Electronic Journal of Statistics
Vol. 17 (2023) 2343–2385
ISSN: 1935-7524
https://doi.org/10.1214/23-EJS2156

Estimation of the Hurst parameter from continuous noisy data

Pavel Chigansky

Department of Statistics and Data Science
The Hebrew University of Jerusalem, Israel
e-mail: Pavel.Chigansky@mail.huji.ac.il

Marina Kleptsyna

Laboratoire de Statistique et Processus
Le Mans Université, France
e-mail: Marina.Kleptsyna@univ-lemans.fr

Abstract: This paper addresses the problem of estimating the Hurst exponent of the fractional Brownian motion from continuous time noisy sample. When the Hurst parameter is greater than 3/4, consistent estimation is possible only if either the length of the observation interval increases to infinity or intensity of the noise decreases to zero. The main result is a proof of the Local Asymptotic Normality (LAN) of the model in these two regimes which reveals the optimal minimax estimation rates.

MSC2020 subject classifications: Primary 60G22, 62F12.
Keywords and phrases: Fractional Brownian motion, Local Asymptotic Normality, Hurst parameter estimation.

Received June 2022.

Contents

arXiv: 2205.11092

1. Introduction

Estimation of the Hurst parameter is an old problem in statistics of time series. A benchmark model is the fractional Brownian motion (fBm) $B^H = (B^H_t,\ t \in \mathbb{R}_+)$, that is, the centered Gaussian process with covariance function

$$\mathsf{E} B^H_t B^H_s = \tfrac{1}{2} t^{2H} + \tfrac{1}{2} s^{2H} - \tfrac{1}{2}|t-s|^{2H}$$

where $H \in (0,1)$ is its Hurst exponent. The fBm is a well studied stochastic process with a variety of interesting and useful properties, see, e.g., [34, 16]. Its increments are stationary and, for $H > \frac{1}{2}$, positively correlated with the long-range dependence

$$\sum_{n=1}^{\infty} \mathsf{E} B^H_1 (B^H_n - B^H_{n-1}) = \infty.$$

It is this feature which makes the fBm relevant to statistical modeling in many applications.

A basic problem is to estimate the Hurst parameter $H \in (0,1)$ and the additional scaling parameter $\sigma^2 \in \mathbb{R}_+$ given the data

$$X^T := (\sigma B^H_t,\ t \in [0,T]).$$

Since both parameters can be recovered from X^T exactly for any $T > 0$, a meaningful statistical problem is to estimate them from the discretized data

$$X^{T,\Delta} := \big(\sigma B^H_\Delta, \dots, \sigma B^H_{n\Delta}\big) \tag{1.1}$$

where $\Delta > 0$ is the discretization step and $n = [T/\Delta]$. The two relevant regimes, in which consistent estimation from (1.1) is feasible, are the large time asymptotics with a fixed $\Delta > 0$ and $T \to \infty$, and the high frequency asymptotics with $\Delta \to 0$ and a fixed $T > 0$.

Often it is more realistic to consider the partially observed setup in which the trajectory is contaminated by additive noise. One possibility is to assume that the noise is added after discretization so that the available data is given by

$$X^{T,\Delta} := \big(\sigma B^H_\Delta + \xi_{1,n},\ \dots,\ \sigma B^H_{n\Delta} + \xi_{n,n}\big), \tag{1.2}$$

where $\xi_{j,n}$ are i.i.d. random variables independent of B^H. Such observation scenario fits the situation when a signal, e.g., the position of a particle or a stock price, is measured periodically by a noisy sensor. Statistical properties of this model are relatively well understood (see some details in Sect. 4.2 below).

Another possibility is to assume that the noise is added directly in continuous time. In this case a natural setup to consider is the mixture of the fBm with an independent standard Brownian motion B_t:

$$X^T := \big(\sigma B^H_t + \sqrt{\varepsilon} B_t, t \in [0,T]\big) \tag{1.3}$$

where $\varepsilon > 0$ is the known noise intensity. The formal derivative of this process is the basic noise model in engineering applications, such as astronomical data processing [45], GPS communications [31], analysis of seismic data [22]. The estimation problem for the observation model (1.3) corresponds to calibration of the parameters of its fractional component [1, 2, 44, 3, 4].

From the statistical standpoint, a peculiar feature of the process (1.3), called the *mixed fBm* in the probabilistic literature, is that consistent estimation in the high frequency regime is possible only for $H \leqslant 3/4$, see [13]. It was shown in [7] that for $H > 3/4$ the probability measures induced by the the process (1.3) and the Brownian motion $\sqrt{\varepsilon}B$ are mutually absolutely continuous. This implies that the parameters in question cannot be recovered exactly from the sample X^T for any finite T and, a fortiori, from its discretization.

In this paper we consider estimation problem for the parameters H and σ^2 from the sample (1.3) when $H \in (3/4, 1)$. Our objective is to identify the best achievable minimax rates in the large time ($\varepsilon > 0$ is fixed and $T \to \infty$) and small noise ($T < \infty$ is fixed and $\varepsilon \to 0$) asymptotic regimes. To this end, we prove the Local Asymptotic Normality (LAN) property in both cases and discuss the construction of the rate optimal estimators.

The rest of the paper is organized as follows. Section 2 outlines the essential background needed to formulate the main results in Sect. 3. The results are discussed and compared to the relevant literature in Sect. 4. The proofs appear in Sects. 5–7.

2. The LAN property and Hájek's bound

Let us briefly recall Le Cam's LAN property and its role in the asymptotic theory of estimation. A comprehensive account on the subject can be found in, e.g., [21]. An abstract parametric statistical experiment consists of a measurable space $(\mathcal{X}, \mathcal{A})$, where $\mathcal{A}$ is a σ-algebra of subsets of $\mathcal{X}$, a family of probability measures $(\mathsf{P}_\theta)_{\theta\in\Theta}$ on $\mathcal{A}$ with the parameter space $\Theta \subseteq \mathbb{R}^k$ and the sample $X \sim \mathsf{P}_{\theta_0}$ for some fixed true unknown value $\theta_0 \in \Theta$ of the parameter variable. Asymptotic theory is concerned with a family of statistical experiments $(\mathcal{X}^h, \mathcal{A}^h, (\mathsf{P}^h_\theta)_{\theta\in\Theta})$ indexed by a real valued variable $h > 0$.

Definition 2.1. A family of probability measures $(\mathsf{P}^h_\theta)_{\theta\in\Theta}$ is Locally Asymptotically Normal (LAN) at a point θ_0 as $h \to 0$ if there exist nonsingular $k \times k$ matrices $\phi(h) = \phi(h, \theta_0)$ such that, for any $u \in \mathbb{R}^k$, the Radon-Nikodym derivatives (likelihood ratios) satisfy the scaling property

$$\log \frac{d\mathsf{P}^h_{\theta_0+\phi(h)u}}{d\mathsf{P}^h_{\theta_0}}(X^h) = u^\top Z_{h,\theta_0} - \tfrac{1}{2}\|u\|^2 + r_h(u, \theta_0) \tag{2.1}$$

where the random vector Z_{h,θ_0} converges weakly under $\mathsf{P}^h_{\theta_0}$ to the standard normal law on $\mathbb{R}^k$ and $r_h(u, \theta_0)$ vanishes in $\mathsf{P}^h_{\theta_0}$-probability as $h \to 0$.

Define a set $W_{2,k}$ of loss functions $\ell : \mathbb{R}^k \mapsto \mathbb{R}_+$, which are continuous and symmetric with $\ell(0) = 0$, have convex sub-level sets $\{u : \ell(u) < c\}$ for all $c > 0$ and satisfy the growth condition $\lim_{\|u\|\to 0} \exp(-a\|u\|^2)\ell(u) = 0$, $\forall a > 0$. The following theorem establishes asymptotic lower bound for the corresponding local minimax risks of estimators in LAN families.

Theorem 2.2 (Hájek)**.** *Let $(\mathsf{P}^h_\theta)_{\theta\in\Theta}$ satisfy the LAN property at θ_0 with matrices $\phi(h, \theta_0) \to 0$ as $h \to 0$. Then for any family of estimators $\widehat{\theta}_h$, a loss function $\ell \in W_{2,k}$ and any $\delta > 0$,*

$$\varliminf_{h\to 0} \sup_{\|\theta-\theta_0\|<\delta} \mathsf{E}^h_\theta \ell\big(\phi(h, \theta_0)^{-1}(\widehat{\theta}_h - \theta)\big) \geqslant \int_{\mathbb{R}^k} \ell(x)\gamma_k(x)dx,$$

where γ_k is the standard normal density on $\mathbb{R}^k$.

Proof. [21, Theorem 12.1]. □

Estimators which achieve Hájek's lower bound are called asymptotically efficient in the local minimax sense. Usually likelihood based estimators, such as the Maximum Likelihood or the Bayes estimators with positive prior densities, are asymptotically efficient. However, they can be excessively complicated and thus it often makes sense to construct simpler estimators, which are at least rate optimal. In complex models this is sometimes done separately for each component of the parameter vector, following some ad-hoc heuristics. Proving rate optimality of the obtained estimators requires finding the best minimax rates for each entry of the parameter vector.

