Surprisingly, the asymptotic Fisher information has two different scaling regimes. In fact we will see that for any pair $(X_i)_i$ and $(Y_i)_i$ there is a positive characteristic $\lozenge$ such that (up to sub-polynomial factors) $I^n_{\sigma^2} \propto n^{1-\lozenge\beta}$ if $\lozenge < 4$ and $I^n_{\sigma^2} \propto n^{1-4\beta}$ if $\lozenge \geq 4$. The latter appears to be the same rate as in our heuristic argument above. Altogether, the different scaling behavior becomes visible as elbow effect in the convergence rate of σ^2. As a curious fact, let us mention that the spectral densities of the processes do not need to be known explicitly in order to compute the proposed estimator or the asymptotic Fisher information.

It is a classical result that if we observe a sample of a stationary Gaussian process with a spectral density $h(\theta, \cdot)$, the asymptotic Fisher information I^n_θ for estimation of a one-dimensional parameter θ is given by (cf. Davies [10] and Dzhaparidze [11] for the general case as well as Fox and Taqqu [12], Dahlhaus [9], Giraitis and Surgailis [13] for long-range dependent processes)

$$I^n_\theta = \frac{n}{2\pi} \int_0^\pi \big(\partial_\theta \log h(\theta, \lambda)\big)^2 d\lambda + o(n). \tag{1.4}$$

In Theorem 2, we prove that under fairly general conditions on $(X_i)_i$, a result of the type (1.4) holds for $\theta = \sigma^2$ in model (1.1). One should note that our setting is non-standard and not covered within the existing literature. In contrast, due to the factor $n^{-\beta}$, we cannot work with a fixed h but rather have to consider a sequence of spectral densities $(h_n)_n$ with degenerate limit. Furthermore, we are not in the classical parametric estimation setting, i.e. $I^n_{\sigma^2}$ may diverge with a rate which is much slower than n. As for example in (1.3), we need therefore to prove (1.4) with an approximation error which is of smaller order than $o(n)$. This in turn implies that very precise control on the (large) noise process $(Y_i)_i$ has to be imposed (cf. Assumption 2). Let us also mention that we cover both cases, long and short-range dependence of $(X_i)_i$. In particular this allows to treat model (1.2) for all $H \in (0,1)$.

The work is organized as follows. In Section 2.1, we construct the estimator and investigate its theoretical properties. Closed-form expressions for the Fisher information are derived in Section 2.2. In particular, we give some heuristic arguments why different scaling regimes appear. To illustrate the results some examples are provided in Section 3. Proofs are deferred to the appendix.

Notation: We write $\mathbf{X} := \mathbf{X}_n := (X_1, \ldots, X_n)$, $\mathbf{Y} := \mathbf{Y}_n := (Y_1, \ldots, Y_n)$ and $\mathbf{Z} := \mathbf{Z}_n := (Z_{1,n}, \ldots, Z_{n,n})$. For two sequence $(a_k)_k$ and $(b_k)_k$, we say that $a_k \sim b_k$ iff $\lim_{k\to\infty} a_k/b_k = 1$. Similar, for two functions g_1 and g_2, we write $g_1(\lambda) \sim g_2(\lambda)$ (for $\lambda \downarrow 0$) iff $\lim_{\lambda\downarrow 0} g_1(\lambda)/g_2(\lambda) = 1$.

2 Main results

Let $\mathbf{U} = \mathbf{U}_n$ be an n-dimensional, centered Gaussian vector with positive definite covariance matrix Σ_θ, depending on a one-dimensional parameter $\theta \in \mathbb{R}$. The log-likelihood function is

$$L(u|\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\Sigma_\theta|) - \frac{1}{2}u^t\Sigma_\theta^{-1}u,$$

with $|\Sigma_\theta|$ the determinant of Σ_θ. Let $\partial_\theta\Sigma_\theta$ denote the entrywise derivative of Σ_θ with respect to θ (which we assume to exist). Since $\partial_\theta \log(|\Sigma_\theta|) = \operatorname{tr}(\Sigma_\theta^{-1}\partial_\theta\Sigma_\theta)$ and $\partial_\theta\Sigma_\theta^{-1} = -\Sigma_\theta^{-1}(\partial_\theta\Sigma_\theta)\Sigma_\theta^{-1}$, we find for the score function

$$\dot{L}(\mathbf{U}|\theta) := \partial_\theta L(\mathbf{U}|\theta) = -\frac{1}{2}\operatorname{tr}(\Sigma_\theta^{-1}\partial_\theta\Sigma_\theta) + \frac{1}{2}\mathbf{U}^t\Sigma_\theta^{-1}(\partial_\theta\Sigma_\theta)\Sigma_\theta^{-1}\mathbf{U}.$$

In distribution, $\mathbf{U} = \Sigma_\theta^{1/2}\xi$ for an n-dimensional standard normal vector ξ. Together with some algebra this shows that the Fisher information for θ is $I_\theta^n = \frac{1}{2}\operatorname{tr}\left([(\partial_\theta\Sigma_\theta)\Sigma_\theta^{-1}]^2\right)$ (cf. also Porat and Friedlander [22]). In particular, for model (1.1) we obtain

$$I_{\sigma^2}^n = \tfrac{1}{2}\operatorname{tr}\left([n^{-2\beta}\operatorname{Cov}(\mathbf{X})\operatorname{Cov}(\mathbf{Z})^{-1}]^2\right). \tag{2.1}$$

To simplify the notation, we will view $I_{\sigma^2}^n$ in the following always as a sequence in n.

2.1 An asymptotically Cramer-Rao efficient estimator

In this section, we construct an explicitly computable estimator which mimics the MLE. Furthermore, we prove that the mean squared error (MSE) of this estimator is first order optimal (cf. Theorem 1). All the results in this section, work under fairly general conditions.

In fact, we only require that $\operatorname{Cov}(\mathbf{X})$ and $\operatorname{Cov}(\mathbf{Y})$ are known and positive definite for all n as well as divergence of the Fisher information. In particular, we neither have to impose stationarity on $(X_i)_i$ or $(Y_i)_i$ nor do we assume that $\operatorname{Cov}(\mathbf{X})$ and $\operatorname{Cov}(\mathbf{Y})$ have the same set of eigenvectors.

The construction will be done in several steps. First, we can find $n \times n$ matrices A and D such that $\operatorname{Cov}(\mathbf{Y}) = A^t A$, $D^t D = \mathrm{id}_n$ (the identity), and D diagonalizes $(A^{-1})^t \operatorname{Cov}(\mathbf{X}) A^{-1}$. Hence, $\Lambda := (A^{-1}D)^t \operatorname{Cov}(\mathbf{X})(A^{-1}D)$ is diagonal and the diagonal entries are denoted by $\lambda_1, \ldots, \lambda_n$. The maximum likelihood equation in the transformed model $(\widetilde{Z}_{1,n}, \ldots, \widetilde{Z}_{n,n})^t := (A^{-1}D)^t \mathbf{Z}$ motivates to consider the oracle estimator

$$\widehat{\sigma}^2_{\text{oracle}} := (2I^n_{\sigma^2})^{-1} \sum_{i=1}^n \frac{\lambda_i n^{-2\beta}(\widetilde{Z}^2_{i,n} - 1)}{(\sigma^2 n^{-2\beta}\lambda_i + 1)^2} = \sigma^2 + (2I^n_{\sigma^2})^{-1} \sum_{i=1}^n \frac{\lambda_i n^{-2\beta}(\widetilde{Z}^2_{i,n} - \mathbb{E}\widetilde{Z}^2_{i,n})}{(\sigma^2 n^{-2\beta}\lambda_i + 1)^2}. \tag{2.2}$$

To verify the equality one should note that by rewriting (2.1)

$$I^n_{\sigma^2} = \frac{1}{2} \sum_{i=1}^n \frac{\lambda_i^2 n^{-4\beta}}{(\sigma^2 \lambda_i n^{-2\beta} + 1)^2}.$$

Observe that the oracle estimator (2.2) is unbiased and attains the Cramer-Rao bound since $\operatorname{Var}(\widehat{\sigma}^2_{\text{oracle}}) = (I^n_{\sigma^2})^{-1}$. Oracle estimators depend on the unknown quantities itself, and are thus not computable. Below, we derive a simple construction for a statistical estimator which mimics $\widehat{\sigma}^2_{\text{oracle}}$ and is asymptotically sharp. Similar as in [7], we use a sample splitting technique. First, we take a small part of the data in the transformed model, which are used for a preliminary estimate, say $\widetilde{\sigma}^2$, of σ^2. In a second step, we plug $\widetilde{\sigma}^2$ into (2.2). Discarding all indices, which were already used for $\widetilde{\sigma}^2$ gives an estimator, which as we show has asymptotically the same properties as $\widehat{\sigma}^2_{\text{oracle}}$. This implies then the first order Cramer-Rao efficiency.

The next lemma ensures that sample splitting can be done.

Lemma 1. *For $u > 0$ and $B \subset \{1, \ldots, n\}$, let*

$$I^B_u := \frac{1}{2} \sum_{i \in B} \frac{\lambda_i^2 n^{-4\beta}}{(u\lambda_i n^{-2\beta} + 1)^2}.$$

Then there is a sequence of index sets $(A_n)_n$ with $A_n \subset \{1, \ldots, n\}$ such that

$$I_1^{A_n} \to \infty \quad \textit{and} \quad \frac{I_1^{A_n}}{I_1^n} \to 0 \quad \textit{as} \quad n \to \infty.$$

Proof. Note that $0 \leq c_{i,n} := \lambda_i^2 n^{-4\beta}(n^{-2\beta}\lambda_i + 1)^{-2} < 1$. Consider the partial sums $S_k = \sum_{i=1}^k c_{i,n}$. and observe that $S_{k+1} - S_k \leq 1$ as well as $S_n = I_1^n \to \infty$. Therefore, we can find

$k^*(n)$, s.t. $\sqrt{I_1^n} \leq S_{k^*(n)} \leq \sqrt{I_1^n} + 1$. The result follows with $A_n = \{1, \dots, k^*(n)\}$. $\square$

Throughout this section let $(A_n)_n$ be as in the previous lemma and pick a sequence $(\delta_n)_n$, satisfying $\delta_n \leq 1$, $\delta_n \to 0$, and $\delta_n^4 I_1^{A_n} \to \infty$. A possible choice is $\delta_n = (I_1^{A_n})^{-1/8}$. Observe that

$$V := (2I_1^{A_n})^{-1} \sum_{i \in A_n} \frac{\lambda_i n^{-2\beta}(\widetilde{Z}_{i,n}^2 - 1)}{(n^{-2\beta}\lambda_i + 1)^2} = \sigma^2 + (2I_1^{A_n})^{-1} \sum_{i \in A_n} \frac{\lambda_i n^{-2\beta}(\widetilde{Z}_{i,n}^2 - \mathbb{E}\widetilde{Z}_{i,n}^2)}{(n^{-2\beta}\lambda_i + 1)^2}.$$

has expectation σ^2 and variance bounded by $(\sigma^4 \vee 1)(I_1^{A_n})^{-1}$. Now, we define the preliminary estimator $\widetilde{\sigma}^2$, as the truncated version of V,

$$\widetilde{\sigma}^2 := (V \vee \delta_n) \wedge \delta_n^{-1}. \tag{2.3}$$

This allows us to construct the final estimator $\widehat{\sigma}_n^2$ for σ^2. Let $A_n^c = \{1, \dots, n\} \setminus A_n$ and set

$$\widehat{\sigma}_n^2 := (2I_{\widetilde{\sigma}^2}^{A_n^c})^{-1} \sum_{i \in A_n^c} \frac{\lambda_i n^{-2\beta}(\widetilde{Z}_{i,n}^2 - 1)}{(\widetilde{\sigma}^2 n^{-2\beta}\lambda_i + 1)^2} = \sigma^2 + (2I_{\widetilde{\sigma}^2}^{A_n^c})^{-1} \sum_{i \in A_n^c} \frac{\lambda_i n^{-2\beta}(\widetilde{Z}_{i,n}^2 - \mathbb{E}\widetilde{Z}_{i,n}^2)}{(\widetilde{\sigma}^2 n^{-2\beta}\lambda_i + 1)^2}. \tag{2.4}$$

One should note the similarity to the oracle $\widehat{\sigma}_{\mathrm{oracle}}^2$ as introduced in (2.2). As the following theorem shows, $\widehat{\sigma}_n^2$ has in fact the same asymptotic MSE as the oracle, implying Cramer-Rao efficiency.

Theorem 1. *Suppose that the Fisher information diverges and* $\mathrm{Cov}(\mathbf{Y})$ *is positive definite. The estimator* $\widehat{\sigma}^2$ *defined in* (2.4) *attains the Cramer-Rao bound asymptotically over every compact set, not containing zero, i.e. for* $0 < \sigma_{\min} < \sigma_{\max} < \infty$,

$$\lim_{n \to \infty} \sup_{\sigma \in [\sigma_{\min}, \sigma_{\max}]} I_{\sigma^2}^n \cdot \mathrm{MSE}(\widehat{\sigma}_n^2) = 1.$$

2.2 Closed-form expressions for the Fisher information

So far we have seen that there are estimators which are asymptotically Cramer-Rao efficient. However, in order to get some understanding of the asymptotics, we need to study the behavior of the Fisher information. In this section, we derive explicit closed-form expressions.

To state the results, some definitions, in particular from regular variation theory, are unavoidable. For the notion of quasi-monotone and slowly varying functions see the monograph [4]. A positive sequence $(r_j)_j$ is called O-regularly varying if for any $\lambda > 1$,

$$0 < \underline{\lim}_{n \to \infty} \frac{r_{\lfloor \lambda n \rfloor}}{r_n} < \overline{\lim}_{n \to \infty} \frac{r_{\lfloor \lambda n \rfloor}}{r_n} < \infty,$$

with $\lfloor \cdot \rfloor$ the Gauss bracket. For real sequences $(a_j)_j$, O-regularly varying quasi-monotonicity is equivalent to the existence of a positive, non-decreasing, and O-regularly varying sequence $(r_j)_j$ such that the sequence $(a_j/r_j)_j$ is decreasing. We say that a sequence $(a_j)_j$ is general monotone if there are finite constants C, J_0, such that for any positive integer $J \geq J_0$, $\sum_{j=J}^{2J-1} |a_{j+1} - a_j| \leq C|a_J|$. The class of general monotone sequences will be denoted by GM. It was introduced and studied recently by Belov [2] and Tikhonov [27]. To simplify some arguments, we have relaxed the original definition slightly by introducing J_0 (this does not cause any trouble and all results on GM sequences can be transfered with obvious changes). In particular GM is fairly general in the sense that it includes all well-known generalizations of monotone sequences, such as quasi-monotonicity, regularly quasi-monotonicity, O-regularly-quasi-monotonicity and sequences of rest bounded variation.

In order to deal with boundary problems (cf. the second example in Section 3), we assume that $(X_i)_i$ is only approximately stationary in the following sense.

Assumption 1 (Assumptions on X)**.** *Suppose that there is a stationary process $(X_i')_i$ and a process $(R_i)_i$ such that in distribution*

$$X_i = X_i' + R_i, \quad \textit{for all } i \in \mathbb{N}$$

and $(X_i')_i$ and $(R_i)_i$ have the following properties.

(i) There is a positive and quasi-monotone slowly varying function $\ell : (0, \infty) \to \mathbb{R}^+$, satisfying

$$\frac{\ell(x\ell^\kappa(x))}{\ell(x)} \to 1, \ x \to \infty \quad \textit{for all} \ \ \kappa \in \mathbb{R} \tag{2.5}$$

such that for an index $\alpha \in (-1/2, 1/2)$,

$$\gamma_k := \operatorname{Cov}(X_1', X_{1+k}') \sim \operatorname{sign}(-\alpha) k^{-2\alpha-1} \ell(k) \quad \textit{as} \ \ k \to \infty. \tag{2.6}$$

(ii) With $\mathbf{X}_n' := (X_1', \ldots, X_n')$, $\mathbf{R}_n := (R_1, \ldots, R_n)$ and $\| \cdot \|_2$ the Frobenius norm, we have the uniform bound

$$\sup_n \| \operatorname{Cov}(\mathbf{X}_n', \mathbf{R}_n)\|_2 + \| \operatorname{Cov}(\mathbf{R}_n)\|_2 < \infty. \tag{2.7}$$

Throughout the following, we interpret the autocovariance $(\gamma_k)_{k \in \mathbb{Z}}$ as a sequence on $\mathbb{Z}$ via $\gamma_k = \gamma_{-k}$. An example for a (quasi-)monotone slowly varying ℓ is the logarithm $\log(1 + \cdot)$. However, (2.5) does not hold for every slowly varying function. A stronger condition, which implies (2.5) is

$$\lim_{\lambda\to 0}\left(\frac{\ell(a\lambda)}{\ell(\lambda)}-1\right)\log(\lambda)=0 \quad \text{for an } a>1$$

(cf. Theorem 1 in [5]). As a consequence (2.5) holds whenever $\lim_{\lambda\to\infty}\ell(\lambda)\in(0,\infty)$. Moreover, if it is true for ℓ then also for ℓ^μ, $\mu\geq 0$.

The $n\times n$ matrix Δ denotes the backward difference operator, i.e.

$$\Delta=\begin{pmatrix}1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1\end{pmatrix}. \tag{2.8}$$

For a vector $v=(v_1,\ldots,v_n)^t$, $\Delta v=(v_1-v_0,v_2-v_1,\ldots,v_n-v_{n-1})^t$ with $v_0:=0$ is the backwards difference process. Furthermore, the transposed matrix Δ^t is the negative forward difference operator $\Delta^t v=-(v_2-v_1,v_3-v_2,\ldots,v_{n+1}-v_n)^t$ with $v_{n+1}:=0$. We assume that for a non-negative integer K, the process $\mathbf{Y}$ is generated by taking the K-th finite difference of a white noise process (alternating between forward and backward differences).

Assumption 2. *Given a non-negative integer K and $\tau>0$, assume that $\mathbf{Y}$ is an n-dimensional, centered Gaussian random vector with covariance matrix $\tau^2\left(\Delta\Delta^t\right)^K$ or $\tau^2\left(\Delta^t\Delta\right)^K$.*

Assumption 2 imposes in fact a very serious restriction, but seems to be somehow unavoidable in order to prove the statement (cf. also the discussion in the introduction). Our results could be worked out under more general boundary conditions of the difference operator, of course. It is indeed sufficient that $\mathrm{Cov}(\mathbf{Y})$ can be perfectly diagonalized by a discrete since or cosine transform. However, since the assumption above is somehow the most natural one and allows to treat the fBM+WN model, we will restrict ourselves to it for sake of simplicity.

Let throughout the paper $f=\sum_{k=-\infty}^{\infty}\gamma_k\cos(k\cdot)$ denote the spectral density of $(X_i')_{i\in\mathbb{N}}$. Although $\mathbf{X}$ and $\mathbf{Y}$ are stationary only up to boundary values, we will refer occasionally to

$$f,\quad 4^K\tau^2\sin^{2K}\left(\tfrac{\cdot}{2}\right),\quad \text{and}\quad h_n=\sigma^2 n^{-2\beta}f+4^K\tau^2\sin^{2K}\left(\tfrac{\cdot}{2}\right) \tag{2.9}$$

as the spectral density of the processes $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$, respectively. Because of the imposed independence of $(X_i)_i$ and $(Y_i)_i$, h_n is the sum of the spectral densities of $\mathbf{X}$ and $\mathbf{Y}$.

Define

$$r_n:=\begin{cases} n^{1-\frac{\beta}{K-\alpha}}\left(\ell(n^{\frac{\beta}{K-\alpha}})\right)^{\frac{1}{2K-2\alpha}}+n^{1-4\beta}, & \text{if } K-\alpha\neq\frac{1}{4},\\ n^{1-4\beta}\log(n)\ell^2(n), & \text{if } K-\alpha=\frac{1}{4}.\end{cases} \tag{2.10}$$

Now, we are ready to state the main results of the paper. Surprisingly, it turns out that the rates depend on K and α only through their (inverted) difference, i.e. the problem is characterized by

$$\Diamond := \frac{1}{K-\alpha}.$$

Theorem 2. *Work in model* (1.1) *under Assumptions 1 and 2 and suppose that* $K-\alpha > \max\{\beta, (4\alpha+1)\beta, 1/4\}$. *Further assume that either*

1. $\alpha \in (0, 1/2)$, $(\gamma_k)_k$ *is O-regularly varying quasi-monotone, and* $\sum_{k=-\infty}^{\infty} \gamma_k = 0$, *or*
2. $\alpha \in (-1/4, 0)$ *and* $(\gamma_k)_k \in \mathrm{GM}$, *or*
3. $\alpha \in (-1/2, -1/4)$ *and there exists a constant* C_1, *such that for any* $p \in \mathbb{N}$, $|\gamma_{p+1} - \gamma_p| \leq C_1 |\gamma_p| p^{-1}$.

Then, the Fisher information of σ^2 *based on* n *observations is given by*

$$I_{\sigma^2}^n = \frac{n^{1-4\beta}}{2\pi} \int_0^\pi \frac{f^2(\lambda)}{h_n^2(\lambda)} d\lambda (1+o(1)) + o(r_n). \tag{2.11}$$

If the condition $K-\alpha > \max\{\beta, (4\alpha+1)\beta, 1/4\}$ *is replaced by the weaker assumption* $K-\alpha > \max\{\beta, (4\alpha+1)\beta\}$, *imposing additionally* $\log(n)\ell^2(n) \to \infty$ *in the critical case* $K-\alpha = 1/4$, *then* (2.11) *holds as well, provided there exists a constant* C_f, *such that*

$$|f(\lambda) - f(\mu)| \leq C_f \lambda^{2\alpha-2} |\lambda - \mu|, \quad \textit{for all} \ \ 0 < \lambda \leq \mu \leq \pi. \tag{2.12}$$

Let

$$C(\Diamond, \alpha) := \frac{(2-\Diamond)\Diamond}{8\sin(\Diamond\frac{\pi}{2})} \big(2\,\mathrm{sign}(-\alpha)\Gamma(-2\alpha)\cos(\pi\alpha)\big)^{\frac{1}{2}\Diamond}.$$

Theorem 3. *Under the assumptions of Theorem 2, the asymptotic Fisher information is explicitly given by*

$$I_{\sigma^2}^n \sim n^{1-\Diamond\beta} \big(\ell(n^{\Diamond\beta})\big)^{\Diamond/2} \sigma^{\Diamond-4} \tau^{-\Diamond} C(\Diamond, \alpha), \quad \textit{if} \ \ \Diamond < 4, \tag{2.13}$$

and

$$I_{\sigma^2}^n \sim \frac{n^{1-4\beta}}{2\pi\tau^4} \int_0^\pi f^2(\lambda) d\lambda = \frac{n^{1-4\beta}}{2\tau^4} \sum_{k=-\infty}^{\infty} \gamma_k^2, \quad \textit{if} \ \ \Diamond > 4. \tag{2.14}$$

For $\ell = |\log^\rho(\cdot)|, \ \rho > -1/2$,

$$I_{\sigma^2}^n \sim n^{1-4\beta} (\log n)^{2\rho+1} \ \tau^{-4} \frac{(4\beta)^{2\rho+1}}{2\rho+1}, \qquad \textit{if} \ \ \Diamond = 4. \tag{2.15}$$

Corollary 1. *In the fBM+WN model* (1.2) *with* $H \in (0,1)$, *it holds that*

$$\Diamond = \frac{2}{2H+1} < 4$$

and the asymptotic Fisher information for σ^2 *is given by* (1.3) *with*

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

where $\Gamma(\cdot)$ *denotes the Gamma function.*

Proofs of the statements are deferred to Appendix B. Let us conclude the section by some comments on Theorems 2 and 3.

First, observe that (2.11) is of the type (1.4) for $\theta = \sigma^2$. The surprising fact is that one can compute the integral $\int_0^\pi f^2(\lambda)/h_n^2(\lambda)d\lambda$ obtaining expressions which for $\Diamond < 4$ do not depend on f anymore. Let us shortly explain this. Suppose that Assumption 1 holds with $\ell = 1$. By classical results, we can conclude that for $\lambda \downarrow 0$, $f(\lambda) \sim C_\alpha \lambda^{2\alpha}$ and $h_n(\lambda) \sim C_\alpha \sigma^2 n^{-2\beta}\lambda^{2\alpha} + \tau^2\lambda^{4K}$ with $C_\alpha = 2\,\mathrm{sign}(-\alpha)\Gamma(-2\alpha)\cos(\pi\alpha)$. Next, observe that for small λ, $f^2(\lambda)/h_n^2(\lambda) \approx n^{4\beta}$ whereas for large values the integrand behaves like $\lambda^{4\alpha-4K}$. Now, $\Diamond \leq 4$ is equivalent to $4\alpha - 4K < -1$ and in this case the integral will be determined in first order by $f^2(\lambda)/h_n^2(\lambda)$ for small λ. Therefore, one expects that f and h_n can be replaced by their corresponding small value approximations $C_\alpha\lambda^{2\alpha}$ and $C_\alpha\sigma^2 n^{-2\beta}\lambda^{2\alpha} + \tau^2\lambda^{4K}$, respectively and

$$\int_0^\pi \tfrac{f^2(\lambda)}{h_n^2(\lambda)} d\lambda \approx \sigma^{-4} n^{4\beta} \int_0^\pi \left(1 + C_\alpha^{-1}\sigma^{-2}\tau^2 n^{2\beta}\lambda^{2K-2\alpha}\right)^{-2} d\lambda.$$

The r.h.s. can be explicitly solved and does not depend on f. In contrast to that, for $\Diamond > 4$, the Fisher information depends also asymptotically on the whole spectrum $(0,\pi]$. This is why we need the additional assumption (2.12) which controls the continuity of f globally.

The phase transition for $\Diamond = 4$ does not only affect the asymptotic constant but also leads to an elbow phenomenon in the rate of convergence for estimation of σ^2. If $\Diamond \leq 4$ the optimal rate (neglecting sub-polynomial factors in the following) is $n^{\frac{1}{2}\Diamond\beta - \frac{1}{2}}$ whereas for $\Diamond > 4$, it is $n^{2\beta - \frac{1}{2}}$. The latter only depends on β. Typically, if in an estimation problem an elbow effect occurs there are different sources of errors which cannot be balanced and therefore the best attainable rate is given by the maximum of the single error rates. However, in our situation the optimal rate turns out to be the *minimum*, more precisely it is $\min(n^{\frac{1}{2}\Diamond\beta - \frac{1}{2}}, n^{2\beta - \frac{1}{2}})$.

Let us also shortly remark on the dependence of the asymptotic Fisher information on the scaling coefficient σ. Write $p = \Diamond/4$ and $\theta = \sigma^2$. Choosing $Q_n(\tau, \alpha, \Diamond)$ appropriately, we find that the Fisher information for θ observing $U \sim \mathcal{N}(\theta^p, Q_n)$ if $p < 1$ and $U \sim \mathcal{N}(\theta, Q_n)$ if

$p \geq 1$ coincides in first order with (2.13)-(2.15) (standardizing X such that $\sum_{k=-\infty}^{\infty} \gamma_k^2 = 1$ if $p > 1$). Hence, in an asymptotic sense our original statistical estimation problem is related to a Gaussian shift model where we want to estimate the p-th power ($p \leq 1$) of the mean value.

Note that our results also cover the case $\alpha \in (-1/2, -1/4)$ for which the autocovariance function is not (square) summable. In fact the proof turns out to be very subtle and requires quite restrictive conditions. In particular, we have to impose an assumption on the increments of the autocovariance which is much stronger than GM.

One should also note that in the critical case $\diamondsuit = 4$ an additional log-factor appears in the rate of convergence. In Theorem 3, we have restricted ourselves to the (most important) case where ℓ is a power of the logarithm, which allows to evaluate the asymptotic Fisher information in closed form. However, from the proof one can follow a slightly more general version, namely that under the assumptions of Theorem 2 (in particular $\log(n)\ell^2(n) \to \infty$) and with $q_n := n^{-4\beta}\ell^2(n^{4\beta})$,

$$I_{\sigma^2}^n = n^{1-4\beta}\tau^{-4} \int_{q_n}^1 \ell^2\big(\tfrac{1}{\lambda}\big) \frac{d\lambda}{\lambda}(1 + o(1)) + o(nq_n \log(n)).$$

Theorem 2 and Theorem 3 are derived for all $\alpha \in (-1/2, 1/2), \alpha \neq 0$. The case $\alpha = 0$ is indeed special since as $\alpha \to 0$, $C(\diamondsuit, \alpha) \to \infty$. However, in the fBM+WN model (1.2) (recall that $H = 1/2 - \alpha$) this phenomenon does not play a role because of $\ell(\lambda)$ =const.= $H|2H - 1|$, which converges to zero (fast enough) as $H \to 1/2$. This explains why c_H is continuous for all $H \in (0, 1)$, whereas $\alpha \mapsto C(\diamondsuit, \alpha)$ is not.

Besides the classical case $H = 1/2$, which was mentioned already in the introduction, one can easily simplify the asymptotic Fisher information in the fBM+WN model (1.2) for $H \in \{1/4, 3/4\}$. Indeed, as a consequence of Corollary 1, we obtain for the multiplicative inverse (which is the asymptotic variance of our estimator)

$$\begin{aligned}
H = 1/4: &\quad (I_{\sigma^2}^n)^{-1} \sim \frac{27}{\sqrt{3}\ \pi^{1/3}}\sigma^{8/3}\tau^{4/3}n^{-2/3} \approx 10.64\sigma^{8/3}\tau^{4/3}n^{-2/3}, \\
H = 1/2: &\quad (I_{\sigma^2}^n)^{-1} \sim 8\sigma^3\tau n^{-1/2}, \\
H = 3/4: &\quad (I_{\sigma^2}^n)^{-1} = \frac{25\sqrt{5+\sqrt{5}}}{\sqrt{2}\ 3^{7/5}\ \pi^{1/5}}\sigma^{16/5}\tau^{4/5}n^{-2/5} \approx 8.12\sigma^{16/5}\tau^{4/5}n^{-2/5}.
\end{aligned}$$

Finally, one should note that the elbow effect observed in Gloter and Jacod [16] does not relate to our results. In fact they have studied the fBM+WN model for $H = 1/2$ (i.e. BM+WN), where the variance of the noise is allowed to depend on n. With the notation of model (1.1), the change in the rate appears as $\beta \downarrow 0$. In particular, they also discuss the case $\beta < 0$ in which the classical $n^{-1/2}$-rate can be achieved. In our framework, $\beta < 0$ corresponds to estimation of the scaling parameter of $(Y_i)_i$.

3 Examples

In the introduction, we have already discussed the main example of estimating the scale parameter of fractional Brownian motion under Gaussian measurement noise. The solution is given in (1.3) (cf. also Corollary 1). In order to provide some further illustration of the derived results, we discuss two estimation problems for which the Fisher information can be explicitly computed.

Large measurement error: Let $(X_i)_i$ denote a stationary process with long-range dependence. More precisely, assume that for constants A, C, and self-similarity parameter $H \in (1/2, 1)$, the autocovariance satisfies $\gamma_k \sim Ak^{2H-2}$, $|\gamma_{k+1} - \gamma_k| \leq C\gamma_k/k$ and for $H \in (1/2, 3/4)$, additionally $\sum_{k=-\infty}^{\infty} \gamma_k^2 = 1$. Suppose that we observe the scaled process $(X_i)_i$ under large noise, i.e.

$$Z_{i,n} = \sigma X_i + \tau n^{\beta} \epsilon_i, \quad i = 1, \ldots, n, \quad 0 < \beta < H - 1/2.$$

Now, with the notation of Theorem 3, $\alpha = 1/2 - H$ and $\Diamond = 2/(2H-1)$. In particular $\Diamond > 4$ for $H \in (1/2, 3/4)$ and $\Diamond < 4$ for $H \in (3/4, 1)$. Therefore, an elbow effect occurs at $H = 3/4$ and the Fisher information is determined in first order by

$$\begin{aligned}
&I_{\sigma^2}^n \sim \frac{n^{1-4\beta}}{2\tau^4}, && H \in (\tfrac{1}{2}, \tfrac{3}{4}),\\
&I_{\sigma^2}^n \sim 4\beta \frac{n^{1-4\beta}\log(n)}{\tau^4}, && H = \tfrac{3}{4},\\
&I_{\sigma^2}^n \sim n^{\frac{2H-1-2\beta}{2H-1}} \sigma^{-\frac{8H-6}{2H-1}} \tau^{-\frac{2}{2H-1}} \frac{\left(2A\Gamma(2H-1)\sin(\pi H)\right)^{\frac{1}{2H-1}}}{(2H-1)^2 \sin(\frac{\pi}{2H-1})}, && H \in (\tfrac{3}{4}, 1).
\end{aligned}$$

Integrated fractional Brownian motion: Suppose that we are interested in efficient estimation of σ^2 given observations $(V_{1,n}, \ldots, V_{n,n})$,

$$V_{i,n} = \sigma \int_{i/n}^{1} B_s^H ds + \tau \epsilon_i, \quad i = 1, \ldots, n, \quad \epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1), \quad H \in (0, 1/4),$$

and $(B_t^H)_t$ is a fBM, which is independent of the WN. With Δ as in (2.8), $X_i := \int_{i-1}^{i} B_s^H ds - \int_{0\vee(i-2)}^{i-1} B_s^H ds$, and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)$, consider

$$\Delta\Delta^t (V_{1,n}, \ldots, V_{n,n})^t \overset{\mathcal{D}}{=} (\sigma n^{-1-H} X_i)_{i=1,\ldots,n} + \tau \Delta\Delta^t \epsilon^t$$

and note that $(X_i)_{i\geq 2}$ is stationary. By defining R_1 appropriately, it is straightforward to verify Assumption 1. In particular, we find that (2.7) is bounded by $\lesssim n^{4H-1} \leq 1$ and

$\gamma_k \sim H(2H-1)k^{2H-2}$. Therefore, $\ell = H(1-2H)$, $\alpha = 1/2 - H, K = 2$ and $\beta = 1 + H$ imply

$$\diamondsuit = \frac{2}{2H+3} < 4$$

and the Fisher information is given by

$$I_{\sigma^2}^n \sim n^{\frac{1}{2H+3}} \sigma^{-\frac{8H+10}{2H+3}} \tau^{-\frac{2}{2H+3}} \frac{(H+1)\sin^{\frac{1}{2H+3}}(\pi H)\,\Gamma(2H+1)^{\frac{1}{2H+3}}}{(2H+3)^2 \sin\left(\frac{\pi}{2H+3}\right)}.$$

Acknowledgments. Till Sabel was supported by DFG grant FOR 916. Johannes Schmidt-Hieber was funded by DFG postdoctoral fellowship SCHM 2807/1-1. We would like to thank Axel Munk for helpful discussions.