Let us explain how such entrywise rates can be derived using the bound of Theorem 2.2. Analysis of the likelihood ratio in (2.1) typically shows that in LAN families $\phi(h, \theta_0)$ must satisfy the condition

$$\phi(h, \theta_0)^\top M(h, \theta_0) I(\theta_0) M(h, \theta_0)^\top \phi(h, \theta_0) \xrightarrow[h\to 0]{} \mathrm{Id}, \tag{2.2}$$

where the matrices $M(h, \theta_0)$ and $I(\theta_0)$ are determined by the statistical model under consideration. The matrix $I(\theta_0)$ is positive definite and independent of

h, and it can be often regarded as the analog of the usual Fisher information matrix.

Consider the Cholesky decomposition

$$L(h,\theta_0)L(h,\theta_0)^\top = M(h,\theta_0)I(\theta_0)M(h,\theta_0)^\top$$

where $L(h,\theta_0)$ is the unique lower triangular matrix with positive diagonal entries. Then (2.2) holds with $\phi(h,\theta_0)^\top := L(h,\theta_0)^{-1}$ and the last entry of the vector $\phi(h,\theta_0)^{-1}(\widehat{\theta}_h - \theta)$ is given by

$$\big[\phi(h,\theta_0)^{-1}(\widehat{\theta}_h - \theta)\big]_k = \big[L(h,\theta_0)^\top(\widehat{\theta}_h - \theta)\big]_k = L_{kk}(h,\theta_0)(\widehat{\theta}_{h,k} - \theta_k).$$

Let $\widetilde{\ell} \in W_{2,1}$ be a loss function of a scalar variable, $\widetilde{\ell} : \mathbb{R} \mapsto \mathbb{R}_+$, and define $\ell(x) := \widetilde{\ell}(x_k)$, $x \in \mathbb{R}^k$. This loss function belongs to $W_{2,k}$ and Hájek's bound implies

$$\varliminf_{h\to 0} \sup_{\|\theta-\theta_0\|<\delta} \mathsf{E}^h_\theta \widetilde{\ell}\big(L_{kk}(h,\theta_0)(\widehat{\theta}_{h,k} - \theta_k)\big) \geqslant \int_{\mathbb{R}} \widetilde{\ell}(t)\gamma_1(t)dt > 0. \tag{2.3}$$

This inequality identifies the last diagonal entry of $L(h,\theta_0)$ as the best minimax rate in estimation of θ_k. Similar bound for an arbitrary entry can be obtained by permuting the components of θ so that it becomes the last.

A commonly encountered instance of (2.2) is when the matrix $M(h,\theta_0)$ is diagonal. Then $L(h,\theta_0) = M(h,\theta_0)S(\theta_0)$ where $S(\theta_0)$ is the Cholseky factor of $I(\theta_0)$. Since $I(\theta_0)$ is positive definite, all diagonal entries of $S(\theta_0)$ are positive (and constant in h) and, in view of (2.3), the best minimax rate is determined only by $M_{kk}(h,\theta_0)$. This is the case for our model in the large time asymptotic regime with $h := 1/T$ (see Theorem 3.1). In the small noise regime with $h := \varepsilon$, the matrix $M(h,\theta_0)$ is non-diagonal (see Theorem 3.2), which results in a logarithmic discrepancy between the best minimax rates in estimation of each parameter.

3. Main results

3.1. Large time asymptotics

Covariance function of the fBm with parameter variable $\theta = (H,\sigma^2) \in (3/4,1) \times (0,\infty) =: \Theta$ can be written as

$$\mathrm{cov}(B^H_s, B^H_t) = \int_0^s \int_0^t K_\theta(u-v)dudv$$

where

$$K_\theta(\tau) = \sigma^2 H(2H-1)|\tau|^{2H-2}. \tag{3.1}$$

The Fourier transform of this kernel has the explicit formula

$$\widehat{K}_\theta(\lambda) = \int_{\mathbb{R}} K_\theta(\tau)e^{-\mathrm{i}\lambda\tau}d\tau = \sigma^2 a_H|\lambda|^{1-2H}, \quad \lambda \in \mathbb{R}\backslash\{0\}, \tag{3.2}$$

with the constant $a_H := \Gamma(2H+1)\sin(\pi H)$. The function $\widehat{K}_\theta(\lambda)$ does not decay sufficiently fast to be integrable on $\mathbb{R}$ and hence, strictly speaking, it is not a spectral density of a stochastic process in the usual sense. Roughly, it can be thought of as the spectral density of the *fractional noise*, a formal derivative of the fBm.

Denote by P_θ^T the probability measure on the space of continuous functions $C([0,T],\mathbb{R})$ induced by the mixed fBm (1.3) with parameter θ and a fixed noise intensity $\varepsilon > 0$.

Theorem 3.1. *The family $(\mathsf{P}_\theta^T)_{\theta\in\Theta}$ is LAN at any $\theta_0 \in \Theta$ as $T \to \infty$ with*

$$\phi(T) = T^{-1/2} I(\theta_0, \varepsilon)^{-1/2}$$

where $I(\theta_0,\varepsilon)$ is the Fisher information matrix

$$I(\theta,\varepsilon) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \nabla^\top \log\big(\varepsilon + \widehat{K}_\theta(\lambda)\big)\nabla \log\big(\varepsilon + \widehat{K}_\theta(\lambda)\big)d\lambda > 0 \tag{3.3}$$

with ∇ being the gradient with respect to parameter variable θ.

In view of the discussion in the previous section, this result implies that the rate $T^{-1/2}$ is minimax optimal for both H and σ^2. As explained in Sect. 4.3, this rate is achievable and, moreover, Hájek's lower bound can be approached arbitrarily close by estimators based on sufficiently dense grid of discretized observations. The Fisher information matrix in (3.3) remains finite if and only if $H > 3/4$, in agreement with the absolute continuity of measures [7]. It admits of an explicit though somewhat cumbersome expression.

3.2. Small noise asymptotics

With a convenient abuse of notations, let $\mathsf{P}_\theta^\varepsilon$ now denote the probability measure induced by the mixed fBm (1.3) with parameter θ and a fixed interval length $T > 0$. Define the matrix

$$M(\varepsilon,\theta) = \varepsilon^{-1/(4H-2)}\begin{pmatrix} 1 & -2\sigma^2\log\varepsilon^{-1/(2H-1)} \\ 0 & 1\end{pmatrix}.$$

Theorem 3.2. *Assume that $\phi(\varepsilon,\theta_0)$ satisfies the scaling condition*

$$\phi(\varepsilon,\theta_0)^\top M(\varepsilon,\theta_0) T I(\theta_0,1) M(\varepsilon,\theta_0)^\top \phi(\varepsilon,\theta_0) \xrightarrow[\varepsilon\to 0]{} \mathrm{Id}, \tag{3.4}$$

with $I(\theta_0,1)$ defined in (3.3). Then the family $(\mathsf{P}_\theta^\varepsilon)_{\theta\in\Theta}$ is LAN at $\theta_0 \in \Theta$ as $\varepsilon \to 0$.

Condition (3.4) cannot be satisfied by any diagonal matrix $\phi(\varepsilon,\theta_0)$, since in this case the limit, if exists and finite, must be a singular matrix. Otherwise the choice of $\phi(\varepsilon,\theta_0)$ is not unique. As explained in the previous section, the upper and lower triangular Cholesky factors of the matrix $M(\varepsilon,\theta_0)I(\theta_0,1)M(\varepsilon,\theta_0)^\top$ reveal the optimal minimax estimation rates for H and σ^2.

Corollary 3.3.

1) *For any family of estimators* $\widehat{H}_\varepsilon$*, a loss function* $\ell \in W_{2,1}$ *on* $\mathbb{R}$ *and* $\delta > 0$*,*

$$\varliminf_{\varepsilon\to 0} \sup_{\|\theta-\theta_0\|<\delta} \mathsf{E}_\theta \ell\Big(\varepsilon^{-1/(4H_0-2)}(\widehat{H}_\varepsilon - H)\Big) \geqslant \int_{\mathbb{R}} \ell(x/J(\theta_0))\gamma(x)dx,$$

where γ *is the standard normal density on* $\mathbb{R}$ *and*

$$J(\theta) := \sqrt{T\Big(I_{11}(\theta,1) - \frac{I_{12}(\theta,1)^2}{I_{22}(\theta,1)}\Big)}.$$

2) *For any family of estimators* $\widehat{\sigma}^2_\varepsilon$*, a loss function* $\ell \in W_{2,1}$ *on* $\mathbb{R}$ *and* $\delta > 0$*,*

$$\varliminf_{\varepsilon\to 0} \sup_{\|\theta-\theta_0\|<\delta} \mathsf{E}_\theta \ell\Big(\varepsilon^{-1/(4H_0-2)}\frac{1}{\log \varepsilon^{-1}}\big(\widehat{\sigma}^2_\varepsilon - \sigma^2\big)\Big) \geqslant \int_{\mathbb{R}} \ell(x/S(\theta_0))\gamma(x)dx,$$

where γ *is the standard normal density on* $\mathbb{R}$ *and*

$$S(\theta) := \frac{H-\frac{1}{2}}{\sigma^2} J(\theta).$$

If only one parameter is to be estimated, while the other one is known, the relevant LAN property corresponds to the respective one-dimensional family. The following theorem shows that the optimal minimax rates is these cases improve by a logarithmic factor.