Appendix A Notation and some remarks

Let us first give some notation. Whenever it is clear from the context, we omit the index n. In particular, we suppress the index n of the spectral density $h_n = h$ and the estimator $\widehat{\sigma}_n^2 = \widehat{\sigma}^2$. Inequalities for Hermitian matrices should be understand in the sense of partial Loewner ordering. The matrix norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denote the Frobenius and spectral norm, respectively. Furthermore, we write $\wedge$ and $\vee$ for the minimum/maximum and

$$u_i := u_{i,n} := \pi \frac{2i-1}{2n+1}. \tag{A.1}$$

The last definition occurs frequently in connection with finite-dimensional approximations due to the transformation property of the discrete cosine transform. Let $\lfloor \cdot \rfloor$ be the Gauss bracket. As in (2.10), we denote by $(r_n)_n$ the rate at which the Fisher information tends to infinity (this still needs to be proved, of course). For technical reasons, however, it will be chosen in the following as an integer sequence, i.e.

$$r_n := \begin{cases} \lfloor n^{1-\frac{\beta}{K-\alpha}} \left(\ell(n^{\frac{\beta}{K-\alpha}})\right)^{\frac{1}{2K-2\alpha}} + n^{1-4\beta} \rfloor, & \text{if } \diamondsuit \neq 4, \\ \lfloor n^{1-4\beta} \log(n) \ell^2(n) \rfloor, & \text{if } \diamondsuit = 4. \end{cases}$$

This definition will be used at many places throughout the proofs. In particular, one should keep in mind that $u_{r_n} = O(r_n/n)$ and as a direct consequence of (2.5)

Lemma 2. *If ℓ is as in Assumption 1 and $\diamondsuit < 4$, then*

$$\frac{\ell(u_{r_n}^{-1})}{\ell(n^{\frac{\beta}{K-\alpha}})} \to 1.$$

The projection of a function on $\{\cos(k\cdot) : 0 \leq k \leq n\}$ is denoted by S_n, i.e. if $g = \sum_{k=-\infty}^{\infty} g_k \cos(k\cdot)$ with $g_k = g_{-k}$, then

$$S_n g = \sum_{k=-n}^{n} g_k \cos(k\cdot).$$

We write $T_n(g)$ for the $n \times n$ Toeplitz matrix corresponding to g, i.e.

$$T_n(g) = (g_{|i-j|})_{i,j=1,\dots,n}. \tag{A.2}$$

In particular, $\mathrm{Cov}(\mathbf{X}') = T_n(f)$, with f as in (2.9). Let DCT_8 be the discrete cosine transform

$$\mathrm{DCT}_8 = \left(\tfrac{2}{\sqrt{2n+1}} \cos\left[(i-\tfrac{1}{2})(j-\tfrac{1}{2})\tfrac{2\pi}{2n+1}\right]\right)_{i,j=1,\dots,n} = \left(\tfrac{2}{\sqrt{2n+1}} \cos\left[(i-\tfrac{1}{2})u_j\right]\right)_{i,j=1,\dots,n}$$

(which is DCT-VIII in the notation of [6]). Note that $\mathrm{DCT}_8 = \mathrm{DCT}_8^t$ is orthonormal. Further introduce the matrix $D_n(g)$ as

$$D_n(g) := \mathrm{DCT}_8 \cdot \mathrm{diag}\left(g(u_1), g(u_2), \dots, g(u_n)\right) \cdot \mathrm{DCT}_8. \tag{A.3}$$

Asymptotically, the eigenvalues of $D_n(g)$ and $T_n(g)$ are 'close', provided the symbol g is sufficiently smooth (cf. Lemma 9).

If $\mathrm{Cov}(\mathbf{Y}) = \tau^2(\Delta^t\Delta)^K$, consider the observation vector $\widetilde{Z}$ which is obtained by the sufficient transformation $\widetilde{Z}_{i,n} = Z_{n-i,n}$. These observations satisfy our assumptions with $\mathrm{Cov}(\mathbf{Y}) = \tau^2(\Delta\Delta^t)^K$. Therefore, without loss of generality, we can and will consider only the first case of Assumption 2, i.e. $\mathrm{Cov}(\mathbf{Y}) = \tau^2(\Delta\Delta^t)^K$. By Lemma 7, DCT_8 diagonalizes $\Delta^t\Delta$ and hence also $\mathrm{Cov}(\mathbf{Y}) = \tau^2(\Delta^t\Delta)^K$. The eigenvalues of $\mathrm{Cov}(\mathbf{Y})$ are explicitly given by $4^K\tau^2 \sin^{2K}(u_{i,n}\frac{\pi}{2})$, $i = 1, \dots, n$.

For the subsequent proofs, the following three elementary inequalities turn out to be very useful. Firstly, from (2.6) and Potter's bound (cf. for instance Bingham [4]) it follows that for any $\epsilon > 0$ there exists a k_0 such that for any $k \geq k_0$

$$\frac{1}{4}k^{-2\alpha-1-\epsilon} \leq \frac{1}{2}k^{-2\alpha-1}\ell(k) \leq \gamma_k \leq 2k^{-2\alpha-1}\ell(k) \leq 4k^{-2\alpha-1+\epsilon}. \tag{A.4}$$

Moreover, we can find a constant C_1 such that $\gamma_k \leq C_1 k^{-2\alpha-1+\epsilon}$ for all $k = 1, 2, \dots$ Secondly, if $f(\lambda) \sim Cn^{-2\beta}\lambda^{2\alpha}\ell(1/\lambda)$, then, for every $\epsilon > 0$ we can find a $\delta > 0$, such that for all $\lambda \in (0, \delta]$,

$$\frac{1}{4}C\lambda^{2\alpha+\epsilon} \leq \frac{1}{2}C\lambda^{2\alpha}\ell(\tfrac{1}{\lambda}) \leq f(\lambda) \leq 2C\lambda^{2\alpha}\ell(\tfrac{1}{\lambda}) \leq 4C\lambda^{2\alpha-\epsilon}. \tag{A.5}$$

Additionally, under the assumptions of Theorem 2, we know that f is bounded on $[\delta, \pi]$ for every $\delta > 0$ and therefore the upper bound $f(\lambda) \le 4Cn^{-2\beta}\lambda^{2\alpha-\epsilon}$ can be extended to all $\lambda \in (0,\pi]$ by enlarging the constant appropriately. Finally, for all $\lambda \in (0,\pi]$, we have

$$f(\lambda) + 4^{-K}\tau^{-2}\lambda^{2K} \le h(\lambda) \le f(\lambda) + \tau^{-2}\lambda^{2K}. \tag{A.6}$$

Appendix B Proofs

B.1 Proofs of Section 2.1

Lemma 3. *Let $\widetilde{\sigma}^2$ be as defined in* (2.3)*. Given $0 < \underline{\sigma} < \overline{\sigma} < \infty$, there exists an $N = N(\underline{\sigma}, \overline{\sigma})$ such that for all $n \ge N$ and for all $\eta > 0$,*

$$\sup_{\sigma \in [\underline{\sigma}, \overline{\sigma}]} \mathbb{P}\big(|\widetilde{\sigma}^2 - \sigma^2| \ge (I_1^{A_n})^{-1/2} \max(1, \overline{\sigma}^2)\eta\big) \le 2e^{1/4 - \eta/\sqrt{8}}.$$

Proof. Since $\delta_n \to 0$, we can choose N such that for all $n \ge N$, $\delta_n \le \underline{\sigma}^2$ and $\delta_n^{-1} \ge \overline{\sigma}^2$. Then, for $\sigma^2 \in [\underline{\sigma}, \overline{\sigma}]$, $|\widetilde{\sigma}^2 - \sigma^2| \le |\widetilde{\sigma}^2_{\text{pre}} - \sigma^2|$. Let $\alpha_1, \alpha_2, \ldots$ be a sequence of i.i.d. χ_1^2 random variables. By Proposition 6 in Rohde and Duembgen [24] (similar statements have been derived also elsewhere, for another reference see for instance Johnstone [20], p.74), for a vector $(\mu_i)_{i \in A_n}$ of real-valued numbers

$$\mathbb{P}\big(|\sum_{i \in A_n} \mu_i(\alpha_i - 1)| \ge \sqrt{2}\|\mu\|_2 \eta\big) \le 2e^{1/4 - \eta/\sqrt{8}}.$$

Note that in distribution, $\widetilde{\sigma}^2_{\text{pre}} - \sigma^2 = \sum_{i \in A_n} \mu_i(\alpha_i - 1)$ with

$$\mu_i = (2I_1^{A_n})^{-1} \frac{\lambda_i n^{-2\beta}(\sigma^2 n^{-2\beta}\lambda_i + 1)}{(n^{-2\beta}\lambda_i + 1)^2}.$$

Application of the exponential inequality above together with $\|\mu\|_2 \le (2I_1^{A_n})^{-1/2} \max(\sigma^2, 1)$ yields the result. ☐

Proof of Theorem 1. Due to the independence of $(\widetilde{Z}_i)_{i \in A_n}$ and $(\widetilde{Z}_i)_{i \in A_n^c}$, the estimator $\widehat{\sigma}^2$ is unbiased. In addition, using Lemma 1, the theorem is proved once we have established that

$$(I): \sup_{\sigma \in [\underline{\sigma}, \overline{\sigma}]} \frac{|\operatorname{Var}(\widehat{\sigma}^2) - \mathbb{E}(I_{\widehat{\sigma}^2}^{A_n^c})^{-1}|}{(I_{\sigma^2}^n)^{-1}} = o(1),$$

$$\text{and} \quad (II): \sup_{\sigma \in [\underline{\sigma}, \overline{\sigma}]} \frac{|\mathbb{E}(I_{\widehat{\sigma}^2}^{A_n^c})^{-1} - (I_{\sigma^2}^{A_n^c})^{-1}|}{(I_{\sigma^2}^n)^{-1}} = o(1).$$

In the following, we make frequently use of the following observation. For any set $B \subseteq \{1,\ldots,n\}$,

$$\min\big(1, (\tfrac{v}{u})^2\big) I_v^B \le I_u^B \le \max\big(1, (\tfrac{v}{u})^2\big) I_v^B.$$

Proof of (I): Writing $\mathbb{E}[\cdot] = \mathbb{E}[\mathbb{E}[\cdot|(\widetilde{Z}_{i,n})_{i\in A_n}]]$,

$$\begin{aligned}\Big|\operatorname{Var}(\widehat{\sigma}^2) - \mathbb{E}(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big| &= 2\Big|\mathbb{E}\Big[(2I_{\widetilde{\sigma}^2}^{A_n^c})^{-2}\sum_{i\in A_n^c}\frac{\lambda_i^3 n^{-6\beta}(\sigma^2-\widetilde{\sigma}^2)}{(\widetilde{\sigma}^2\lambda_i n^{-2\beta}+1)^4}\Big((\sigma^2+\widetilde{\sigma}^2)\lambda_i n^{-2\beta}+2\Big)\Big]\Big| \\ &\le 2\mathbb{E}\Big[(2I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|\big(3+\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|\big)\Big],\end{aligned}$$

where we used the inequalities

$$\sum_{i\in A_n^c}\frac{\lambda_i^4 n^{-8\beta}}{(\widetilde{\sigma}^2\lambda_i n^{-2\beta}+1)^4} \le \frac{2}{\widetilde{\sigma}^4} I_{\widetilde{\sigma}^2}^{A_n^c} \quad \text{and} \quad \sum_{i\in A_n^c}\frac{\lambda_i^3 n^{-6\beta}}{(\widetilde{\sigma}^2\lambda_i n^{-2\beta}+1)^4} \le \frac{1}{\widetilde{\sigma}^2} I_{\widetilde{\sigma}^2}^{A_n^c}.$$

Therefore

$$\begin{aligned}\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{E}\Big[(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|\Big] &= \sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\int \mathbb{P}\big[|\sigma^2-\widetilde{\sigma}^2| \ge \widetilde{\sigma}^2 I_{\widetilde{\sigma}^2}^{A_n^c} t\big]dt \\ &\le \int \sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{P}\big[|\sigma^2-\widetilde{\sigma}^2| \ge \delta_n I_1^{A_n^c} t\big]dt.\end{aligned}$$

Application of Lemma 3 together with $\int_0^\infty \exp(-At)dt = A^{-1}$ yields

$$\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{E}\Big[(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|\Big] \lesssim (\delta_n^2 I_1^{A_n})^{-1/2}(I_1^{A_n^c})^{-1}.$$

Similar,

$$\begin{aligned}\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{E}\Big[(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|^2\Big] &= \sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\int \mathbb{P}\big[|\sigma^2-\widetilde{\sigma}^2| \ge \widetilde{\sigma}^2 (I_{\widetilde{\sigma}^2}^{A_n^c})^{1/2}\sqrt{t}\big]dt \\ &\le \int \sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{P}\big[|\sigma^2-\widetilde{\sigma}^2| \ge \delta_n (I_1^{A_n^c})^{1/2}\sqrt{t}\big]dt\end{aligned}$$

and because of $\int_0^\infty \exp(-A\sqrt{t})dt = A^{-2}$,

$$\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{E}\Big[(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|\frac{\sigma^2}{\widetilde{\sigma}^2}-1\Big|^2\Big] \lesssim (\delta_n^2 I_1^{A_n})^{-1}(I_1^{A_n^c})^{-1}.$$

By definition $\delta_n^2 I_1^{A_n} \to \infty$. Since for sufficiently large n, by Lemma 1,

$$I_{\sigma^2}^n \le \max(1,\sigma^{-4}) I_1^n \le 2\max(1,\sigma^{-4}) I_1^{A_n^c}, \tag{B.1}$$

it follows that

$$\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\frac{|\operatorname{Var}(\widehat{\sigma}^2)-\mathbb{E}(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}|}{(I_{\sigma^2}^{n})^{-1}}\lesssim I_1^{A_n^c}\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\Big|\operatorname{Var}(\widehat{\sigma}^2)-\mathbb{E}(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}\Big|=o(1).$$

Proof of (II): From Taylor expansion, we find that for positive x, y, $|x^{-2}-y^{-2}|\leq 2(\min(x,y))^{-3}|x-y|$. Therefore, we can bound

$$I_{\sigma^2}^{A_n^c}-I_{\widetilde{\sigma}^2}^{A_n^c}=\frac{1}{2}\sum_{i\in A_n^c}n^{-4\beta}\lambda_i^2\Big(\frac{1}{(\sigma^2n^{-2\beta}\lambda_i+1)^2}-\frac{1}{(\widetilde{\sigma}^2n^{-2\beta}\lambda_i+1)^2}\Big)$$

by

$$\begin{aligned}|I_{\sigma^2}^{A_n^c}-I_{\widetilde{\sigma}^2}^{A_n^c}|&\leq\sum_{i\in A_n^c}\frac{n^{-6\beta}\lambda_i^3}{(\min(\sigma^2,\widetilde{\sigma}^2)n^{-2\beta}\lambda_i+1)^3}|\sigma^2-\widetilde{\sigma}^2|\\&=2\max(\sigma^{-2},\widetilde{\sigma}^{-2})I_{\sigma^2\wedge\widetilde{\sigma}^2}^{A_n^c}|\sigma^2-\widetilde{\sigma}^2|\leq(\underline{\sigma}^{-6}+1)\delta_n^{-2}I_{\widetilde{\sigma}^2}^{A_n^c}|\sigma^2-\widetilde{\sigma}^2|.\end{aligned}$$

Thus,

$$\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}I_{\sigma^2}^{A_n^c}|\mathbb{E}(I_{\widetilde{\sigma}^2}^{A_n^c})^{-1}-(I_{\sigma^2}^{A_n^c})^{-1}|\lesssim\delta_n^{-2}\int\sup_{\sigma\in[\underline{\sigma},\overline{\sigma}]}\mathbb{P}(|\sigma^2-\widetilde{\sigma}^2|\geq t)dt\lesssim(\delta_n^4I_1^{A_n})^{-1/2}\to 0.$$

Using (B.1), the convergence in (II) follows and this completes the proof. ☐

B.2 Proofs of Theorem 2

The proof of the main theorem builds in a very neat way upon an elementary analytical observation (cf. Lemma 4) which leads in a second step to a trace approximation for positive semidefinite matrices (cf. Lemma 5). This approximation result does not require any assumption on the behavior of the smallest or largest eigenvalue. Together with a rather standard but slightly technical Riemann approximation argument, we can then deduce a generalized version of Theorem 2.

Lemma 4. *Let $(x_n)_n,(y_n)_n,(q_n)_n$ be positive sequences such that $|x_n-y_n|=O\big(q_n\,(\sqrt{x_n}+\sqrt{y_n})\,\big)$. Then, $x_n=y_n(1+o(1))+O(q_n^2)$.*

Proof. First, assume that $q_n=o(\sqrt{y_n})$. The assumption $|x_n-y_n|=O\big(q_n\,(\sqrt{x_n}+\sqrt{y_n})\,\big)$ is equivalent to $|\sqrt{x_n}-\sqrt{y_n}|=O(q_n)$ and $|(x_n/y_n)^{1/2}-1|=O(q_n{y_n}^{-1/2})=o(1)$. This implies $(x_n/y_n)^{1/2}\sim 1$ and thus $x_n/y_n\sim 1$, i.e. $x_n=y_n(1+o(1))$. Now, if $q_n=o(\sqrt{x_n})$ the same argument shows that $x_n(1+o(1))=y_n$ (but this is of course equivalent to $x_n=y_n(1+o(1))$). If neither $q_n=o(\sqrt{y_n})$ nor $q_n=o(\sqrt{x_n})$ holds, then, $q_n\gtrsim\sqrt{x_n}+\sqrt{y_n}$ and therefore, $|x_n-y_n|=O(q_n^2)$. ☐

Let A be an $n \times n$ matrix. For convenience, we introduce the notation

$$\langle A \rangle := \operatorname{tr}(A^2).$$

Lemma 5. *Let A_1, A_2, B be (sequences of) positive semidefinite, $n \times n$ matrices and suppose that A_1 is invertible. Then, for $n \to \infty$,*

$$\langle A_1 B \rangle = \langle A_2 B \rangle \big(1 + o(1)\big) + O\big(\langle B(A_1 - A_2) \rangle\big).$$

Furthermore, if $B \leq A_1^{-1}$, then

$$\langle A_1 B \rangle = \langle A_2 B \rangle \big(1 + o(1)\big) + O\big(\langle \operatorname{id}_n - A_2^{1/2} A_1^{-1} A_2^{1/2} \rangle\big).$$

Proof. By Cauchy-Schwarz,

$$\begin{aligned} \big|\langle A_1 B \rangle - \langle A_2 B \rangle\big| &= \big|\operatorname{tr}(B(A_1 - A_2)BA_1) + \operatorname{tr}(BA_2B(A_1 - A_2))\big| \\ &\leq \|B^{1/2}(A_1 - A_2)B^{1/2}\|_2 \Big[\operatorname{tr}((A_1B)^2)^{1/2} + \operatorname{tr}((A_2B)^2)^{1/2}\Big]. \end{aligned}$$

For the last inequality we have rewritten $\operatorname{tr}(B(A_1 - A_2)BA_1)$ and $\operatorname{tr}(BA_2B(A_1 - A_2))$ as $\operatorname{tr}(B^{1/2}(A_1 - A_2)B^{1/2} \cdot B^{1/2}A_1B^{1/2})$ and $\operatorname{tr}(B^{1/2}(A_1 - A_2)B^{1/2} \cdot B^{1/2}A_2B^{1/2})$. Since $\langle B(A_1 - A_2) \rangle = \|B^{1/2}(A_1 - A_2)B^{1/2}\|_2^2$, the result follows with Lemma 4.

To prove the second claim, write

$$B^{1/2}(A_1 - A_2)B^{1/2} = B^{1/2}A_1^{1/2}(\operatorname{id}_n - A_1^{-1/2}A_2A_1^{-1/2})A_1^{1/2}B^{1/2}$$

and note that due to $A_1^{1/2}BA_1^{1/2} \leq \operatorname{id}_n$,

$$\langle B(A_1 - A_2) \rangle \leq \langle \operatorname{id}_n - A_1^{-1/2}A_2A_1^{-1/2} \rangle = \langle \operatorname{id}_n - A_2^{1/2}A_1^{-1}A_2^{1/2} \rangle.$$

□

In the case $\alpha > 0$ the multiplicative inverse of the spectral density h has a singularity at zero. In order to deal with this, we introduce the regularized spectral density $\widetilde{h}$, which is defined as follows: Let (ρ_n) be a sequence of positive integers satisfying $\rho_n \ll r_n$. Then, we define

$$\widetilde{h}(\lambda) := \begin{cases} h(\lambda) \vee h(u_{\rho_n}), & \lambda \leq u_{\rho_n} \\ h(\lambda), & else \end{cases}$$

with u_{ρ_n} as in (A.1). Replacing h by f, define in the same way $\widetilde{f}$. We will prove a generalized version of Theorem 2 for a generic sequence $(\rho_n)_n$. In a second step the different versions of the main theorem are deduced and ρ_n will be chosen according to the specific setting. We may interpret this spectral regularization as adding an asymptotically non-informative (i.e. sufficiently small) WN process to our observation vector. This induces some stability, which becomes important in the bounds for the inverse covariance matrices.

Theorem 4. *Work under Assumption 1 and Assumption 2 in model* (1.1)*. Suppose that* $\alpha \in (-1/2, 1/2)$ *and* $K - \alpha > \beta \vee 1/4$*. If*

(i) $(\gamma_k)_{k\geq 0}$ *is in* GM*,* f *is bounded on any interval* $[\delta, \pi]$ *with* $\delta > 0$ *and there exists a positive, quasi-monotone slowly varying function* ℓ*, such that*

$$f(\lambda) \sim 2\,\mathrm{sign}(-\alpha)\Gamma(-2\alpha)\cos(\pi\alpha)\lambda^{2\alpha}\ell(1/\lambda),$$

(ii) $n^{-4\beta-4\alpha-2+2\epsilon}\sum_{i=1}^{n}(u_{i,n}\widetilde{h}(u_{i,n}))^{-2} = o(r_n)$*, for some* $\epsilon > 0$*,*

(iii) $\langle D_n^{-1}(\widetilde{h})(D_n(S_n f) - T_n(f))\rangle + \sup_{\lambda\in(0,\pi]}\widetilde{h}^{-2}(\lambda) = o(r_n n^{4\beta})$*.*

Then, the asymptotic Fisher information of σ^2 *is*

$$I_{\sigma^2}^n = \frac{n^{1-4\beta}}{2\pi}\int_0^\pi \tfrac{f^2(\lambda)}{h^2(\lambda)}d\lambda(1+o(1)) + o(r_n). \tag{B.2}$$

If the condition $K - \alpha > \beta \vee 1/4$ *is replaced by the weaker assumption* $K - \alpha > \beta$*, imposing additionally* $\log(n)\ell^2(n) \to \infty$ *in the critical case* $K - \alpha = 1/4$*, then* (B.2) *holds, provided there exists a constant* C_f *such that*

$$|f(\lambda) - f(\mu)| \leq C_f n^{-2\beta}\lambda^{2\alpha-2}|\lambda - \mu|, \quad \textit{for all} \;\; 0 < \lambda \leq \mu \leq \pi. \tag{B.3}$$

Remark 1. *Later on we will see that the different parts of Theorem 2 follow from Theorem 4. If* $(X_i')_i$ *has long-memory, condition* (*iii*) *turns out to be quite difficult to verify. Although it would be easier (and more standard) to formulate the condition with respect to squared Frobenius norms, let us shortly explain, why the use of the* $\langle\cdot\rangle$ *notation is essential. By definition,* $\langle A\rangle = \mathrm{tr}(A^2)$ *for an* $n \times n$*, square matrix* A*, which in turn can be upper bounded by the squared Frobenius norm of* A *(cf. Lemma 6 (i)). This is even an identity if* A *is symmetric but can be very rough in general. To see this consider* $(A)_{i,j} = 1/i$*. Then* $\langle A\rangle = \log^2 n$ *but the squared Frobenius norm is of order* n *which is much worse. Since these phenomenons occur in some cases, condition* (*iii*) *is stated for squared traces.*

Proof. Recall the explicit expression for the Fisher information in (2.1). The proof is subdivided into three steps, namely

$$\langle D_n(\widetilde{f}) D_n^{-1}(\widetilde{h})\rangle = \frac{n}{\pi}\int_0^{\pi} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda (1+o(1)) + o(r_n n^{4\beta}),$$
$$\langle T_n(f) D_n^{-1}(\widetilde{h})\rangle = \langle D_n(\widetilde{f}) D_n^{-1}(\widetilde{h})\rangle (1+o(1)) + o(r_n n^{4\beta}),$$
$$2I_{\sigma^2}^n = \langle n^{-2\beta} T_n(f) \operatorname{Cov}(\mathbf{Z})^{-1}\rangle = \langle n^{-2\beta} T_n(f) D_n^{-1}(\widetilde{h})\rangle (1+o(1)) + o(r_n).$$

which are denoted by $(I), (II)$ and (III), respectively.

(I): By the trivial bound $\int_0^\pi f^2(\lambda)/h^2(\lambda) d\lambda = O(n^{4\beta}) = o(r_n n^{4\beta})$, we can replace the normalization factor n/π by $(2n+1)/(2\pi)$. Thus, using (A.3), it is sufficient to show that

$$\Big|\sum_{i=1}^{n} \tfrac{\widetilde{f}^2(u_i)}{\widetilde{h}^2(u_i)} - \tfrac{2n+1}{2\pi}\int_0^{\pi} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda\Big| = o(r_n n^{4\beta}). \tag{B.4}$$

Now, let us treat the cases $K-\alpha > 1/4 \vee \beta$ and $\beta < K-\alpha \leq 1/4$, separately.

If $K-\alpha > 1/4 \vee \beta$ holds: Using Assumption (i) and $r_n \ll n$, we can find integer sequences (r_n^+) and (r_n^-) such that $\rho_n \ll r_n^- \ll r_n \ll r_n^+ \ll n$ and $(q_n + (r_n^-)^{-1}) r_n^+ = o(r_n)$ with

$$q_n := q_n(r_n^+) := \sup_{\substack{0<\lambda\leq u_{r_n^+} \\ 1\leq \mu \leq 2}} \Big|\frac{C\lambda^{2\alpha}\ell(1/\lambda)}{f(\lambda)} - 1\Big| + \Big|1 - \frac{\ell(\frac{\mu}{\lambda})}{\ell(\frac{1}{\lambda})}\Big|. \tag{B.5}$$

Since $\sigma^2 n^{-2\beta} f \leq h$ and $\sigma^2 n^{-2\beta}\widetilde{f} \leq \widetilde{h}$, it follows that

$$\sum_{i=1}^{r_n^-} \tfrac{\widetilde{f}^2(u_i)}{\widetilde{h}^2(u_i)} = o(r_n n^{4\beta}) \quad \text{and} \quad n\int_0^{u_{r_n^-}} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda = o(r_n n^{4\beta}).$$

and together with Proposition 3, we see that in (B.4) the sum over $i = 1,\ldots,r_n^-$ and $i = r_n^+,\ldots,n$ as well as the integral over $(0, u_{r_n^-}] \cup [u_{r_n^+}, \pi]$ are of order $o(r_n)$ and thus negligible. Thus, we have proved (I), once we have verified that

$$\Big|\sum_{i=r_n^-+1}^{r_n^+} \tfrac{f^2(u_i)}{h^2(u_i)} - \tfrac{2n+1}{2\pi}\int_{u_{r_n^-}}^{u_{r_n^+}} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda\Big| = o(r_n n^{4\beta}). \tag{B.6}$$

To see this, write

$$\Big|\sum_{i=r_n^-+1}^{r_n^+} \tfrac{f^2(u_i)}{h^2(u_i)} - \tfrac{2n+1}{2\pi}\int_{u_{r_n^-}}^{u_{r_n^+}} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda\Big| \leq \sum_{i=r_n^-+1}^{r_n^+} \sup_{\xi_i\in[u_{i-1},u_i]} \Big|\tfrac{f^2(u_i)}{h^2(u_i)} - \tfrac{f^2(\xi_i)}{h^2(\xi_i)}\Big|$$
$$\leq \frac{2n^{4\beta}}{\sigma^2}\sum_{i=r_n^-+1}^{r_n^+} \sup_{\xi_i\in[u_{i-1},u_i]} \Big|\tfrac{n^{-2\beta}f(u_i)}{h(u_i)} - \tfrac{n^{-2\beta}f(\xi_i)}{h(\xi_i)}\Big|. \tag{B.7}$$

Fix $i \in \{r_n^- + 1, \ldots, r_n^+\}$ and let $a_1 = \sigma^2 n^{-2\beta} f(u_i), a_2 = \sigma^2 n^{-2\beta} f(\xi_i), b_1 = 4^K \tau^2 \sin^{2K}(u_i/2)$, $b_2 = 4^K \tau^2 \sin^{2K}(\xi_i/2), c_1 = h(u_i), c_2 = h(\xi_i)$. Since $0 \le a_1 \le c_1 = a_1 + b_1$ and $0 \le a_2 \le c_2 = a_2 + b_2$, we find

$$\Big|\frac{a_1}{c_1} - \frac{a_2}{c_2}\Big| \le \frac{|a_1 - a_2| + |b_1 - b_2|}{c_1 \vee c_2}. \tag{B.8}$$

Thus, for sufficiently large n,

$$\begin{aligned} |a_1 - a_2| &\le (a_1 + a_2) q_n + \sigma^2 C n^{-2\beta} \big| u_i^{2\alpha} \ell(\tfrac{1}{u_i}) - \xi_i^{2\alpha} \ell(\tfrac{1}{\xi_i}) \big| \\ &\le (3a_1 + a_2) q_n + \sigma^2 C n^{-2\beta} \ell(\tfrac{1}{\xi_i}) |u_i^{2\alpha} - \xi_i^{2\alpha}| \\ &\le (3a_1 + a_2) q_n + \tfrac{\pi}{n} \sigma^2 C n^{-2\beta} \ell(\tfrac{1}{\xi_i}) \xi_i^{2\alpha - 1} \le (3a_1 + a_2) q_n + 6 (r_n^-)^{-1} a_2. \end{aligned} \tag{B.9}$$

On the other hand, we find $|b_1 - b_2| \le \frac{2K\pi}{2n+1} 4^K \tau^2 \sin^{2K-1}\big(\frac{u_i}{2}\big) \le 8K (r_n^-)^{-1} b_1$, and therefore

$$\Big|\frac{a_1}{c_1} - \frac{a_2}{c_2}\Big| \le 4 q_n + (6 + 8K)(r_n^-)^{-1}.$$

Due to $(q_n + (r_n^-)^{-1}) r_n^+ = o(r_n)$, we see by (B.7) that (B.6) is bounded by $o(r_n n^{4\beta})$. This completes the proof for part (I) if $K - \alpha > 1/4 \vee \beta$.