Theorem 3.4.

1) *For any fixed* $\sigma_0^2 \in \mathbb{R}_+$*, the family* $\Big(\mathsf{P}^\varepsilon_{(H,\sigma_0^2)}\Big)_{H\in(3/4,1)}$ *is LAN at any* $H_0 \in (3/4,1)$ *as* $\varepsilon \to 0$ *with*

$$\phi(\varepsilon, H_0) := \varepsilon^{1/(4H_0-2)} \frac{1}{\log \varepsilon^{-1}} \frac{H_0 - \frac{1}{2}}{\sigma_0^2} \frac{1}{\sqrt{T I_{22}(\theta_0,1)}}. \tag{3.5}$$

2) *For any fixed* $H_0 \in (3/4,1)$*, the family* $\Big(\mathsf{P}^\varepsilon_{(H_0,\sigma^2)}\Big)_{\sigma^2\in(0,\infty)}$ *is LAN at any* $\sigma_0^2 \in \mathbb{R}_+$ *as* $\varepsilon \to 0$ *with*

$$\phi(\varepsilon, \sigma_0) := \varepsilon^{1/(4H_0-2)} \frac{1}{\sqrt{T I_{22}(\theta_0,1)}}.$$

4. A discussion

4.1. On the information matrix

The expression for the Fisher information matrix in (3.3) is known as Whittle's formula. It was discovered by P. Whittle [42, 43] and was originally derived for discrete time stationary Gaussian processes with continuous spectral densities, see also [41]. Its validity was extended in [11, 12] to sequences with long range dependence, for which the spectral density has an integrable singularity at the origin.

Whittle's formula in continuous time is a more subtle matter due to complexity of the absolute continuity relation between Gaussian measures on function spaces. In fact, according to the survey [15], it was never rigorously verified beyond processes with rational spectra. One important class for which further generalization is plausible are processes observed with additive "white noise", that is,

$$X_t = Z_t + B_t, \quad t \in [0, T],$$

where B is a standard Brownian motion and Z is a centered Gaussian process with stationary increments. The mixed fBm is a special case from this class.

Results in [37] imply that the probability measure induced by X is equivalent to the Wiener measure if and only if

$$\mathsf{E} Z_t Z_s = \int_0^t \int_0^s K_\theta(u - v) du dv$$

for some kernel $K_\theta \in L^2([0, T])$. In this case, the Radon-Nikodym derivative has the same form as in (5.4). Using the theory of finite section approximation from [19] it is indeed possible to prove Theorem 3.1 for such processes under the additional, crucial to the approach of [19], assumption $K \in L^1(\mathbb{R})$.

This condition is violated by the kernel (3.1), which makes the method of [19] inapplicable. This is not entirely surprising in view of the difficulties, needed to be overcome in [11] to extend Whittle's theory to discrete time processes with the long range dependence. The results in our paper are proved using a different approach, based on the ideas from [39] and their recent applications to processes with the fractional covariance structure [8].

4.2. On the joint and separate estimation

Logarithmic discrepancy in the minimax rates between joint and separate estimation as in Corollary 3.3 and Theorem 3.4 is known to occur in the high-frequency regime in experiments with discrete data such as (1.1). The optimal rates for the separate estimation of H and σ^2 for $\Delta = T/n$ are

$$n^{-1/2} \frac{1}{\log n} \quad \text{and} \quad n^{-1/2}$$

respectively, see [26] and references therein. These rates are achievable, e.g., by estimators based on discrete power variations as in [24, 27, 10].

It was long noticed that analogous estimators achieve slower rates, degraded by logarithmic factor:

$$n^{-1/2} \quad \text{and} \quad n^{-1/2} \log n \tag{4.1}$$

when both parameters are unknown. These rates were recently proved minimax optimal in [5] where the LAN property was shown to hold with a non-diagonal matrix M in (2.2).

High frequency estimation from the noisy data (1.2) was considered in [18], where the optimal minimax rates for joint estimation of $H > 1/2$ and σ^2 were found to be

$$n^{-1/(4H+2)} \quad \text{and} \quad n^{-1/(4H+2)} \log n,$$

respectively. These rates are slower than those in (4.1), confirming the intuition that noise should make the estimation problem harder. For further developments in the minimax theory of this and related models see [35, 36]. The same rates are shown to remain optimal for $H < \frac{1}{2}$ in the recent preprint [38].

Another important direction of research is concerned with construction of consistent estimators with explicit asymptotic distribution. Such results can be useful, e.g., for construction of asymptotic confidence intervals. In [29] the authors consider a model more general than (1.2) where σB_t^H is replaced with the process

$$\int_0^t b_s ds + \int_0^t \sigma_s dB_s^H$$

with unknown, possibly random functions $b = (b_t, t \in [0,T])$ and $\sigma = (\sigma_t, t \in [0,T])$. They construct a family of consistent estimators for H and prove their asymptotic normality. Inference in presence of jumps is studied in [30].

Hurst parameter estimation for the mixed fBm (1.3) with $H < 3/4$ is addressed in [13], where estimators, consistent in the high frequency regime, are constructed using the power variations technique. Recently it was shown in [9] that if B^H and B in (1.3) are correlated, H becomes identifiable and can be estimated consistently in the high frequency setup for the whole range $H \in (0,1)$. The case of complete correlation is studied in [14].

4.3. On the rate optimal estimators

It is typical for the LAN models in general that Hájek's asymptotic bound (Theorem 2.2) is attained by the Maximum Likelihood estimator and Bayes estimators with positive prior densities, see [21]. However, this is not automatic and, in our case, the proof would require estimates on the solution to the integral equation (5.3), more delicate than those needed for the LAN analysis presented in this paper. Such estimates currently remain out of reach and thus the question of exact attainability of Hájek's bound remains open.

On the more practical side, likelihood based estimators for the model under consideration are of limited interest, since their realization needs a high precision numerical solution of (5.3) and approximation of the stochastic integrals in (5.4). A less ambitious but still meaningful objective is to construct simpler estimators whose asymptotic risk exceeds the bound only by a constant.

In our case, such a rate optimal estimator in the large time asymptotic regime can be constructed using increments of the observed continuous path on a discrete grid of points with a step $\delta > 0$. These increments form a stationary sequence to which, e.g., Whittle's spectral estimator applies directly. The theory from [11] then tells that it achieves the rate $T^{-1/2}$, optimal by Theorem 3.1. Moreover, its limit risk can be made arbitrarily close to Hájek's bound with the Fisher information matrix (3.3) if δ is chosen small enough.

In the small noise regime, rate optimal estimators can be constructed by means of the method suggested in [18], see some details in Appendix A. These estimators attain the best possible minimax rates derived in Corollary 3.3 and Theorem 3.4. However, they can hardly be expected to attain Hájek's bound exactly.

5. The proof roadmap

In this section we detail the principle steps of the proof, deferring its more technical parts to the next sections. Let $B = (B_t, t \in \mathbb{R}_+)$ and $B^H = (B_t^H, t \in \mathbb{R}_+)$ be independent standard and fractional Brownian motions on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. The mixed fBm (1.3) with $\theta = (H, \sigma^2) \in (3/4, 1) \times \mathbb{R}_+$ satisfies the canonical innovation representation [20]

$$X_t = \int_0^t \rho_t(X, \theta) dt + \sqrt{\varepsilon}\, \overline{B}_t, \quad t \in [0, T], \tag{5.1}$$

where $\overline{B}$ is a Brownian motion with respect to $\mathcal{F}_t^X = \sigma\{X_s, s \leqslant t\}$, and

$$\rho_t(X, \theta) = \int_0^t g(t, t-s; \theta) dX_s, \tag{5.2}$$

and the function $g(t, s; \theta)$ solves the integral equation

$$\varepsilon g(t, s; \theta) + \int_0^t K_\theta(r-s) g(t, r; \theta) dr = K_\theta(s), \quad 0 < s < t, \tag{5.3}$$

with the kernel $K_\theta(\cdot)$ defined in (3.1). This equation has the unique solution in $L^2([0, t])$ since its kernel is Hilbert-Schmidt for $H > 3/4$. The stochastic integral in (5.2) can therefore be defined in the usual way, see [33].

Let P^T and P_θ^T be the probability measures on $C([0, T], \mathbb{R})$ induced by the Brownian motion $\sqrt{\varepsilon}\, \overline{B}$ and the mixed fBm with parameter θ, respectively. By the Girsanov theorem, applied to the innovation representation (5.1), these measures are mutually absolutely continuous $\mathsf{P}^T \sim \mathsf{P}_\theta^T$ with the Radon-Nikodym derivative

$$\frac{d\mathsf{P}_\theta^T}{d\mathsf{P}^T}(X^T) = \exp\left(\frac{1}{\varepsilon} \int_0^T \rho_t(X, \theta) dX_t - \frac{1}{2} \frac{1}{\varepsilon} \int_0^T \rho_t(X, \theta)^2 dt\right). \tag{5.4}$$

5.1. The large time asymptotics

In view of (5.1) and (5.4) the likelihood ratio in Definition 2.1 takes the form

$$\begin{aligned}\log \frac{d\mathsf{P}^T_{\theta_0+\phi(T)u}}{d\mathsf{P}^T_{\theta_0}}(X^T) =& \frac{1}{\sqrt{\varepsilon}}\int_0^T \big(\rho_t(X,\theta_0+\phi(T)u)-\rho_t(X,\theta_0)\big)d\overline{B}_t \\ & -\frac{1}{2}\frac{1}{\varepsilon}\int_0^T \big(\rho_t(X,\theta_0+\phi(T)u)-\rho_t(X,\theta_0)\big)^2 dt,\end{aligned} \tag{5.5}$$

where X is the mixed fBm with parameter θ_0 and, cf. Theorem 3.1,

$$\phi(T) = T^{-1/2} I(\theta_0,\varepsilon)^{-1/2}.$$

The matrix $I(\theta_0,\varepsilon)$, defined in (3.3) is invertible, and establishing the LAN property claimed in Theorem 3.1 amounts to proving that for any $u \in \mathbb{R}^2$

$$\frac{1}{\varepsilon}\int_0^T \big(\rho_t(X,\theta_0+u/\sqrt{T})-\rho_t(X,\theta_0)\big)^2 dt \xrightarrow[T\to\infty]{\mathsf{P}} u^\top I(\theta_0,\varepsilon)u, \tag{5.6}$$

since by the CLT for stochastic integrals [28, Theorem 1.19], (5.6) implies the convergence in distribution

$$\frac{1}{\sqrt{\varepsilon}}\int_0^T \big(\rho_t(X,\theta_0+u/\sqrt{T})-\rho_t(X,\theta_0)\big)d\overline{B}_t \xrightarrow[T\to\infty]{d(\mathsf{P})} u^\top I(\theta_0,\varepsilon)^{1/2} Z$$

where $Z\sim N(0,\mathrm{Id})$.

Let us denote partial derivatives with respect to the entries of parameter vector θ by $\partial_1 := \partial_H$ and $\partial_2 := \partial_{\sigma^2}$. The kernel in (3.1) has partial derivatives of all orders for $\tau\neq 0$ and

$$\partial_i K_\theta(\cdot),\ \partial_i\partial_j K_\theta(\cdot) \in L^2([0,t]),\quad i,j\in\{1,2\}.$$

This implies that the solution to equation (5.3) also has partial derivatives

$$\partial_i g(t,\cdot;\theta),\ \partial_i\partial_j g(t,\cdot;\theta) \in L^2([0,t]),\quad i,j\in\{1,2\},$$

which can be interchanged with the stochastic integral in (5.2). Consequently $\rho_t(X,\theta)$ has partial derivatives and

$$\nabla\rho_t(X,\theta) = \int_0^t \nabla g(t,s;\theta)dX_s,\quad \nabla^2\rho_t(X,\theta) = \int_0^t \nabla^2 g(t,s;\theta)dX_s,$$

where ∇ stands for the gradient and ∇^2 denotes the Hessian with respect to θ. Therefore,

$$\begin{aligned}&\rho_t(X,\theta_0+u/\sqrt{T})-\rho_t(X,\theta_0) = \\ &\frac{1}{\sqrt{T}}\nabla\rho_t(X,\theta_0)u + \frac{1}{T}\int_0^1\int_0^\tau u^\top\nabla^2\rho_t(X,\theta_0+su/\sqrt{T})u\,ds\,d\tau,\end{aligned} \tag{5.7}$$

and (5.6) will be true if we show that

$$\frac{1}{\varepsilon}\frac{1}{T}\int_0^T \nabla^\top \rho_t(X,\theta_0)\nabla\rho_t(X,\theta_0)dt \xrightarrow[T\to\infty]{L^2(\Omega)} I(\theta_0,\varepsilon) \tag{5.8}$$

and, for all sufficiently small $\delta > 0$,

$$\frac{1}{T^2}\int_0^T \sup_{\theta:\|\theta-\theta_0\|\leqslant\delta} \mathsf{E}\big\|\nabla^2\rho_t(X,\theta)\big\|^2 dt \xrightarrow[T\to\infty]{} 0. \tag{5.9}$$

The main challenge in the proof consists of establishing the properties of the gradient process $\nabla\rho_t(X,\theta)$ which guarantee these two limits. By definition (5.2),

$$\begin{aligned}
&\mathsf{E}\,\partial_i\rho_s(X,\theta_0)\partial_j\rho_t(X,\theta_0) = \\
&\int_0^t\int_0^s \partial_i g(s,s-x;\theta_0)\partial_j g(t,t-y;\theta_0)K_{\theta_0}(x-y)dxdy+ \\
&\varepsilon\int_0^s \partial_i g(s,s-x;\theta_0)\partial_j g(t,t-x;\theta_0)dx = \\
&\int_0^t\int_0^s \partial_i g(s,x;\theta_0)\partial_j g(t,y;\theta_0)K_{\theta_0}(y-(x+t-s))dxdy+ \\
&\varepsilon\int_0^s \partial_i g(s,x;\theta_0)\partial_j g(t,x+t-s;\theta_0)dx.
\end{aligned}$$

Extending the domain of $s\mapsto g(t,s;\theta_0)$ outside the interval $(0,t)$ by zero, define the Fourier transform

$$\widehat{g}_t(\mathrm{i}\lambda,\theta_0) = \int_{\mathbb{R}} g(t,s;\theta_0)e^{-\mathrm{i}\lambda s}ds, \quad \lambda\in\mathbb{R},$$

and the function

$$\Lambda(\mathrm{i}\lambda) = \varepsilon + \widehat{K}_{\theta_0}(\lambda) \tag{5.10}$$

where $\widehat{K}_{\theta_0}(\lambda)$ is the Fourier transform (3.2). Then by Plancherel's theorem

$$\begin{aligned}
&\mathsf{E}\,\partial_i\rho_s(X,\theta_0)\partial_j\rho_t(X,\theta_0) = \\
&\qquad \frac{1}{2\pi}\int_{-\infty}^{\infty} \partial_j\widehat{g}_t(\mathrm{i}\lambda,\theta_0)\overline{\partial_i\widehat{g}_s(\mathrm{i}\lambda,\theta_0)}\Lambda(\mathrm{i}\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda.
\end{aligned} \tag{5.11}$$

Using this formula and suitable estimates for the Fourier transform of the solution to (5.3) we will derive the following decomposition.

Lemma 5.1. *The covariance function of the gradient process satisfies*

$$\mathsf{E}\,\nabla^\top\rho_s(X;\theta_0)\nabla\rho_t(X;\theta_0) = Q(t-s)+R(s,t)$$

where the matrices in the right hand side admit the bounds

$$\begin{aligned}
&\big\|Q(t-s)\big\| \leqslant C\wedge|t-s|^{-1}\big|\log|t-s|\big|^3, \quad \forall s,t\in\mathbb{R}_+, \\
&\big\|R(s,t)\big\| \leqslant C\Big(t^{-1/2}+s^{-1/2}+(st)^{-b}\Big), \quad \forall s,t\in[T_{\min},\infty),
\end{aligned} \tag{5.12}$$

with some constants $b \in (0, \frac{1}{2})$, $C > 0$ *and* $T_{\min} > 0$. *Moreover,*

$$Q(0) = \varepsilon I(\theta_0, \varepsilon), \tag{5.13}$$

where $I(\theta, \varepsilon)$ *is defined in* (3.3).

This lemma implies that

$$\mathsf{E}\nabla^\top \rho_t(X, \theta_0)\nabla\rho_t(X, \theta_0) \xrightarrow[t\to\infty]{} Q(0) = \varepsilon I(\theta_0, \varepsilon)$$

and

$$\frac{1}{T^2}\int_0^T\int_0^T \left\|\mathsf{E}\nabla^\top \rho_s(X, \theta_0)\nabla\rho_t(X, \theta_0)\right\|^2 dsdt \xrightarrow[T\to\infty]{} 0.$$

Since $\nabla\rho_t(X, \theta_0)$ is a centered Gaussian process, these two limits and Isserlis' theorem ensure (5.8). In addition, the convergence in (5.9) holds due to the following bound.