If $\beta < K - \alpha \le 1/4$: The proof is very similar to the one for the first case. Note that the assumptions imply $r_n n^{4\beta} \gtrsim n$, $K = 0$, and $\alpha \in [-1/4, -\beta)$. Similar as above we see that it is sufficient to prove (B.6) for $r_n^+ = n$ and any sequence $r_n^- = o(r_n)$. We may assume that $r_n^- \to \infty$. Since by (B.3), f and h are continuous, we can apply the mean value theorem, i.e. for any i there is a $\xi_i \in (u_{i-1}, u_i]$ with

$$\tfrac{f^2(\xi_i)}{h^2(\xi_i)} = \tfrac{2n+1}{2\pi} \int_{u_{i-1}}^{u_i} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda \tag{B.10}$$

and

$$\Big| \sum_{i=r_n^-+1}^{n} \tfrac{f^2(u_i)}{h^2(u_i)} - \tfrac{2n+1}{2\pi} \int_{u_{r_n^-}}^{u_n} \tfrac{f^2(\lambda)}{h^2(\lambda)} d\lambda \Big| \le \sum_{i=r_n^-+1}^{n} \Big(\tfrac{f(u_i)}{h(u_i)} + \tfrac{f(\xi_i)}{h(\xi_i)} \Big) \Big| \tfrac{f(u_i) - f(\xi_i)}{h(u_i) \vee h(\xi_i)} \Big|. \tag{B.11}$$

Pick an integer sequence w_n such that $w_n = o(n)$ and for some $\epsilon > 0$, $n^{-4\alpha+\epsilon} w_n^{4\alpha-\epsilon-1} = o(1)$. Let $q_n(w_n)$ be as in (B.5) with r_n^+ replaced by w_n. Note that since $u_{w_n} \to 0$, the sequence $(q_n(w_n))_n$ tends to zero. As in (B.9), we find for sufficiently large n and for all $i = r_n^-, \ldots, w_n$,

$$|f(u_i) - f(\xi_i)| \le (3f(u_i) + f(\xi_i)) q_n(w_n) + 2\sigma^2 \pi C n^{-1} \xi_i^{2\alpha - 1 - \epsilon}. \tag{B.12}$$

For large indices, we use the estimate (B.3), i.e. $|f(u_i) - f(\xi_i)| \le C_f \pi n^{-1} \xi_i^{2\alpha-2}$ for $i = w_n + 1, \ldots, n$. Split the sum in (B.11) into $\sum_{i=r_n^-+1}^{w_n} + \sum_{i=w_n+1}^{n}$. It is easy to bound the second sum, using the definition of $(w_n)_n$, which in turn implies that $n^{-1}\sum_{i=w_n+1}^{n}\xi_i^{4\alpha-2-\epsilon}\lesssim n^{1-4\alpha+\epsilon}w_n^{4\alpha-1-\epsilon}=o(n)$. Similar, computing the sum $\sum_{i=r_n^-+1}^{w_n}$, over the second term in (B.12) yields for small ϵ, $n^{-1}\sum_{i=r_n^-+1}^{w_n}\xi_i^{4\alpha-1-2\epsilon}\lesssim n^{-4\alpha+2\epsilon}w_n^{4\alpha-2\epsilon}=o(n)$. For the first term in (B.12) we see that (B.11) can be further bounded by a multiple of $q_n(w_n)(\sum_{i=1}^n f^2(u_i)/h^2(u_i)+(2n+1)(2\pi)^{-1}\int_0^\pi f^2(\lambda)/h^2(\lambda)d\lambda)$. Putting all estimates together, we have derived

$$|\sum_{i=r_n^-+1}^{n}\tfrac{f^2(u_i)}{h^2(u_i)}-\tfrac{2n+1}{2\pi}\int_{u_{r_n^-}}^{u_n}\tfrac{f^2(\lambda)}{h^2(\lambda)}d\lambda|\lesssim\tfrac{2n+1}{2\pi}\int_0^\pi\tfrac{f^2(\lambda)}{h^2(\lambda)}d\lambda+o(n).$$

This finishes the proof of part (I).

To prove (II) and (III) it will not be necessary to distinguish whether $K-\alpha>1/4$ or $K-\alpha\leq 1/4$.

(II): In Lemma 5, set $A_1=T_n(f)$, $A_2=D_n(\widetilde{f})$ and $B=D_n^{-1}(\widetilde{h})$. We have to show that

$$\langle B(A_1-A_2)\rangle=\langle D_n^{-1}(\widetilde{h})(T_n(f)-D_n(\widetilde{f}))\rangle=o(r_nn^{4\beta}).\tag{B.13}$$

First, note that due to (A.3), $\sigma^2n^{-2\beta}f\leq\sigma^2n^{-2\beta}\widetilde{f}\leq\widetilde{h}$, and $\rho_n\ll r_n$,

$$\langle D_n^{-1}(\widetilde{h})(D_n(f)-D_n(\widetilde{f}))\rangle=\sum_{j=1}^{\rho_n}\frac{(f(u_j)-\widetilde{f}(u_j))^2}{(\widetilde{h}(u_j))^2}\leq 2\frac{\rho_nn^{4\beta}}{\sigma^4}=o(r_nn^{4\beta}).$$

By Assumption $(\gamma_k)_k\in\mathrm{GM}$ and $\alpha\in(-1/2,1/2)$. Therefore, we can use the estimate from Lemma 8 (i) together with (A.4), i.e. there exists a constant C_1, such that

$$|f(x)-S_nf(x)|\leq C_1\frac{1}{x}(|\gamma_n|+\sum_{k=n+1}^{\infty}\tfrac{|\gamma_k|}{k})\lesssim\frac{1}{x}n^{-2\alpha-1+2\epsilon},\quad\text{for all }x\in[\tfrac{1}{n},\pi),$$

where the second inequality holds for sufficiently large n and ϵ small. With (A.3) and Assumption (ii) this yields

$$\langle D_n^{-1}(\widetilde{h})(D_n(S_nf)-D_n(f))\rangle=o(r_nn^{4\beta}).\tag{B.14}$$

Decompose $T_n(f)-D_n(\widetilde{f})=(T_n(f)-D_n(S_nf))+(D_n(S_nf)-D_n(f))+(D_n(f)-D_n(\widetilde{f}))$. Now by Lemma 6, (iv) and Assumption (iii), (B.13) follows.

(III): Set $A_1=\mathrm{Cov}(\mathbf{Z})^{-1}$, $A_2=D_n^{-1}(\widetilde{h})$ and $B=\sigma^2n^{-2\beta}\,\mathrm{Cov}(\mathbf{X})$ and apply Lemma 5. Since $B\leq\mathrm{Cov}(\mathbf{Z})=A_1^{-1}$ it is sufficient to show

$$\langle\,\mathrm{id}_n-A_2^{1/2}A_1^{-1}A_2^{1/2}\rangle=\langle\,\mathrm{id}_n-D_n^{-1/2}(\widetilde{h})\,\mathrm{Cov}(\mathbf{Z})D_n^{-1/2}(\widetilde{h})\rangle=o(r_n).\tag{B.15}$$

By the perfect diagonalization property of $\mathrm{Cov}(\mathbf{Y})$ (cf. the remarks in Section A), we have

$\operatorname{Cov}(\mathbf{Y}) = D_n(h - \sigma^2 n^{-2\beta} f)$ and

$\operatorname{Cov}(\mathbf{Z}) = \sigma^2 n^{-2\beta}\big[\operatorname{Cov}(\mathbf{R}) + \operatorname{Cov}(\mathbf{X}', \mathbf{R}) + \operatorname{Cov}(\mathbf{R}, \mathbf{X}')\big] + \sigma^2 n^{-2\beta}(T_n(f) - D_n(f)) + D_n(h).$

Together with Lemma 6 (iv),

$$\begin{aligned}\big\langle \operatorname{id}_n - D_n^{-1/2}(\widetilde{h}) \operatorname{Cov}(\mathbf{Z}) D_n^{-1/2}(\widetilde{h})\big\rangle \lesssim & \big\langle \operatorname{id}_n - D_n(\widetilde{h}^{-1} h)\big\rangle \\ & + n^{-4\beta} \sup_{\lambda\in(0,\pi]} \widetilde{h}^{-2}(\lambda) \| \operatorname{Cov}(\mathbf{R}) + \operatorname{Cov}(\mathbf{X}', \mathbf{R}) + \operatorname{Cov}(\mathbf{R}, \mathbf{X}')\|_2^2 \\ & + n^{-4\beta} \langle D_n^{-1}(\widetilde{h})(D_n(f) - D_n(S_n f))\rangle \\ & + n^{-4\beta} \langle D_n^{-1}(\widetilde{h})(D_n(S_n f) - T_n(f))\rangle. \end{aligned} \tag{B.16}$$

For the first term note that because of (A.3) and $0 \le 1 - h/\widetilde{h} \le 1$,

$$\big\langle \operatorname{id}_n - D_n(\widetilde{h}^{-1} h)\big\rangle = \sum_{i=1}^{\rho_n} \Big(1 - \frac{h(u_i)}{\widetilde{h}(u_i)}\Big)^2 = o(r_n).$$

The other three terms on the r.h.s. of (B.16) can be seen to be of order $o(r_n)$ as well, by Assumption 1, Assumption (ii), (B.14) and Assumption (iii). Therefore, (B.15) holds and the proof is completed. □

Proof of Theorem 2, Part 1. We check the conditions of Theorem 4. Note that the special case, i.e. $K-\alpha \le 1/4$ implies together with $K-\alpha > \beta$ that $K = 0$ and $\alpha \in [-1/4, -\beta)$. Hence this case only plays a role in Part 2 and Part 3 (in the latter only if $K = 0$ and $\alpha = -1/4$). All the derived estimates will work for both situations and thus, in the following, we do not distinguish between these two cases explicitly.

Let $\rho_n = n^{1-(4\alpha)^{-1}(1-\frac{\beta}{K-\alpha})+\delta} \ll r_n$ for some $\delta > 0$. Such a δ always exists thanks to the assumption $K - \alpha > (4\alpha + 1)\beta$. This assures that

$$(\rho_n/n)^{-4\alpha-2\epsilon} = o(r_n) \quad \text{for } \epsilon \text{ small enough.} \tag{B.17}$$

(i): By [27], $(\gamma_k)_k \in \mathrm{GM}$. The second part follows from Lemma 11.

(ii): Making use of inequalities (A.5) and (A.6),

$$\begin{aligned} & n^{-4\beta-4\alpha-2+2\epsilon} \sum_{i=1}^{n} (u_{i,n}\widetilde{h}(u_{i,n}))^{-2} \\ \lesssim & n^{-4\beta-4\alpha-2+2\epsilon}\Big[\sum_{i=1}^{\rho_n} (\tfrac{n}{i}(\tfrac{n}{\rho_n})^{2\alpha+\epsilon} n^{2\beta})^2 + \sum_{i=\rho_n+1}^{r_n} (\tfrac{n}{i})^{4\alpha+2\epsilon+2} n^{4\beta} + \sum_{i=r_n+1}^{n} (\tfrac{n}{i})^{4K+2}\Big] \\ \lesssim & n^{4\epsilon}(\rho_n)^{-4\alpha-2\epsilon} + n^{4(K-\alpha-\beta)+2\epsilon} r_n^{-4K-1} = o(r_n), \end{aligned}$$

if ϵ is chosen small enough.

(iii): By Lemma 6 (iii) and (i), Lemma 9, and $T_n(f) = T_n(S_n f)$,

$$\begin{aligned}\langle D_n^{-1}(\widetilde{h})(D_n(S_n f) - T_n(f))\rangle &\leq \|D_n^{-1}(\widetilde{h})\|_\infty^2 \langle D_n(S_n f) - T_n(f)\rangle \\ \leq &\|D_n^{-1}(\widetilde{h})\|_\infty^2 \|D_n(S_n f) - T_n(f)\|_2^2 \lesssim \left(\frac{\rho_n}{n}\right)^{-4\alpha-2\epsilon} n^{4\beta} = o(r_n n^{4\beta}),\end{aligned}$$

if δ and ϵ are chosen appropriately. The bound for $\sup_{\lambda\in(0,\pi]} \widetilde{h}^{-2}(\lambda)$ is the same and this completes the proof of the claim. □

Proof of Theorem 2, Part 2. Let $\rho_n = 1$, i.e. $\widetilde{h}(u_i) = h(u_i)$. The proof of this part is similar to the one for (i). Again, we check the conditions of Theorem 4:

(i): This follows from Lemma 11.

(ii): Splitting the sum $\sum_{i=1}^{n} = \sum_{i=1}^{r_n} + \sum_{i=r_n+1}^{n}$, we find for small ϵ, by inequalities (A.5) and (A.6),

$$n^{-4\beta-4\alpha-2+2\epsilon} \sum_{i=1}^{n} (u_{i,n}\widetilde{h}(u_{i,n}))^{-2} \lesssim n^{4\epsilon} + n^{4(K-\alpha-\beta)+2\epsilon} r_n^{-4K-1} = o(r_n).$$

(iii): Observe that by (A.5) and (A.6), $h(\lambda) \gtrsim n^{-\frac{2K\beta}{K-\alpha-\epsilon/2}}$. Together with Lemma 9,

$$\begin{aligned}\langle D_n^{-1}(\widetilde{h})(D_n(S_n f) - T_n(f))\rangle + \sup_{\lambda\in(0,\pi]} \widetilde{h}^{-2}(\lambda) &\leq \sup_{\lambda\in(0,\pi]} \widetilde{h}^{-2}(\lambda)(\|D_n(S_n f) - T_n(f)\|_2^2 + 1) \\ &\lesssim n^{\frac{4K\beta}{K-\alpha-\epsilon/2}-4\alpha+2\epsilon} = o(r_n n^{4\beta}).\end{aligned}$$

□

Proof of Theorem 2, Part 3. We apply Theorem 4 with $\widetilde{f}(u_i) = f(u_i)$ and $\widetilde{h}(u_i) = h(u_i)$, i.e. $\rho_n = 1$.

(i): This follows from Lemma 10 and Lemma 11.

For the following parts we make frequently use of the inequalities (A.4)-(A.6) and the subsequent comments.

(ii): Since $u_{r_n} \to 0$, we find, if n is sufficiently large,

$$\sum_{i=1}^{n} \frac{1}{u_i^2 h(u_i)^2} \lesssim \sum_{i=1}^{r_n} \frac{n^{4\beta}}{u_i^2 f(u_i)^2} + \sum_{i=r_n+1}^{n} \frac{1}{u_i^{2+4K}} \lesssim n^{2+4\beta+4\alpha+2\epsilon} r_n^{-1-4\alpha} + n^{2+4K} r_n^{-1-4K}.$$

Therefore,

$$n^{-4\beta-4\alpha-2+2\epsilon}\sum_{i=1}^{n}\frac{1}{u_i^2 h(u_i)^2}\lesssim n^{4\epsilon}r_n^{-1-4\alpha}+n^{-4\beta-4\alpha+4K+2\epsilon}r_n^{-1-4K}=o(r_n)$$

for ϵ sufficiently small.

(iii): It is straightforward to bound $\sup_{\lambda\in(0,\pi]}\widetilde{h}^{-2}(\lambda)$ by a multiple of $n^{\frac{4K\beta}{K-\alpha-\epsilon/2}}=o(r_n n^{4\beta})$, which immediately implies that the second term has the right order. However, to show the same rate for the first term turns out to be the most difficult part of the proof. Let us shortly remark on that. The crucial point is that although we have good control on the spectral density h, this gives no direct link to entries of the inverse of $D_n(h)$. In contrast to the proofs above, estimating $\langle D_n^{-1}(h)(D_n(S_nf)-T_n(f))\rangle$ by $(\sup_\lambda 1/h(\lambda))^2\|D_n(S_nf)-T_n(f)\|_2^2$ is too rough (cf. also Remark 1). Therefore, we look for a new function, say g, with the properties that $1/(gh)$ behaves like a constant for small λ and $D_n(S_ng)$ is explicitly known. It turns out that $g=f_{1/2+\alpha}$ is a good choice, where $f_{1/2+\alpha}$ denotes the spectral density of a fractional Gaussian noise process with Hurst index $1/2+\alpha\le 1/4$, cf. Lemma 12 for details. Furthermore, $r_{1/2+\alpha}$ denotes the corresponding autocovariance function. From (A.5) and (A.6), we obtain

$$\sup_{\lambda\in\left[\frac{1}{n},\pi\right)}\frac{1}{f_{1/2+\alpha}(\lambda)h(\lambda)}\lesssim n^{2\beta+\epsilon}. \tag{B.18}$$

Define

$$\begin{aligned}
I &:= \langle D_n^{-1}(hf_{1/2+\alpha})(D_n(f_{1/2+\alpha})-D_n(S_nf_{1/2+\alpha}))(D_n(S_nf)-T_n(f))\rangle,\\
II &:= \langle (D_n(S_nf_{1/2+\alpha})-T_n(f_{1/2+\alpha}))(D_n(S_nf)-T_n(f))\rangle,\\
III &:= \langle T_n(f_{1/2+\alpha})(D_n(S_nf)-T_n(f))\rangle
\end{aligned}$$

and note that by Lemma 6 (iii),

$$\begin{aligned}
\frac{1}{4}\langle D_n^{-1}(h)(D_n(S_nf)-T_n(f))\rangle &\le I+\|D_n^{-1}(hf_{1/2+\alpha})\|_\infty^2(II+III)\\
&\lesssim I+n^{4\beta+2\epsilon}(II+III).
\end{aligned} \tag{B.19}$$

In the following we will bound the terms I, II and III, separately.