Lemma 5.2. *For all sufficiently small* $\delta > 0$, *there exist constants* $C > 0$ *and* $T_{\min} > 0$ *such that*

$$\sup_{\theta:\|\theta-\theta_0\|\leqslant\delta} \mathsf{E}\left\|\nabla^2\rho_t(X, \theta)\right\|^2 \leqslant C, \quad \forall t \geqslant T_{\min}.$$

To recap, the proof of Theorem 3.1 now reduces to verifying Lemmas 5.1–5.2. This is done by means of asymptotic analysis of the integral equation (5.3) as $T \to \infty$, see Sect. 6.1. In essence, it yields quantitative bounds on the deviation of $g(t, s; \theta)$ from the the solution to the Wiener-Hopf equation on the semi-axis:

$$\varepsilon g(s;\theta) + \int_0^\infty K_\theta(r - s)g(r;\theta)dr = K_\theta(s), \quad s \in (0, \infty).$$

These bounds are obtained directly in terms of the Laplace transforms, which turns out to be particularly convenient in view of the formula (5.11).

5.2. The small noise asymptotics

The relevant likelihood ratio in this case is, cf. (5.5),

$$\begin{aligned}\log \frac{d\mathsf{P}^\varepsilon_{\theta_0+\phi(\varepsilon)u}}{d\mathsf{P}^\varepsilon_{\theta_0}}(X^T) =& \frac{1}{\sqrt{\varepsilon}}\int_0^T \big(\rho^\varepsilon_t(X^\varepsilon, \theta_0 + \phi(\varepsilon)u) - \rho^\varepsilon_t(X^\varepsilon, \theta_0)\big)d\overline{B}_t \\ &-\frac{1}{2}\frac{1}{\varepsilon}\int_0^T \big(\rho^\varepsilon_t(X^\varepsilon, \theta_0 + \phi(\varepsilon)u) - \rho^\varepsilon_t(X^\varepsilon, \theta_0)\big)^2 dt,\end{aligned} \tag{5.14}$$

where T is fixed and dependence on ε is emphasized by superscripts. Here, cf. (5.7),

$$\begin{aligned}&\rho^\varepsilon_t(X^\varepsilon, \theta_0 + \phi(\varepsilon)u) - \rho^\varepsilon_t(X^\varepsilon, \theta_0) = \\ &\nabla\rho^\varepsilon_t(X^\varepsilon, \theta_0)\phi(\varepsilon)u + \int_0^1\int_0^\tau u^\top\phi(\varepsilon)^\top\nabla^2\rho^\varepsilon_t(X^\varepsilon, \theta_0 + s\phi(\varepsilon)u)\phi(\varepsilon)udsd\tau.\end{aligned}$$

We will argue that for an appropriate choice of $\phi(\varepsilon) := \phi(\varepsilon, \theta_0)$

$$\frac{1}{\varepsilon} u^\top \phi(\varepsilon)^\top \left(\int_0^T \nabla^\top \rho_t^\varepsilon(X^\varepsilon, \theta_0) \nabla \rho_t^\varepsilon(X^\varepsilon, \theta_0) dt \right) \phi(\varepsilon) u \xrightarrow[\varepsilon \to 0]{\mathsf{P}_{\theta_0}} \|u\|^2 \tag{5.15}$$

and

$$\mathsf{E} \frac{1}{\varepsilon} \int_0^T \left(\int_0^1 \int_0^\tau u^\top \phi(\varepsilon)^\top \nabla^2 \rho_t^\varepsilon \big(X^\varepsilon, \theta_0 + s\phi(\varepsilon)u\big) \phi(\varepsilon) u ds d\tau \right)^2 dt \xrightarrow[\varepsilon \to 0]{} 0. \tag{5.16}$$

Then the second term in (5.14) converges to $-\frac{1}{2}\|u\|^2$ in probability and the stochastic integral converges in distribution to $u^\top Z$ with $Z \sim N(0, \mathrm{Id})$, see [25, Ch. IX.5].

Equation (5.3) degenerates as $\varepsilon \to 0$ to the integral equation of the first kind

$$\int_0^t K_\theta(r - s) g(t, r; \theta) dr = K_\theta(s), \quad 0 < s < t,$$

which does not have a classic solution. This makes the direct proof of (5.15)–(5.16) complicated. The main tool in proving these limits is a certain scaling property of the solution to (5.3) (see Lemma 7.2), which relates the small noise to the large time asymptotics from the previous subsection. This scaling stems from the structure of kernel (3.1), corresponding to self-similarity of the fBm.

6. Proof of Theorem 3.1

As argued in Sect. 5.1, the assertion of Theorem 3.1 follows once we prove Lemmas 5.1 and 5.2. This is done in Sects. 6.2 and 6.3, respectively. The proofs are based on representation of the solution to equation (5.3) derived in Sect. 6.1.

6.1. Equation (5.3)

In this subsection we show that solution to (5.3) can be decomposed into a sum of the main term independent of t and the residual term which vanishes as $t \to \infty$, see Lemma 6.4 below. Our approach is inspired by the method, pioneered in [39] in the context of spectral analysis of the integral operator with weakly singular kernel (3.1). Recently it was generalized to covariance operators of related stochastic processes [8, 32]. Here we will adapt this method to a different problem, namely solving an integral equation of the second kind. For brevity, θ will be omitted from the notations in this section.

Let us first sketch the main ideas. Consider the Laplace transform of the solution to (5.3)

$$\widehat{g}_t(z) = \int_{-\infty}^{\infty} g(t, s) e^{-zs} ds = \int_0^t g(t, s) e^{-zs} ds, \quad z \in \mathbb{C}, \tag{6.1}$$

where the domain of $g(t,\cdot)$ is extended to the real whole axis by zero.

Using the specific structure of kernel (3.1) it is possible (Lemma 6.1) to derive the representation

$$\widehat{g}_t(z) - 1 = \frac{\Phi_0(z) + e^{-tz}\Phi_1(-z)}{\Lambda(z)}. \tag{6.2}$$

Here $\Phi_0(z)$ and $\Phi_1(z)$ are functions, holomorphic on the cut plane $\mathbb{C}\backslash\mathbb{R}_+$ with a discontinuity across the cut $\mathbb{R}_+$ with

$$\Phi_i^{\pm}(t) = \lim_{z\to t^{\pm}} \Phi_i(z), \quad t \in \mathbb{R}_+,$$

where $+$ and $-$ correspond to the limits taken in the upper and lower half-planes, respectively. Such functions are called sectionally holomorphic, [17]. The function $\Lambda(z)$ is defined by an explicit formula, see (6.4) below, it is non-vanishing and sectionally holomorphic on $\mathbb{C}\backslash\mathbb{R}$.

Since integration in (6.1) is carried out over a bounded interval, $\widehat{g}_t(z)$ is an entire function. This implies that the discontinuity in the right hand side of (6.2) must be removable, i.e.,

$$\lim_{z\to\tau^+} \frac{\Phi_0(z) + e^{-tz}\Phi_1(-z)}{\Lambda(z)} = \lim_{z\to\tau^-} \frac{\Phi_0(z) + e^{-tz}\Phi_1(-z)}{\Lambda(z)}, \quad \forall\tau \in \mathbb{R}. \tag{6.3}$$

A calculation shows that this is equivalent to a boundary condition, see (6.13), which must be satisfied by $\Phi_0(z)$ and $\Phi_1(z)$ on the cut $\mathbb{R}_+$. It turns out that this condition along with certain a priori growth estimates (see Lemma 6.1) determine these functions uniquely and they can be expressed in terms of solutions to a system of auxiliary integral equations on $\mathbb{R}_+$, see (6.21). Plugging back these expressions into (6.2) yields the desired decomposition for the Laplace transform $\widehat{g}_t(z)$, see Lemma 6.4. Moreover, using the auxiliary equations (6.21) it is possible to derive useful bounds for the residual term in this decomposition, see Lemma 6.6.

The rest of this section details the implementation of this program.

6.1.1. The Laplace transform

The following lemma derives representation (6.2) for the Laplace transform of solution to (5.3).

Lemma 6.1. *The Laplace transform* (6.1) *satisfies* (6.2) *where*

$$\Lambda(z) = \varepsilon + \frac{\sigma^2}{2}\Gamma(2H+1)\big(z^{1-2H} + (-z)^{1-2H}\big) \tag{6.4}$$

and the functions $\Phi_0(z)$ *and* $\Phi_1(z)$ *are sectionally holomorphic on* $\mathbb{C}\backslash\mathbb{R}_+$ *and satisfy*

$$\Phi_0(z) = -\varepsilon + O(z^{1-2H}) \quad \text{and} \quad \Phi_1(z) = O(z^{1-2H}), \quad z \to \infty, \tag{6.5}$$

and

$$\Phi_0(z) = O(z^{1-2H}) \quad \text{and} \quad \Phi_1(z) = O(z^{1-2H}), \quad z \to 0. \tag{6.6}$$

Proof. By the definition of Euler's gamma function, the kernel in (3.1) satisfies the integral formula

$$K_\theta(u) = \int_0^\infty \kappa(\tau)e^{-|u|\tau}d\tau, \quad u \in \mathbb{R}, \tag{6.7}$$

where

$$\kappa(\tau) = \sigma^2\frac{H(2H-1)}{\Gamma(2-2H)}\tau^{1-2H}, \quad \tau \in \mathbb{R}_+.$$

Replacing the kernel in equation (5.3) with expression (6.7) gives

$$\varepsilon g(t,s) + \int_0^t g(t,r)\int_0^\infty \kappa(\tau)e^{-|s-r|\tau}d\tau dr = \int_0^\infty \kappa(\tau)e^{-s\tau}d\tau. \tag{6.8}$$

The Laplace transform of the integral in the left hand side is

$$\begin{aligned}
&\int_0^t \left(\int_0^t g(t,r)\int_0^\infty \kappa(\tau)e^{-|s-r|\tau}d\tau dr\right)e^{-sz}ds = \\
&\int_0^t g(t,r)\int_0^\infty \kappa(\tau)\left(\int_0^t e^{-|s-r|\tau}e^{-sz}ds\right)d\tau dr = \\
&\int_0^t g(t,r)\int_0^\infty \kappa(\tau)\left(\frac{e^{-rz}-e^{-r\tau}}{\tau-z}+\frac{e^{-rz}-e^{-tz-(t-r)\tau}}{\tau+z}\right)d\tau dr = \\
&\widehat{g}_t(z)\big(\mu(z)+\mu(-z)\big) - \int_0^\infty \frac{\kappa(\tau)}{\tau-z}\widehat{g}_t(\tau)d\tau - e^{-tz}\int_0^\infty \frac{\kappa(\tau)}{\tau+z}\breve{g}_t(\tau)d\tau,
\end{aligned}$$

where $\breve{g}_t(z) := \int_0^t g(t,t-r)e^{-zr}dr$ is the Laplace transform of time reversed solution and

$$\mu(z) = \int_0^\infty \frac{\kappa(x)}{x-z}dx = \frac{\sigma^2}{2}\Gamma(2H+1)(-z)^{1-2H}.$$

Similarly,

$$\int_0^t \left(\int_0^\infty \kappa(\tau)e^{-s\tau}d\tau\right)e^{-sz}ds = \mu(-z) - e^{-tz}\int_0^\infty \kappa(\tau)\frac{e^{-t\tau}}{\tau+z}d\tau.$$

Thus applying the Laplace transform to (6.8) we obtain (6.2) with

$$\begin{aligned}
\Phi_0(z) &:= -\varepsilon - \mu(z) + \int_0^\infty \frac{\kappa(\tau)}{\tau-z}\widehat{g}_t(\tau)d\tau, \\
\Phi_1(z) &:= -\int_0^\infty \frac{\kappa(\tau)}{\tau-z}e^{-t\tau}d\tau + \int_0^\infty \frac{\kappa(\tau)}{\tau-z}\breve{g}_t(\tau)d\tau,
\end{aligned}$$

and $\Lambda(z) = \varepsilon+\mu(z)+\mu(-z)$. The functions $\widehat{g}_t(\tau)$ and $\breve{g}_t(\tau)$ are bounded over $\tau \in \mathbb{R}_+$ and the estimates (6.5)–(6.6) are derived from these formulas by standard calculations. ∎

The next lemma gathers some useful properties of $\Lambda(z)$.

Lemma 6.2. *The function $\Lambda(z)$ defined in (6.4) is non-vanishing and sectionally holomorphic on $\mathbb{C}\backslash\mathbb{R}$ with the limits*

$$\Lambda^{\pm}(\tau) = \varepsilon + \sigma^2 a_H |\tau|^{1-2H} \begin{cases} e^{\pm(H-\frac{1}{2})\pi \mathrm{i}}, & \tau \in \mathbb{R}_+, \\ e^{\mp(H-\frac{1}{2})\pi \mathrm{i}}, & \tau \in \mathbb{R}_-, \end{cases}$$

where $a_H = \Gamma(2H+1)\sin(\pi H)$. These functions satisfy the symmetries

$$\Lambda^+(\tau) = \overline{\Lambda^-(\tau)}, \tag{6.9}$$

$$\frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} = \frac{\Lambda^-(-\tau)}{\Lambda^+(-\tau)}, \tag{6.10}$$

and the principal branch of the argument $\alpha(\tau) := \arg\{\Lambda^+(\tau)\}$,

$$\alpha(\tau) = \arctan \frac{\sigma^2 a_H \sin\big((H-\frac{1}{2})\pi\big)}{\varepsilon|\tau|^{2H-1} + \sigma^2 a_H \cos\big((H-\frac{1}{2})\pi\big)} \operatorname{sign}(\tau), \tag{6.11}$$

is an odd decreasing function, continuous on $\mathbb{R}\backslash\{0\}$, satisfying

$$\alpha(0+) = \pi(H - \tfrac{1}{2}) \quad \text{and} \quad \alpha(\tau) = O(\tau^{1-2H}) \quad \text{as } \tau \to \infty. \tag{6.12}$$

Proof. All the claims are derived by direct calculations using (6.4). □

6.1.2. An equivalent representation

In this section we will use (6.2) to show that Laplace transform (6.1) can be expressed in terms of solutions to certain auxiliary equations (6.21), see Lemma 6.4. The key observation to this end is that $\widehat{g}_t(z)$ is an entire function and hence discontinuity in the right hand side of (6.2) must be removable, that is, (6.3) must hold. Due to the symmetries in (6.10), this condition reduces to

$$\begin{aligned} \Phi_0^+(\tau) - \frac{\Lambda^+(\tau)}{\Lambda^-(\tau)}\Phi_0^-(\tau) &= e^{-t\tau}\Phi_1(-\tau)\Big(\frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} - 1\Big), \\ \Phi_1^+(\tau) - \frac{\Lambda^+(\tau)}{\Lambda^-(\tau)}\Phi_1^-(\tau) &= e^{-t\tau}\Phi_0(-\tau)\Big(\frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} - 1\Big), \end{aligned} \qquad \forall \tau \in \mathbb{R}_+, \tag{6.13}$$

where, in view of (6.9),

$$\frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} = \exp(2\mathrm{i}\alpha(\tau)).$$

The functions $\Phi_0(z)$ and $\Phi_1(z)$ are sectionally holomorphic on $\mathbb{C}\backslash\mathbb{R}_+$, satisfy the boundary conditions (6.13) and the growth estimates (6.5). Using the usual technique of solving the Hilbert boundary value problems, such functions can be expressed in terms of solutions to certain auxiliary integral equations as follows.

The first step consists of finding a function $X(z)$, sectionally holomorphic on $\mathbb{C}\backslash\mathbb{R}_+$ and satisfying the *homogeneous* boundary condition, cf. (6.13),

$$X^+(\tau) - \frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} X^-(\tau) = 0, \quad \forall \tau \in \mathbb{R}_+.$$

This is a standard instance of the homogeneous Hilbert boundary value problem [17]. Since the function

$$\log \frac{\Lambda^+(\tau)}{\Lambda^-(\tau)} = 2\mathrm{i}\alpha(\tau)$$

satisfies the Hölder condition on $\mathbb{R}_+ \cup \{\infty\}$, all solutions to this problem, which do not vanish on $\mathbb{C}\backslash\{0\}$, have the form $X(z) = z^k X_c(z)$ for some integer $k \in \mathbb{Z}$, where the canonical part is found by the Sokhotski–Plemelj formula

$$\begin{aligned} X_c(z) = \exp\Big(\frac{1}{2\pi\mathrm{i}}\int_0^\infty \frac{\log \Lambda^+(\tau)/\Lambda^-(\tau)}{\tau - z} d\tau\Big) = \\ \exp\Big(\frac{1}{\pi}\int_0^\infty \frac{\alpha(\tau)}{\tau - z} d\tau\Big), \quad z \in \mathbb{C}\backslash\mathbb{R}_+. \end{aligned} \tag{6.14}$$

The following lemma summarizes some of its useful properties.