(I): By (ii) and (iv) of Lemma 12, we know $(r_{1/2+\alpha}(k))_k\in$ GM and that $f_{1/2+\alpha}$ behaves like a multiple of $\lambda^{-2\alpha}$ for $\lambda\downarrow 0$. Using Lemma 8 (i) and (B.18),

$$\sup_{\lambda\in\left[\frac{1}{n},\pi\right)}\frac{|f_{1/2+\alpha}(\lambda)-S_nf_{1/2+\alpha}(\lambda)|}{h(\lambda)f_{1/2+\alpha}(\lambda)}\lesssim n^{2\beta+1+\epsilon}\Big(|r_{\alpha+1/2}(n)|+\sum_{k=n+1}^{\infty}\frac{|r_{\alpha+1/2}(k)|}{k}\Big)\lesssim n^{2\beta+2\alpha+\epsilon}.$$

Thus, with (A.3) and Lemma 6 (iii),

$$I \leq \|D_n^{-1}(hf_{1/2+\alpha})(D_n(f_{1/2+\alpha}) - D_n(S_n f_{1/2+\alpha}))\|_\infty^2 \langle D_n(S_n f) - T_n(f)\rangle \lesssim n^{4\epsilon+4\beta}.$$

(II): By Lemma 12 (i), and the boundedness of the sequence $(r_{1/2+\alpha}(k))_k$, we see that there exists a constant C_α such that for all $k \in \mathbb{N}$, $|r_{1/2+\alpha}(k)| \leq C_\alpha k^{2\alpha-1}$. Using Lemma 7 and (A.4),

$$\begin{aligned} e_{i,j} &:= \Big|\Big[(D_n(S_n f_{1/2+\alpha}) - T_n(S_n f_{1/2+\alpha}))(D_n(S_n f) - T_n(f))\Big]_{i,j}\Big| \\ &\leq C_\alpha C_r \sum_{k=1}^n (|i+k|^{2\alpha-1} + |2n+2-k-i|^{2\alpha-1})(|j+k|^{-2\alpha-1+\delta'} + |2n+2-j-k|^{-2\alpha-1+\delta'}). \end{aligned}$$

Define $F(i,j) := \sum_{k=1}^n |i+k|^{2\alpha-1}|j+k|^{-2\alpha-1}$. It is well-known that if $(a_k)_k$ and $(b_k)_k$ are non-negative sequences which are monotone increasing and decreasing, respectively, then $\sum_{k=1}^n a_k b_k \leq \sum_{k=1}^n a_k b_{n+1-k}$. Thus,

$$e_{i,j} \leq C_\alpha C_r (2n)^{\delta'} \Big(F(i,j) + F(i, n+1-j) + F(n+1-i, j) + F(n+1-i, n+1-j)\Big).$$

From the monotonicity of $x \mapsto x^{-2\alpha-1}$ and $x \mapsto x^{2\alpha-1}$ for $x > 0$,

$$F(i,j) \leq j^{-2\alpha-1} \sum_{k=1}^n |i+k|^{2\alpha-1} \leq j^{-2\alpha-1} \int_i^\infty x^{2\alpha-1} dx = \frac{1}{2|\alpha|} j^{-2\alpha-1} i^{2\alpha}.$$

This allows to bound $e_{i,j}e_{j,i}$ by a multiple of $n^{2\epsilon}(\min(i, n+1-i)\min(j, n+1-j))^{-1}$. Hence, $II \leq \sum_{i,j=1}^n e_{i,j}e_{j,i} \lesssim n^{2\epsilon}\log^2 n$.

(III): First let us introduce the projection $\Pi = \Pi_n$ defined for an $n \times n$ matrix $A = (a_{i,j})_{i,j=1,\dots,n}$ by $(\Pi A)_{i,j} := a_{i,j}$ if $i + j \leq n+1$ and zero otherwise. Further let E denote the $n \times n$ matrix $(E)_{i,j} := 1$ if $i + j = n+1$ and zero otherwise. In particular $E^2 = \mathrm{id}_n$. Note that by Lemma 6 (iv),

$$\begin{aligned} &\langle T_n(f_{1/2+\alpha})(D_n(S_n f) - T_n(f))\rangle \\ &\leq 2\langle T_n(f_{1/2+\alpha})\Pi(D_n(S_n f) - T_n(f))\rangle + 2\langle T_n(f_{1/2+\alpha})(\mathrm{id}_n - \Pi)(D_n(S_n f) - T_n(f))\rangle, \end{aligned}$$

where by a slight abuse of language id_n denotes here the identity operator on the space of

$n \times n$ matrices. To bound the first term, we decompose

$$\begin{aligned}(T_n(f_{1/2+\alpha})\Pi(D_n(S_nf) - T_n(f)))_{i,j} &= \sum_{k=1}^{2i-1} r_{1/2+\alpha}(i-k)[\gamma_{k+j-1} - \gamma_{i+j-1}] \\ &\quad + \gamma_{i+j-1}\sum_{k=1}^{2i-1} r_{1/2+\alpha}(i-k) \\ &\quad + \sum_{k=2i}^{n+1-j} r_{1/2+\alpha}(i-k)\gamma_{k+j-1} \\ &=: A_1(i,j) + A_2(i,j) + A_3(i,j),\end{aligned}$$

with the convention $\sum_{k=2i}^{n+1-j} = -\sum_{k=n+1-j}^{2i}$ if $2i > n+1-j$. By assumption $|\gamma_{q+p} - \gamma_p| \leq \sum_{v=p}^{q+p-1} |\gamma_{v+1} - \gamma_v| \lesssim n^\epsilon qp^{-2\alpha-1}$. Hence, uniformly in i, j,

$$|A_1(i,j)| \lesssim n^\epsilon \sum_{k=1}^{2i-1} \min(|i-k|^{2\alpha}, 1) j^{-2\alpha-2} \lesssim n^\epsilon i^{2\alpha+1} j^{-2\alpha-2}. \tag{B.20}$$

If $2j < i$, we can split the sum $\sum_{k=1}^{2i-1} = \sum_{k=1}^{\lfloor i/2 \rfloor} + \sum_{k=\lfloor i/2 \rfloor+1}^{2i-1}$. Then, the first part of $|A_1(i,j)|$ can be bounded also by

$$\Big|\sum_{k=1}^{\lfloor i/2 \rfloor} r_{1/2+\alpha}(i-k)[\gamma_{k+j-1} - \gamma_{i+j-1}]\Big| \lesssim n^\epsilon i^{2\alpha-1} \sum_{k=1}^{\lfloor i/2 \rfloor} (k+j-1)^{-2\alpha-1} \lesssim n^\epsilon i^{-1}$$

and the second part is by the same arguments as in (B.20) (j can now be replaced by $i+j$) of the order $n^\epsilon i^{2\alpha+1}(i+j-1)^{-2\alpha-2} \leq n^\epsilon i^{-1}$. Together, this shows that

$$|A_1(i,j)| \lesssim \begin{cases} n^\epsilon i^{2\alpha+1} j^{-2\alpha-2} & \text{if } 2j \geq i, \\ n^\epsilon i^{-1} & \text{if } 2j < i. \end{cases}$$

With Lemma 12 (i) and telescoping, $\sum_{k=1}^{2i-1} r_{1/2+\alpha}(i-k) = i^{2\alpha+1} - (i-1)^{2\alpha+1}$ and therefore,

$$|A_2(i,j)| \lesssim n^\epsilon (i+j-1)^{-2\alpha-1} i^{2\alpha}.$$

The last term of the expansion can be simply bounded by

$$|A_3(i,j)| \lesssim n^\epsilon (i+j-1)^{-2\alpha-1} \sum_{k=2i}^{n} |i-k|^{2\alpha-1} \lesssim n^\epsilon (i+j-1)^{-2\alpha-1} i^{2\alpha}.$$

In particular, the bounds for $A_2(i,j)$ and $A_3(i,j)$ are uniformly in i,j as well. Hence by elementary computations

$$
\begin{aligned}
&|\langle T_n(f_{1/2+\alpha})(D_n(S_nf) - T_n(f))\rangle| \\
&\leq \sum_{i,j} |A_1(i,j) + A_2(i,j) + A_3(i,j)||A_1(j,i) + A_2(j,i) + A_3(j,i)| \lesssim n^{2\epsilon}\log^2 n.
\end{aligned}
$$

Finally, note that $ET_n(f_{1/2+\alpha})E = T_n(f_{1/2+\alpha})$ and $E((\mathrm{id}_n - \Pi)[D_n(S_nf) - T_n(f)])E$ is a matrix with entries $-\gamma_{i+j}$ for $i+j \leq n$ and zero otherwise. Therefore, we have by rewriting

$$
\begin{aligned}
&\langle T_n(f_{1/2+\alpha})(\mathrm{id}_n - \Pi)(D_n(S_nf) - T_n(f))\rangle \\
=&\Big\langle T_n(f_{1/2+\alpha})E\Big[(\mathrm{id}_n - \Pi)(D_n(S_nf) - T_n(f))\Big]E\Big\rangle
\end{aligned}
$$

the same structure as above (up to an index shift by one) and all arguments apply. This shows that $III \lesssim n^{2\epsilon}\log^2 n$.

The estimates in $(I), (II)$, and (III) show that the r.h.s. of (B.19) can be upper bounded by $n^{4\epsilon+4\beta}\log^2 n$, and hence Assumption (iii) of Theorem 4 follows by choosing ϵ sufficiently small.

Since we have verified $(i), (ii)$, and (iii) this allows to apply Theorem 4 and therefore the conclusion of the theorem follows. □

B.3 Proof of Theorem 3

We use the superscript $\downarrow 0$ to indicate the approximation of a function at zero, i.e. define for $\lambda \in (0, \pi]$,

$$
f^{\downarrow 0}(\lambda) := C_\alpha \lambda^{2\alpha}\ell(\tfrac{1}{\lambda}), \quad \text{and} \quad \overline{f}^{\downarrow 0}(\lambda) := C_\alpha \lambda^{2\alpha}\ell(u_{r_n}^{-1}), \tag{B.21}
$$

with $C_\alpha = 2\,\mathrm{sign}(-\alpha)\Gamma(-2\alpha)\cos(\pi\alpha)$. Similarly

$$
h^{\downarrow 0}(\lambda) := \sigma^2 f^{\downarrow 0}(\lambda) + \tau^2\lambda^{2K} \quad \text{and} \quad \overline{h}^{\downarrow 0}(\lambda) := \sigma^2 \overline{f}^{\downarrow 0}(\lambda) + \tau^2\lambda^{2K}. \tag{B.22}
$$

In order to prove Theorem 3, we will show in the following propositions that up to negligible terms

$$
\int_0^{u_{r_n^+}} \big(\tfrac{f(\lambda)}{h(\lambda)}\big)^2 d\lambda = \int_0^{u_{r_n^+}} \big(\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)}\big)^2 d\lambda = \int_0^{u_{r_n^+}} \big(\tfrac{\overline{f}^{\downarrow 0}(\lambda)}{\overline{h}^{\downarrow 0}(\lambda)}\big)^2 d\lambda,
$$

where $(r_n^+)_n$ is a generic integer sequence satisfying $r_n \ll r_n^+ \ll n$.

Proposition 1. *Work under the assumptions of Theorem 2. For any non-negative sequence*

$(\nu_n)_n$, *tending to zero,*

$$\int_0^{\nu_n} (\tfrac{f(\lambda)}{h(\lambda)})^2 d\lambda = \int_0^{\nu_n} (\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda (1+o(1)).$$

Proof. By Lemma 11, $f(\lambda) \sim f^{\downarrow 0}(\lambda)$ for $\lambda \downarrow 0$. Since $\nu_n \to 0$, it follows from (A.5), (A.6), and Taylor expansion of $\sin^{2K}(\cdot)$ that

$$q_n := \sup_{\lambda \in (0,\nu_n]} \big|\tfrac{f(\lambda)}{f^{\downarrow 0}(\lambda)} - 1\big| + \big|\tfrac{h(\lambda)}{h^{\downarrow 0}(\lambda)} - 1\big| \le \sup_{\lambda \in (0,\nu_n]} 2\big|\tfrac{f(\lambda)}{f^{\downarrow 0}(\lambda)} - 1\big| + \big|\tfrac{4^K \sin^{2K}(\lambda/2) - \lambda^{2K}}{\lambda^{2K}}\big| = o(1).$$

Therefore, we can estimate

$$\Big|\int_0^{\nu_n} (\tfrac{f(\lambda)}{h(\lambda)})^2 - \int_0^{\nu_n} (\tfrac{f^{\downarrow 0}(\lambda)}{h(\lambda)})^2 d\lambda\Big| \le 3\int_0^{\nu_n} \tfrac{|f(\lambda) - f^{\downarrow 0}(\lambda)|}{(h(\lambda))^2} f(\lambda) d\lambda \le 6 q_n \int_0^{\nu_n} (\tfrac{f(\lambda)}{h(\lambda)})^2 d\lambda$$

and the proof is complete, due to

$$\begin{aligned}\Big|\int_0^{\nu_n} (\tfrac{f^{\downarrow 0}(\lambda)}{h(\lambda)})^2 d\lambda - \int_0^{\nu_n} (\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda\Big| &\le 3\int_0^{\nu_n} \Big|\tfrac{(f^{\downarrow 0}(\lambda))^2 \big(h^{\downarrow 0}(\lambda) - h(\lambda)\big) h^{\downarrow 0}(\lambda)}{(h(\lambda))^2 (h^{\downarrow 0}(\lambda))^2}\Big| d\lambda \\ &\le 12\int_0^{\nu_n} (\tfrac{f(\lambda)}{h(\lambda)})^2 \big|\tfrac{h(\lambda)}{h^{\downarrow 0}(\lambda)} - 1\big| d\lambda \le 12 q_n \int_0^{\nu_n} (\tfrac{f(\lambda)}{h(\lambda)})^2 d\lambda.\end{aligned}$$

□

Proposition 2. *Let $(r_n^+)_n$ be as above. Then,*

$$\int_0^{u_{r_n^+}} (\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda = \int_0^{u_{r_n^+}} (\tfrac{\overline{f}^{\downarrow 0}(\lambda)}{\overline{h}^{\downarrow 0}(\lambda)})^2 d\lambda + o(u_{r_n} n^{4\beta}).$$

Proof. We define $L(\cdot) = (\ell(\cdot))^2$, which is again slowly varying. The lemma is proved, once we have established that $|I + II + III + IV| = o(u_{r_n})$, with

$$\begin{aligned} I &= \int_0^{u_{r_n}} (\tfrac{n^{-2\beta} f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda - \int_0^{u_{r_n}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda, \\ II &= \int_0^{u_{r_n}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda - \int_0^{u_{r_n}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{\overline{h}^{\downarrow 0}(\lambda)})^2 d\lambda, \\ III &= \int_{u_{r_n}}^{u_{r_n}^+} (\tfrac{n^{-2\beta} f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda - \int_{u_{r_n}}^{u_{r_n^+}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda, \\ IV &= \int_{u_{r_n}}^{u_{r_n^+}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)})^2 d\lambda - \int_{u_{r_n}}^{u_{r_n^+}} (\tfrac{n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)}{\overline{h}^{\downarrow 0}(\lambda)})^2 d\lambda. \end{aligned}$$

(I): Note that by substituting $s = u_{r_n} \lambda^{-1}$,

$$|I| \le \sigma^{-4} \int_0^{u_{r_n}} \tfrac{|L(1/\lambda) - L(u_{r_n}^{-1})|}{L(1/\lambda)} d\lambda = u_{r_n} \sigma^{-4} \int_1^{\infty} \tfrac{|L(s u_{r_n}^{-1}) - L(u_{r_n}^{-1})|}{L(s u_{r_n}^{-1})} s^{-2} ds.$$

Since $u_{r_n}^{-1} \to \infty$, we find by Potter's bound that $|L(u_{r_n}^{-1})/L(su_{r_n}^{-1})| \le 2s^{1/2}$, for sufficiently large n and all $s \in [1, \infty)$. Therefore, by dominated convergence and the definition of slowly varying functions,

$$\lim_{n\to\infty} \int_1^\infty \tfrac{|L(su_{r_n}^{-1})-L(u_{r_n}^{-1})|}{L(su_{r_n}^{-1})} s^{-2} ds = \int_1^\infty \lim_{n\to\infty} \tfrac{|L(su_{r_n}^{-1})-L(u_{r_n}^{-1})|}{L(su_{r_n}^{-1})} s^{-2} ds = 0$$

and therefore $I = o(u_{r_n})$.

(II): Let $0 < a_1, a_2, b$. Then,

$$\left|\tfrac{1}{(a_1+b)^2} - \tfrac{1}{(a_2+b)^2}\right| = \left(\tfrac{1}{a_1+b} + \tfrac{1}{a_2+b}\right)\tfrac{|a_1-a_2|}{(a_1+b)(a_2+b)} \le \left(\tfrac{1}{a_1} + \tfrac{1}{a_2}\right)\tfrac{|a_1-a_2|}{a_1 a_2} = \left|\tfrac{1}{a_1^2} - \tfrac{1}{a_2^2}\right|.$$

With $a_1 = \sigma^2 n^{-2\beta} \overline{f}^{\downarrow 0}(\lambda)$, $a_2 = \sigma^2 n^{-2\beta} f^{\downarrow 0}(\lambda)$ and $b = h^{\downarrow 0}(\lambda) - a_2$,

$$|II| \le \sigma^{-4} \int_0^{u_{r_n}} \left|\tfrac{(\overline{f}^{\downarrow 0}(\lambda))^2}{(f^{\downarrow 0}(\lambda))^2} - 1\right| d\lambda = \sigma^{-4} \int_0^{u_{r_n}} \tfrac{|L(1/\lambda)-L(u_{r_n}^{-1})|}{L(1/\lambda)} d\lambda = o(u_{r_n}),$$

where the last step follows from (I).

(III): By estimating $h^{\downarrow 0}(\lambda) \ge \tau^2 \lambda^{2K}$,

$$\begin{aligned}
|III| &\le \int_{u_{r_n}}^{u_{r_n^+}} \left(C\tau^{-2} n^{-2\beta} \lambda^{2\alpha-2K}\right)^2 |L(1/\lambda) - L(u_{r_n}^{-1})| d\lambda \\
&= C^2 \tau^{-4} n^{-4\beta} u_{r_n}^{1+4\alpha-4K} L(u_{r_n}^{-1}) \int_1^{u_{r_n^+}/u_{r_n}} \lambda^{4\alpha-4K} \left|\tfrac{L(u_{r_n}^{-1}/\lambda)-L(u_{r_n}^{-1})}{L(u_{r_n}^{-1})}\right| d\lambda \\
&\lesssim u_{r_n} L(u_{r_n}^{-1}) (L(n^{\frac{\beta}{K-\alpha}}))^{-1} \int_1^{u_{r_n^+}/u_{r_n}} \lambda^{4\alpha-4K} \left|\tfrac{L(u_{r_n}^{-1}/\lambda)-L(u_{r_n}^{-1})}{L(u_{r_n}^{-1})}\right| d\lambda.
\end{aligned}$$

By arguing similar as in (I) and (II), and since $4\alpha - 4K < -1$, we see that the integral tends to zero and therefore, using Lemma 2, $|III| = o(u_{r_n})$ follows.

(IV): Using $(\overline{h}^{\downarrow 0}(\lambda))^{-1} + (h^{\downarrow 0}(\lambda))^{-1} \le \sigma^{-2} n^{2\beta} (\overline{f}^{\downarrow 0}(\lambda))^{-1} + \sigma^{-2} n^{2\beta} (f^{\downarrow 0}(\lambda))^{-1}$ and the elementary inequality $(ab)^{-1} \le a^{-2} + b^{-2}$ for $a, b > 0$, we obtain

$$\begin{aligned}
|IV| &\le \sigma^{-2} n^{2\beta} \int_{u_{r_n}}^{u_{r_n^+}} (\overline{f}^{\downarrow 0}(\lambda))^2 \Big(\tfrac{1}{\overline{f}^{\downarrow 0}(\lambda)} + \tfrac{1}{f^{\downarrow 0}(\lambda)}\Big) \tfrac{|f^{\downarrow 0}(\lambda) - \overline{f}^{\downarrow 0}(\lambda)|}{\overline{h}^{\downarrow 0}(\lambda) h^{\downarrow 0}(\lambda)} d\lambda \\
&\le \sigma^{-2} n^{2\beta} \int_{u_{r_n}}^{u_{r_n^+}} \tfrac{\left(\overline{f}^{\downarrow 0}(\lambda)\right)^2}{\overline{h}^{\downarrow 0}(\lambda) h^{\downarrow 0}(\lambda)} |(f^{\downarrow 0}(\lambda))^2 - (\overline{f}^{\downarrow 0}(\lambda))^2| ((f^{\downarrow 0}(\lambda))^{-2} + (\overline{f}^{\downarrow 0}(\lambda))^{-2}) d\lambda.
\end{aligned}$$

By expanding the last sum and treating each of the two terms separately, we can argue along the lines of (III). $\square$

Proof of Theorem 3. We consider two cases. First assume that $K - \alpha > 1/4$. Let $(r_n^+)_n$ be an integer sequence such that $r_n^+ \ll r_n$. Define

$$C_n := n^{-2\beta}\ell(u_{r_n}^{-1})2\Gamma(-2\alpha)\cos(\pi\alpha) \quad \text{and} \quad C_n' := (C_n^{-1}\sigma^{-2}\tau^2)^{1/(2K-2\alpha)}.$$

Application of Propositions 3, 1, and 2 and substitution show that

$$\begin{aligned}\int_0^\pi \tfrac{f^2(\lambda)}{h^2(\lambda)}d\lambda &= \sigma^{-4}n^{4\beta}\int_0^{u_{r_n^+}} \big(1+C_n^{-1}\sigma^{-2}\tau^2\lambda^{2K-2\alpha}\big)^{-2}d\lambda(1+o(1)) + o(r_n n^{4\beta}) \\ &= (C_n')^{-1}\sigma^{-4}n^{4\beta}\int_0^{u_{r_n^+}/C_n'} (1+\lambda^{2K-2\alpha})^{-2}d\lambda(1+o(1)) + o(r_n n^{4\beta}).\end{aligned} \tag{B.23}$$

Observe that by Lemma 2, $u_{r_n^+}/C_n' \to \infty$. Since $K-\alpha > 1/4$ the integral $\int_0^\infty (1+\lambda^{2K-2\alpha})^{-2}d\lambda$ is finite and we can approximate the r.h.s. of (B.23) further by $\int_0^\infty (1+\lambda^{2K-2\alpha})^{-2}d\lambda(1+o(1))$. By using formula (3.251.6) in [18]

$$\int_0^\infty \Big(\frac{1}{1+\lambda^{2K-2\alpha}}\Big)^2 d\lambda = \frac{1}{K-\alpha}\int_0^\infty \frac{\lambda^{\frac{1}{K-\alpha}-1}}{(1+\lambda^2)^2}d\lambda = \frac{\pi(\frac{1}{K-\alpha}-2)}{4(K-\alpha)\sin(\pi(\frac{1}{K-\alpha}-2)/2)}$$

and for $K-\alpha = 1/2$ the r.h.s. is defined by 1. Combining the last identities (recall that for the Fisher information, the integrals are scaled by $n^{1-4\beta}/(2\pi)$) and application of Lemma 2 completes the proof for the first case.

Now, let us prove the second statement of Theorem 3, i.e. the case $K-\alpha < 1/4$, i.e. $K=0$ and $\alpha \in (-1/4, -\beta)$. With the upper bound in (A.5) (extended for all $\lambda \in (0,\pi]$), we have $f(\lambda) \lesssim \lambda^{2\alpha-\epsilon}$ and $\int_0^{n^{\beta/\alpha}} f^2(\lambda)(h^{-2}(\lambda)+1)d\lambda = o(1)$. Since, $1/h - \tau^{-2} = \sigma^2\tau^{-2}n^{-2\beta}f/h$ we obtain also

$$\int_{n^{\frac{\beta}{\alpha}}}^\pi \Big|\tfrac{f^2(\lambda)}{h^2(\lambda)} - \tau^{-4}f^2(\lambda)\Big|d\lambda \le 2\sigma^2\tau^{-6}n^{-2\beta}\int_{n^{\frac{\beta}{\alpha}}}^\pi f^3(\lambda)d\lambda \lesssim n^{-2\beta}\int_{n^{\frac{\beta}{\alpha}}}^\pi \lambda^{6\alpha-3\epsilon}d\lambda = o(1)$$

for small ϵ. Combining the last approximations yields

$$I_{\sigma^2}^n = \frac{n^{1-4\beta}}{2\pi\tau^4}\int_0^\pi f^2(\lambda)d\lambda + o(n^{1-4\beta}) = \frac{n^{1-4\beta}}{2\tau^4}\sum_{k=-\infty}^\infty \gamma_k^2 + o(n^{1-4\beta})$$

since $f = \sum_{k=-\infty}^\infty \cos(k\cdot)\gamma_k$. This completes the proof of (2.14).

Finally, we evaluate the integral in the critical case, i.e. we prove (2.15). Because of $\log(x)\ell^2(x) \to \infty$ for $x \to \infty$, one can find a non-negative sequence $(\nu_n)_n$, tending to zero, such that

$$\int_{\nu_n}^\pi \big(\tfrac{f(\lambda)}{h(\lambda)}\big)^2 d\lambda \lesssim \int_{\nu_n}^\pi \lambda^{-1-2\epsilon}d\lambda \lesssim \nu_n^{-2\epsilon} = o\big(\log(n)\ell^2(n^{4\beta})\big), \tag{B.24}$$

where we used the global upper bound for f, cf. (A.5), in the first inequality. Since ν_n tends to zero, Proposition 1 applies. With $f^{\downarrow 0}$ and $h^{\downarrow 0}$ as defined in (B.21) and (B.22),

$$\begin{aligned}
I^n_{\sigma^2} &= \frac{n^{1-4\beta}}{2\pi}\int_0^{\nu_n} \big(\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)}\big)^2 d\lambda(1+o(1)) + o\big(n^{1-4\beta}\log(n)\ell^2(n^{4\beta})\big) \\
&= \frac{n^{1-4\beta}}{2\pi}\int_{q_n}^{\nu_n} \big(\tfrac{f^{\downarrow 0}(\lambda)}{h^{\downarrow 0}(\lambda)}\big)^2 d\lambda(1+o(1)) + o\big(n^{1-4\beta}\log(n)\ell^2(n^{4\beta})\big).
\end{aligned}$$

Note that $K-\alpha = 1/4$ implies $\alpha = -1/4$. Observe also that by Potter's bound,

$$\begin{aligned}
\int_{q_n}^{\nu_n} \big(f^{\downarrow 0}(\lambda)\big)^2 \big|\big(h^{\downarrow 0}(\lambda)\big)^{-2} - \tau^{-4}\big| d\lambda &\lesssim n^{-2\beta}\int_{q_n}^{\nu_n} \big(f^{\downarrow 0}(\lambda)\big)^3 d\lambda \\
&= n^{-2\beta}\ell^3(q_n^{-1})q_n^{-1/2}\int_1^{\nu_n/q_n} \lambda^{-3/2}\frac{\ell^3(\frac{1}{\lambda q_n})}{\ell^3(\frac{1}{q_n})}d\lambda = O\big(\ell^2(n^{4\beta})\big).
\end{aligned}$$

For $\alpha=-1/4$, $f^{\downarrow 0}(\lambda) = \sqrt{2\pi}\lambda^{-1/2}\ell(1/\lambda)$. This shows

$$I^n_{\sigma^2} = n^{1-4\beta}\tau^{-4}\int_{q_n}^1 \ell^2\big(\tfrac{1}{\lambda}\big)\frac{d\lambda}{\lambda}(1+o(1)) + o\big(n^{1-4\beta}\log(n)\ell^2(n^{4\beta})\big),$$

where the integral $\int_{\nu_n}^1$ is estimated as in (B.24). □

B.4 Proof of Corollary 1

Proof. For $H=1/2$ the result follows from [16]. If $H\neq 1/2$, we check the conditions of Theorem 2. Note that the three versions correspond to $0<H<1/2$, $1/2<H<3/4$, and $3/4\leq H<1$. By using the properties of fGN, listed in Lemma 12 (for instance in order to show $\sum_{k=-\infty}^{\infty}\gamma_k = 0$ for $0<H<1/2$, use Lemma 12 (i)), Versions 1 and 2 are obviously fulfilled, but some work has to be done for the third part.

To bound the increments of the autocovariance function, i.e. $|\gamma_{p+1}-\gamma_p| = |r_H(p+1)-r_H(p)| \lesssim r_H(p)/p = \gamma_p/p$, define on $\mathbb{R}^+$, $x\mapsto g(x) = |x|^{2H}$. Then there exists $\xi_1\in[x,x+1]$ and $\xi_2\in[x-1,x]$ such that by Taylor expansion $g(x+1) = \sum_{k=0}^3 (k!)^{-1}g^{(k)}(x) + 24^{-1}g^{(4)}(\xi_1)$ and $g(x-1) = \sum_{k=0}^3(-1)^k(k!)^{-1}g^{(k)}(x) + 24^{-1}g^{(4)}(\xi_2)$. Thus,

$$r_H(q) = g''(q) + \frac{1}{24}\big(g^{(4)}(\xi_1)+g^{(4)}(\xi_2)\big).$$

By another Taylor expansion and $p\geq 1$, $|r_H(p+1)-r_H(p)|\lesssim p^{2H-3}$. □

Appendix C Further technicalities

First, let us summarize some facts about Frobenius norms.

Lemma 6. *Given $n \times n$ matrices $A = (a_{i,j})_{i,j=1,\dots,n}, B, C$. Let $\langle A \rangle = \operatorname{tr}(A^2)$ and denote by $\|A\|_2 = (\sum_{i,j} a_{i,j}^2)^{1/2}$ the Frobenius (or Hilbert-Schmidt) norm. Then,*

(i) $\langle A \rangle \leq \|A\|_2^2$, *and equality holds if A is symmetric.*

(ii) *A and B are circular commuting, i.e.* $\langle AB \rangle = \langle BA \rangle$

(iii) *If $A \geq 0$ and B Hermitian, then,* $\langle AB \rangle \leq \|A\|_\infty^2 \langle B \rangle$.

(iv) *If $C \geq 0$ as well as A and B Hermitian, then*

$$\langle C(A+B) \rangle = \|C^{1/2}(A+B)C^{1/2}\|_2^2 \leq 2\langle CA \rangle + 2\langle CB \rangle.$$

Proof. *(i):* Due to $2uv \leq u^2 + v^2$, $\operatorname{tr}(A^2) = \big(\sum_{i,j} a_{i,j}a_{j,i}\big)^2 \leq \sum_{i,j} a_{i,j}^2 = \|A\|_2^2$. *(iii):* Write $A = D^t \Lambda D$ with Λ diagonal. ☐

The next result is well-known, cf. Sánchez et *al.* [25], p. 2635. For sake of completeness we give a short proof.

Lemma 7. *Assume that $g = \sum_{m=-n}^n g_m \cos(m\cdot)$ with $g_m = g_{-m}$ and let $D_n(g)$ be defined as in* (A.3). *Then, we have for the (i,j)-th entry $(D_n(g))_{i,j} = g_{|i-j|} + g_{i+j-1}$ if $i + j \leq n+1$ and $(D_n(g))_{i,j} = g_{|i-j|} - g_{2n+2-i-j}$ if $i+j > n+1$.*

Proof. Recall (A.1) and (A.3). Expand

$$\begin{aligned}(D_n(g))_{ij} = \tfrac{1}{2n+1} \sum_{m=-n}^{n} g_m \sum_{k=1}^{n} \big[&\cos((i+j-1+m)\,u_k) + \cos((i+j-1-m)\,u_k) \\ &+ \cos((i-j+m)\,u_k) + \cos((i-j-m)\,u_k) \big]\end{aligned}$$

using $g(u_k) = \sum_{m=-n}^n g_m \cos(mu_k)$ and $2\cos(x)\cos(y) = \cos(x+y) + \cos(x-y)$. With the well-known summation formula for Dirichlet kernels, we find that for an integer ℓ, $\sum_{k=1}^n \cos(\ell u_k)$ equals $n(-1)^{\ell/(2n+1)}$, if ℓ is divisible by $2n+1$, and $\frac{1}{2}(-1)^{\ell+1}$ otherwise. If we rewrite in the first case $n(-1)^{\ell/(2n+1)}$ as $(n+\frac{1}{2})(-1)^{\ell/(2n+1)} + \frac{1}{2}(-1)^{\ell+1}$, we find that $(D_n(g))_{i,j} = g_{|i-j|} + g_{i+j-1}$ if $i+j \leq n+1$ and $(D_n(g))_{i,j} = g_{|i-j|} - g_{2n+2-i-j}$ if $i+j > n+1$. ☐

In the next lemma, we collect a number of properties of general monotone sequences.