Lemma 6.3. *The function defined in* (6.14) *satisfies the asymptotics*

$$X_c(z) = \begin{cases} O(z^{\frac{1}{2}-H}), & z \to 0, \\ 1, & z \to \infty, \end{cases} \tag{6.15}$$

and is related to $\Lambda(z)$*, defined in* (6.4)*, by the identity*

$$X_c(z)X_c(-z) = \frac{1}{\varepsilon}\Lambda(z), \quad z \in \mathbb{C}\backslash\mathbb{R}. \tag{6.16}$$

Proof. Asymptotics (6.15) readily follows from (6.12). To prove (6.16), we can write

$$\begin{aligned} &\log X_c(z)X_c(-z) = \\ &\frac{1}{2\pi\mathrm{i}}\int_0^\infty \frac{1}{\tau - z}\log\frac{\varepsilon^{-1}\Lambda^+(\tau)}{\varepsilon^{-1}\Lambda^-(\tau)} d\tau + \frac{1}{2\pi\mathrm{i}}\int_0^\infty \frac{1}{\tau + z}\log\frac{\varepsilon^{-1}\Lambda^+(\tau)}{\varepsilon^{-1}\Lambda^-(\tau)} d\tau. \end{aligned}$$

By changing the integration variable and using the symmetry (6.10), the second integral can be written as

$$\frac{1}{2\pi\mathrm{i}}\int_0^\infty \frac{1}{\tau + z}\log\frac{\varepsilon^{-1}\Lambda^+(\tau)}{\varepsilon^{-1}\Lambda^-(\tau)} d\tau = -\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^0 \frac{1}{\tau - z}\log\frac{\varepsilon^{-1}\Lambda^-(\tau)}{\varepsilon^{-1}\Lambda^+(\tau)} d\tau.$$

Since $\log\big(\Lambda^\pm(\tau)/\varepsilon\big) = O(\tau^{1-2H})$, this implies

$$\log X_c(z)X_c(-z) = \frac{1}{2\pi\mathrm{i}}\int_{-\infty}^\infty \frac{\log\big(\Lambda^+(\tau)/\varepsilon\big)}{\tau - z} d\tau - \frac{1}{2\pi\mathrm{i}}\int_{-\infty}^\infty \frac{\log\big(\Lambda^-(\tau)/\varepsilon\big)}{\tau - z} d\tau.$$

The function $\Lambda(z)$ is non-vanishing and holomorphic on the lower and upper half-planes, hence each of the integrals can be computed by the standard contour integration. When $\mathrm{Im}(z) > 0$ the first integral gives $\log(\Lambda(z)/\varepsilon)$ and the second vanishes, which proves validity of (6.16) in the upper half-plane. The same argument applies to the lower half-plane. □

Now let us define

$$\begin{aligned} S(z) &:= \frac{\Phi_0(z) + \Phi_1(z)}{2X(z)}, \\ D(z) &:= \frac{\Phi_0(z) - \Phi_1(z)}{2X(z)}. \end{aligned} \tag{6.17}$$

These functions are also sectionally holomorphic on $\mathbb{C}\backslash\mathbb{R}_+$ and, in view of (6.13), satisfy the *decoupled* boundary conditions

$$\begin{aligned} S^+(\tau) - S^-(\tau) &= \quad 2\mathrm{i}h(\tau)e^{-t\tau}S(-\tau), \\ D^+(\tau) - D^-(\tau) &= -\,2\mathrm{i}h(\tau)e^{-t\tau}D(-\tau), \end{aligned} \qquad \forall \tau \in \mathbb{R}_+, \tag{6.18}$$

where we defined

$$h(\tau) := \frac{1}{2\mathrm{i}}\Big(\frac{X^+(\tau)}{X^-(\tau)} - 1\Big)\frac{X(-\tau)}{X^+(\tau)}.$$

This function is, in fact, real valued:

$$\begin{aligned} h(\tau) =& \frac{1}{2\mathrm{i}}\Big(e^{2\mathrm{i}\alpha(\tau)} - 1\Big)\exp\Big(-\frac{2\tau}{\pi}⨍_0^\infty \frac{\alpha(s)}{s^2-\tau^2}ds - \mathrm{i}\alpha(\tau)\Big) = \\ & \exp\Big(-\frac{1}{\pi}\int_0^\infty \alpha'(s)\log\Big|\frac{\tau+s}{\tau-s}\Big|ds\Big)\sin\alpha(\tau), \end{aligned} \tag{6.19}$$

where the dashed integral is the Cauchy principle value.

In view of estimates (6.6) and (6.15), the functions $S(-\tau)$ and $D(-\tau)$ will have at most square integrable singularities at the origin if we choose $k \leqslant 0$. From here on we will fix $k = 0$ so that $X(z) = X_c(z)$. This choice is not the only possible, but it makes further calculations simpler. Thus the expressions in the right hand side of (6.18) satisfy the Hölder condition on $\mathbb{R}_+ \cup \{\infty\}$ and therefore, by the Sokhotski-Plemelj theorem, the functions (6.17) satisfy

$$\begin{aligned} S(z) &= \quad \frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau - z}S(-\tau)d\tau - \frac{\varepsilon}{2}, \\ D(z) &= -\frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau - z}D(-\tau)d\tau - \frac{\varepsilon}{2}, \end{aligned} \qquad z \in \mathbb{C}\backslash\mathbb{R}_+. \tag{6.20}$$

Constants in the right hand side match the growth of $S(z)$ and $D(z)$ as $z \to \infty$ in view of estimates (6.5) and (6.15).

Consider now a pair of auxiliary integral equations

$$\begin{aligned} p_t(s) &= \quad \frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau + s}p_t(\tau)d\tau - \frac{1}{2}, \\ q_t(s) &= -\frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau + s}q_t(\tau)d\tau - \frac{1}{2}, \end{aligned} \qquad s \in \mathbb{R}_+. \tag{6.21}$$

In the next subsection we will argue that, for all sufficiently large t, they have unique solutions such that $q_t(\cdot)+\frac{1}{2}$ and $p_t(\cdot)+\frac{1}{2}$ belong to $L^2(\mathbb{R}_+)$. Setting $z:=-\tau$ for $\tau\in\mathbb{R}_+$ in (6.20) shows that $S(-\tau)$ and $D(-\tau)$ solve (6.21) multiplied by ε. Since by construction $S(-\tau)$ and $D(-\tau)$ are square integrable near the origin, due to uniqueness of the solutions to (6.21), they must coincide with $\varepsilon p_t(\tau)$ and $\varepsilon q_t(\tau)$ and, consequently,

$$S(z)=\varepsilon p_t(-z)\quad\text{and}\quad D(z)=\varepsilon q_t(-z),$$

where $q_t(z)$ and $p_t(z)$ are the unique sectionally holomorphic extensions to $\mathbb{C}\setminus\mathbb{R}_-$. Plugging these expressions along with (6.16) and (6.17) into (6.2) we obtain the following result.

Lemma 6.4. *The Laplace transform* (6.1) *satisfies*

$$\widehat{g}_t(z)-1=-\frac{1}{X(-z)}+\widehat{R}_t(z),\qquad z\in\mathbb{C},\tag{6.22}$$

where

$$\widehat{R}_t(z):=\frac{1}{X(-z)}\big(p_t(-z)+q_t(-z)+1\big)+e^{-tz}\frac{1}{X(z)}\big(p_t(z)-q_t(z)\big).\tag{6.23}$$

6.1.3. The auxiliary equations (6.21)

Consider the integral operator in (6.21)

$$(A_tf)(s):=\frac{1}{\pi}\int_0^\infty\frac{h(\tau)e^{-t\tau}}{\tau+s}f(\tau)d\tau.\tag{6.24}$$

The following lemma asserts that it is a contraction on $L^2(\mathbb{R}_+)$ for all sufficiently large t.

Lemma 6.5. *For any closed ball $B\subset\Theta$, there exist $T_{\min}>0$ and $\beta\in(0,1)$ such that*

$$\|A_tf\|\leqslant(1-\beta)\|f\|,\quad\forall f\in L^2(\mathbb{R}_+),\quad\forall t\geqslant T_{\min},\ \theta\in B.$$

Proof. The function $h(\tau)$ defined in (6.19) is continuous, nonnegative, vanishes as $\tau\to\infty$ and satisfies, cf. (6.12),

$$h(0+)=\sin\alpha(0+)=\sin\big(\pi(H-\tfrac{1}{2})\big)\in(0,1).$$

Then $c:=\sup_{\theta\in B}h(0+)\in(0,1)$ and there exists $r>0$ such that $h(\tau)\leqslant\frac{1}{2}c+\frac{1}{2}=:1-\beta\in(0,1)$ for all $\tau\in[0,r]$. Then for any $\tau\geqslant0$,

$$h(\tau)e^{-\tau t}\leqslant(1-\beta)\mathbf{1}_{\{\tau\leqslant r\}}+\|h\|_\infty e^{-rt}\mathbf{1}_{\{\tau>r\}}\leqslant1-\beta,$$

where the last inequality holds for all $t\geqslant\frac{1}{r}\log\frac{\|h\|_\infty}{1-\beta}=:T_{\min}$.