Lemma 8. *Assume that $(a_k)_k \in \mathrm{GM}$ and define $f = a_0 + 2\sum_{k=1}^\infty a_k \cos(k\cdot)$.*

(i) *If $\sum_{k=1}^\infty k^{-1}|a_k| < \infty$, then there exists a constant C such that*

$$|f(x) - S_n f(x)| \leq C \frac{1}{x}\Big(|a_n| + \sum_{k=n+1}^\infty \frac{|a_k|}{k}\Big), \quad \text{if } \; x \in [\tfrac{1}{n}, \pi].$$

(ii) *If additionally* $a_k \sim k^{-2\alpha-1}\ell(k)$ *holds for* $\alpha > -1/2$ *and a slowly varying function* ℓ, *then* $(a_k)_k$ *is of bounded variation, i.e.* $\sum_{k=1}^{\infty} |a_{k+1} - a_k| < \infty$.

(iii) *If* $a_k \sim k^{-2\alpha-1}\ell(k)$ *holds for* $\alpha \in (-1/2, 0)$ *and a quasi-monotone slowly varying function* ℓ, *then*

$$f(\lambda) \sim 2\lambda^{2\alpha}\ell(\tfrac{1}{\lambda})\Gamma(-2\alpha)\cos(\pi\alpha).$$

Proof. (i) and (iii) can be found in the recent paper by Tikhonov [27]. To verify (ii) note that for sufficiently large K, using the definition of GM and (A.4), $\sum_{k=K}^{\infty} |a_{k+1} - a_k| \leq \sum_{q=0}^{\infty} |a_{2^\ell K}| < \infty$. □

The following lemma gives a very useful bound on the Frobenius norm between the matrices $T_n(f)$ and $D_n(f)$ defined in (A.2) and (A.3).

Lemma 9. *Let* $f = \sum_{k=0}^{\infty} a_k \cos(k\cdot)$ *and assume that* $(a_k) \in \mathrm{GM}$. *If* $\sum_{k\geq 1} k^{-1}|a_k| < \infty$, *then, there exists a constant* C, *such that*

$$\|D_n(f) - T_n(f)\|_2 \leq Cn\Big(|a_n| + \sum_{k=n+1}^{\infty} \frac{|a_k|}{k}\Big) + (2\sum_{k=1}^{n} ka_k^2)^{1/2}.$$

Proof. Note that $\|D_n(f) - T_n(f)\|_2 \leq \|D_n(f) - D_n(S_n f)\|_2 + \|D_n(S_n f) - T_n(f)\|_2$. In order to bound the first part, we use the estimate from Lemma 8 (i) together with (A.3) and hence $\|D_n(f) - D_n(S_n f)\|_2^2 = \sum_{k=1}^{n}(f(u_k) - S_n f(u_k))^2$ for u_k as in (A.1). Since $u_k \geq 1/n$, this yields

$$\|D_n(f) - D_n(S_n f)\|_2 \leq C_1 n\Big(|a_n| + \sum_{k=n+1}^{\infty} \frac{|a_k|}{k}\Big)$$

for another constant C_1. Using Lemma 7 and $(T_n(f))_{i,j} = f_{|i-j|}$, the second term can be bounded easily by $(2\sum_{k=1}^{n} kf_k^2)^{1/2}$. □

Lemma 10. *Assume that* $(\gamma_k)_k$ *satisfies the assumptions of Theorem 2, Version 3. Then it is also in* GM.

Proof. The following arguments are for J large enough. Observe that by Potter's bound, the sequence $(b_J)_J$,

$$b_J := \frac{1}{J}\sum_{j=J}^{2J-1} \tfrac{\ell(j)}{\ell(J)}\big(\tfrac{j}{J}\big)^{-2\alpha-2}$$

is uniformly bounded. Using also (A.4)

$$\sum_{j=J}^{2J-1} |\gamma_{j+1}-\gamma_j| \le C_1 \sum_{j=J}^{2J-1} \frac{\gamma_j}{j} \lesssim C_1 b_J \ell(J) J^{-2\alpha-1} \lesssim \gamma_J.$$

□

Lemma 11. *Under the assumptions of Theorem 2, we have that f is bounded on any interval $[\delta,\pi]$ with $\delta>0$ and*

$$f(\lambda) \sim 2\,\mathrm{sign}(-\alpha)\Gamma(-2\alpha)\cos(\pi\alpha)\lambda^{2\alpha}\ell(1/\lambda).$$

Proof. For $\alpha\in(0,1/2)$, the sequence $(\gamma_k)_k$ is absolutely summable, and hence f is bounded on $[0,\pi]$. The behavior at zero, i.e. $f(\lambda)\sim -\lambda^{2\alpha}\ell(1/\lambda)2\Gamma(-2\alpha)\cos(\pi\alpha)$, follows from Theorem 1.2 in [8] and the assumption $\sum_{k=-\infty}^{\infty}\gamma_k = 0$. For $\alpha\in(-1/2,0)$, we have $(\gamma_k)_k \in \mathrm{GM}$ (cf. Lemma 10). Hence, by Lemma 8 (ii), $(\gamma_k)_k$ is of bounded variation and therefore (cf. Zygmund [29], Theorem 2.6) f is bounded on $[\delta,\pi]$ for every $\delta>0$. The second property is a consequence of Lemma 8 (iii) and Assumption 1. □

Proposition 3. *Work under the assumptions of Theorem 4 with $K-\alpha>1/4$ and let $(r_n^+)_n$ be an integer sequence satisfying $r_n \ll r_n^+ \ll n$. Then,*

$$n\int_{u_{r_n^+}}^{\pi} \big(\tfrac{f(\lambda)}{h(\lambda)}\big)^2 d\lambda = o(r_n n^{4\beta}) \quad \textit{and} \quad \sum_{i=r_n^+ +1}^{n} \big(\tfrac{f(u_i)}{h(u_i)}\big)^2 = o(r_n n^{4\beta}).$$

Proof. Because of $K-\alpha>1/4$, there exist $\delta'>0$ and an integer n_0 such that by Potter's bound we find for all n larger than n_0,

$$\frac{\ell(\frac{1}{\mu u_{r_n^+}})}{\ell(\frac{1}{u_{r_n^+}})} \le 2\mu^{K-\alpha-1/4}, \quad \text{for all } \ \mu\in(0,\delta/u_{r_n^+}].$$

Let δ be as in (A.5). Together with (A.6), we obtain

$$\begin{aligned}\frac{f(u_{r_n^+}\mu)}{h(u_{r_n^+}\mu)} &\le 2\cdot 4^K \tau^{-2} u_{r_n^+}^{2\alpha-2K}\mu^{2\alpha-2K}\ell(\tfrac{1}{\mu u_{r_n^+}})\\ &\le 4^{K+1}\tau^{-2}u_{r_n^+}^{2\alpha-2K}\mu^{\alpha-K-1/4}\ell(\tfrac{1}{u_{r_n^+}}), \quad \text{for all } \ \mu\in(0,(\delta\wedge\delta')/u_{r_n^+}].\end{aligned}$$

Since f is bounded on $[\delta\wedge\delta',\pi]$ under the assumptions of Theorem 2 (cf. Lemma 11), we also have $f(\lambda)/h(\lambda)\lesssim \lambda^{-2K}$, for all $\lambda\in(\delta\wedge\delta',\pi]$. Now, we are ready to prove the two statements (denoted by (I) and (II)) of the proposition.

(I): Decomposing $\int_{u_{r_n^+}}^{\pi} = \int_{u_{r_n^+}}^{\delta\wedge\delta'} + \int_{\delta\wedge\delta'}^{\pi}$ and substituting $\mu = \lambda/u_{r_n^+}$ in the first integral yields

$$\int_{u_{r_n^+}}^{\pi} \big(\tfrac{f(\lambda)}{h(\lambda)}\big)^2 d\lambda = u_{r_n^+} \int_1^{(\delta\wedge\delta')/u_{r_n^+}} \Big(\frac{f(u_{r_n^+}\mu)}{h(u_{r_n^+}\mu)}\Big)^2 d\mu + \int_{\delta\wedge\delta'}^{\pi} \big(\tfrac{f(\lambda)}{h(\lambda)}\big)^2 d\lambda$$
$$\leq 4^{2K+2}\tau^{-4} u_{r_n^+}^{4\alpha-4K+1} \ell^2\big(\tfrac{1}{u_{r_n^+}}\big) \int_1^{\infty} \mu^{2\alpha-2K-1/2} d\mu + o(u_{r_n} n^{4\beta}).$$

Note that for two positive sequences $(a_n)_n$ and $(b_n)_n$ with $a_n \ll b_n$, and any $\kappa > 0$, we have for sufficiently large n,

$$\tfrac{\ell(a_n)}{\ell(b_n)} \leq 2\big(\tfrac{b_n}{a_n}\big)^{\kappa/2} \ll \big(\tfrac{b_n}{a_n}\big)^{\kappa}.$$

Apply this for $a_n = 1/u_{r_n^+}$ and $b_n = 1/u_{r_n}$ together with Lemma 2. Then,

$$u_{r_n^+}^{4(\alpha-K)+1} \ell^2\big(\tfrac{1}{u_{r_n^+}}\big) \ll u_{r_n}^{4(\alpha-K)+1} \ell^2\big(\tfrac{1}{u_{r_n}}\big) = O(u_{r_n} n^{4\beta})$$

and this completes the proof of (I).

(II): Let κ_n be the largest integer, s.t. $u_{\kappa_n} \leq \delta \wedge \delta'$. Similarly as for (I), we rewrite

$$\sum_{i=r_n^+ +1}^{n} \big(\tfrac{f(u_i)}{h(u_i)}\big)^2 \leq 4^{2K+2}\tau^{-4} u_{r_n^+}^{4\alpha-4K} \ell^2\big(\tfrac{1}{u_{r_n^+}}\big) \sum_{i=r_n^+ +1}^{\kappa_n} \big(\tfrac{u_i}{u_{r_n^+}}\big)^{2\alpha-2K-1/2} + o(r_n n^{4\beta}).$$

Note that

$$\tfrac{1}{r_n^+} \sum_{i=r_n^+ +1}^{\kappa_n} \big(\tfrac{u_i}{u_{r_n^+}}\big)^{2\alpha-2K-1/2} \leq \sum_{i=r_n^+ +1}^{\infty} \big(\tfrac{2i}{r_n^+}\big)^{2\alpha-2K-1/2} \leq \int_1^{\infty} (2\lambda)^{2\alpha-2K-1/2} d\lambda < \infty.$$

Arguing as above yields the required result. □

Finally, we summarizes a number of facts on fractional Gaussian noise.

Lemma 12. *Let $(N_k^H)_{k\in\mathbb{N}}$ denote a fractional Gaussian noise process, i.e. a centered Gaussian process with covariance*

$$\mathrm{Cov}(N_k^H, N_j^H) = \frac{1}{2}\big(|k-j+1|^{2H} - 2|k-j|^{2H} + |k-j-1|^{2H}\big).$$

Denote by f_H and r_H the spectral density and the autocovariance function, respectively. Then,

- *(i)* $r_H(k) = \frac{1}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}) \sim 2H(2H-1)k^{2H-2}$.
- *(ii)* $(r_H(k))_{k\in\mathbb{Z}}$ *is O-regularly varying quasi-monotone and (therefore also) in* GM
- *(iii)* $f_H(\lambda) = 2\sin(\pi H)\Gamma(2H+1)(1-\cos\lambda)\sum_{j=-\infty}^{\infty} |\lambda+2\pi j|^{-2H-1} \sim \sin(\pi H)\Gamma(2H+1)\lambda^{1-2H}$

(iv) There exist positive constants C_H *and* C'_H*, such that* $C_H\lambda^{1-2H} \leq f_H(\lambda) \leq C'_H\lambda^{1-2H}$ *for all* $\lambda \in (0, \pi]$.

Proof. (i), (iii) and (iv) are well-known (cf. [3]). (ii) follows from the monotone decrease of $(r_H(k))_k$. □

References

[1] O. Barndorff-Nielsen, P. Hansen, A. Lunde, and N. Shephard. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008.

[2] A. S. Belov. Remarks on the convergence (boundedness) in the mean of partial sums of a trigonometric series. *Math. Notes*, 71:739–748, 2002.

[3] J. Beran. *Statistics for Long-Memory Processes.* Chapman & Hall, New York, 1994.

[4] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation.* Cambridge University Press, Cambridge, 1987.

[5] R. Bojanić and E. Seneta. Slowly varying functions and asymptotic relations. *J. Math. Anal.*, 34:302–315, 1971.

[6] V. Britanak, P. C. Yip, and K. R. Rao. *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximation.* Elsevier Academic Press, Oxford, 2007.

[7] T. Cai, A. Munk, and J. Schmidt-Hieber. Sharp minimax estimation of the variance of Brownian motion corrupted with Gaussian noise. *Statist. Sinica*, 20:1011–1024, 2010.

[8] C.-P. Chen and L. Chen. Asymptotic behavior of trigonometric series with O-regularly varying quasimonotone coefficients. *J. Math. Anal. Appl.*, 250:13–26, 2000.

[9] R. Dahlhaus. Efficient parameter estimation for self-similar processes. *Ann. Statist.*, 17(4):1749–1766, 1989.

[10] R. Davies. Asymptotic inference in stationary Gaussian time-series. *Adv. in Appl. Probab.*, 5:469–497, 1973.

[11] K. Dzhaparidze. *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.* Springer, New York, 1986.

[12] R. Fox and M. Taqqu. Large-Sample properties of parameter estimates for strongly dependent stationary time series. *Ann. Statist.*, 14(2):517–532, 1986.

[13] L. Giraitis and D. Surgailis. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probab. Th. Rel. Fiels*, 86:87–104, 1990.

[14] A. Gloter and M. Hoffmann. Stochastic volatility and fractional Brownian motion. *Stochastic Process. Appl.*, 113:143–172, 2004.

[15] A. Gloter and M. Hoffmann. Estimation of the Hurst parameter from discrete noisy data. *Ann. Statist.*, 35(5):1947–1974, 2007.

[16] A. Gloter and J. Jacod. Diffusions with measurement errors. i. local asymptotic normality. *ESAIM Probab. Stat.*, 5:225–242, 2001.

[17] A. Gloter and J. Jacod. Diffusions with measurement errors. ii. optimal estimators. *ESAIM Probab. Stat.*, 5:243–260, 2001.

[18] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Elsevier Academic Press, Burlington, 2007.

[19] J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.*, 119(7):2249–2276, 2009.

[20] I. Johnstone. Wavelet shrinkage for correlated data and inverse problems: adaptivity results. *Statist. Sinica*, 9:51–83, 1999.

[21] M. Podolskij and M. Vetter. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15:634–658, 2009.

[22] B. Porat and B. Friedlander. Computation of the exact information matrix of Gaussian time series with stationary random components. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(1):118–130, 1986.

[23] M. Reiß. Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Appl. Probab.*, 39:772–802, 2011.

[24] A. Rohde and L. Dümbgen. Statistical inference for the optimal approximation model. *Probab. Theory Related Fields, to appear*, 2012.

[25] V. Sánchez, P. García, A. Peinado, J. Segura, and A. Rubio. Diagonalizing properties of the discrete cosine transform. *IEEE Trans. Signal Process.*, 43:2631–2641, 1995.

[26] M. Stein. Minimum norm quadratic estimation of spatial variograms. *J. Amer. Statist. Assoc.*, 82(399):765–772, 1987.

[27] S. Tikhonov. Trigonometric series with general monotone coefficients. *J. Math. Anal. Appl.*, 326:721–735, 2007.

[28] L. Zhang. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12:1019–1043, 2006.

[29] A. Zygmund. *Trigonometric Series*. Cambridge University Press, Cambridge, 2002.