Thus for all $t \geqslant T_{\min}$ and any $f, g \in L^2(\mathbb{R}_+)$, by the Cauchy–Schwarz inequality,

$$\big|\langle g, A_t f\rangle\big| \leqslant \frac{1-\beta}{\pi}\int_0^\infty |g(s)| \int_0^\infty \frac{1}{\tau+s}|f(\tau)|d\tau ds \leqslant$$
$$\frac{1-\beta}{\pi}\left(\int_0^\infty f(\tau)^2\int_0^\infty \frac{\sqrt{\tau/s}}{\tau+s}dsd\tau\right)^{1/2}\left(\int_0^\infty g(s)^2\int_0^\infty \frac{\sqrt{s/\tau}}{\tau+s}d\tau ds\right)^{1/2} =$$
$$(1-\beta)\|g\|\|f\|.$$

Hence $\|A_t f\|^2 = \langle A_t f, A_t f\rangle \leqslant (1-\beta)\|A_t f\|\|f\|$, which proves the clam. □

The equations in (6.21) can be written as

$$f + \tfrac{1}{2} = \pm A_t(f+\tfrac{1}{2}) \mp \tfrac{1}{2}(A_t 1). \tag{6.25}$$

A direct calculation shows that $A_t 1 \in L^2(\mathbb{R}_+)$. Hence these equations have unique solutions in $L^2(\mathbb{R}_+)$ given, e.g., by the Neumann series. The estimates for these solutions, derived in the next lemma, play the key role in the asymptotic analysis.

Lemma 6.6. *For any closed ball $B \subset \Theta$, there exist constants $r_{\max} > 0$, $T_{\min} > 0$ and $C > 0$ such that for any $r \in [0, r_{\max}]$ and all $t \geqslant T_{\min}$*

$$\int_{-\infty}^{\infty} \big|m_t(i\lambda)\big|^2 |\lambda|^{-r} d\lambda \leqslant C t^{r-1},$$

where $m_t(z)$ is any of the functions in

$$\Big\{p_t(z)+\tfrac{1}{2},\ q_t(z)+\tfrac{1}{2},\ \partial_j p_t(z),\ \partial_j q_t(z),\ \partial_i\partial_j p_t(z),\ \partial_i\partial_j q_t(z)\Big\}. \tag{6.26}$$

Proof. Let us start with proving the bound for the first two functions in (6.26). Calculations are similar for both equations in (6.21) and we will consider the first one for definiteness. Rearranging it as in (6.25) and multiplying by s^{-r} shows that the function $\phi(s) := \big(p_t(s)+\frac{1}{2}\big)s^{-r}$ solves the equation

$$\phi = B_t\phi + \psi, \tag{6.27}$$

where $\psi(s) := -\frac{1}{2}(A_t 1)(s)s^{-r}$ with A_t as in (6.24) and

$$(B_t f)(s) := \frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+s}(\tau/s)^r f(\tau)d\tau.$$

By applying the generalized Minkowski inequality we get

$$\begin{aligned}\|\psi\| &= \left(\int_0^\infty\left(\frac{1}{2}\frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+s}s^{-r}d\tau\right)^2 ds\right)^{1/2} \leqslant \\ &\int_0^\infty\left(\int_0^\infty\left(\frac{h(\tau)e^{-t\tau}}{\tau+s}s^{-r}\right)^2 ds\right)^{1/2} d\tau = \\ &\int_0^\infty h(\tau)e^{-t\tau}\tau^{-\frac{1}{2}-r}\left(\int_0^\infty \frac{u^{-2r}}{(u+1)^2}du\right)^{1/2} d\tau \leqslant Ct^{r-\frac{1}{2}}\end{aligned} \tag{6.28}$$

where $r < 1/2$ is assumed. Calculations as in the proof of Lemma 6.5 show that B_t is a contraction on $L^2(\mathbb{R}_+)$ for all $t \geqslant T_{\min}$. Indeed, for any $f, g \in L^2(\mathbb{R}_+)$ and $r < 1/4$,

$$
\begin{aligned}
&\big|\langle g, B_t f\rangle\big| \leqslant \int_0^\infty |g(s)| \frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+s}(\tau/s)^r |f(\tau)| d\tau ds \leqslant \\
&\frac{1-\beta}{\pi}\int_0^\infty\int_0^\infty |g(s)|\frac{(s/\tau)^{\frac14}}{\sqrt{\tau+s}}|f(\tau)|\frac{(\tau/s)^{r+\frac14}}{\sqrt{\tau+s}} d\tau ds \leqslant \\
&\frac{1-\beta}{\pi}\left(\int_0^\infty g(s)^2 \int_0^\infty \frac{(s/\tau)^{\frac12}}{\tau+s} d\tau ds\right)^{1/2}\left(\int_0^\infty f(\tau)^2\int_0^\infty \frac{(\tau/s)^{2r+\frac12}}{\tau+s} ds d\tau\right)^{1/2} = \\
&\frac{1-\beta}{\sqrt{\cos(2\pi r)}}\|g\|\|f\|,
\end{aligned}
$$

where β is given in Lemma 6.5. Hence $\|B_t f\| \leqslant (1-\widetilde{\beta})\|f\|$ with some $\widetilde{\beta} > 0$ if r is small enough. This implies $\|\phi\| \leqslant \widetilde{\beta}^{-1}\|\psi\|$, that is,

$$
\left(\int_0^\infty (p_t(s)+\tfrac12)^2 s^{-2r} ds\right)^{1/2} \leqslant C t^{r-\frac12}. \tag{6.29}
$$

We can now prove the bound for the first function in (6.26),

$$
\begin{aligned}
&\int_{-\infty}^\infty \big|p_t(\mathrm{i}\lambda)+\tfrac12\big|^2 |\lambda|^{-r} d\lambda = \int_{-\infty}^\infty \left|\frac{1}{\pi}\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+\mathrm{i}\lambda} p_t(\tau) d\tau\right|^2 |\lambda|^{-r} d\lambda \leqslant \\
&\int_{-\infty}^\infty \left|\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+\mathrm{i}\lambda}(p_t(\tau)+\tfrac12) d\tau\right|^2 |\lambda|^{-r} d\lambda + \int_{-\infty}^\infty \left|\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+\mathrm{i}\lambda} d\tau\right|^2 |\lambda|^{-r} d\lambda.
\end{aligned}
$$

Due to the generalized Minkowski inequality, the last integral satisfies

$$
\begin{aligned}
\int_{-\infty}^\infty \left|\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+\mathrm{i}\lambda} d\tau\right|^2 |\lambda|^{-r} d\lambda &\leqslant \left(\int_0^\infty h(\tau)e^{-t\tau}\left(\int_{-\infty}^\infty \frac{|\lambda|^{-r}}{\tau^2+\lambda^2} d\lambda\right)^{1/2} d\tau\right)^2 \\
&= C\left(\int_0^\infty e^{-t\tau}\tau^{-r/2-1/2} d\tau\right)^2 \leqslant C t^{r-1}.
\end{aligned}
$$

The other integral can be bounded similarly,

$$
\begin{aligned}
&\int_{-\infty}^\infty \left|\int_0^\infty \frac{h(\tau)e^{-t\tau}}{\tau+\mathrm{i}\lambda}(p_t(\tau)+\tfrac12) d\tau\right|^2 |\lambda|^{-r} d\lambda \leqslant \\
&\left(\int_0^\infty h(\tau)e^{-t\tau}|p_t(\tau)+\tfrac12|\left(\int_{-\infty}^\infty \frac{|\lambda|^{-r}}{\tau^2+\lambda^2} d\lambda\right)^{1/2} d\tau\right)^2 \leqslant \\
&C\left(\int_0^\infty e^{-t\tau}|p_t(\tau)+\tfrac12|\tau^{-r/2-1/2} d\tau\right)^2 \leqslant
\end{aligned}
$$

$$C\int_0^\infty \big(p_t(\tau)+\tfrac12\big)^2\tau^{-2r}d\tau\int_0^\infty e^{-2t\tau}\tau^{r-1}d\tau \leqslant Ct^{r-1},$$

where we used (6.29) and applied the generalized Minkowski and the Cauchy-Schwarz inequalities. This completes the proof for the first two functions in (6.26).

The other two bounds are verified similarly. Note that $\phi(s) := \partial_j p_t(s)s^{-r}$ also solves the equation (6.27) with

$$\psi(s) := s^{-r}\frac{1}{\pi}\int_0^\infty \frac{\partial_j h(\tau)e^{-t\tau}}{\tau+s}p_t(\tau)d\tau.$$

In view of (6.11)

$$\partial_j\alpha(\tau) = \begin{cases} O(1), & \tau\to 0,\\ O(\tau^{1-2H}\log\tau), & \tau\to\infty,\end{cases} \tag{6.30}$$

and, consequently, due to (6.19),

$$\partial_j \log h(\tau) = \begin{cases} O(1), & \tau\to 0,\\ O(\log\tau), & \tau\to\infty.\end{cases}$$

Calculations as in (6.28) then show that $\|\psi\|\leqslant Ct^{r-1/2}$ and the claimed bound for the next two functions in (6.26) are proved as above. The last two bounds for the second order derivatives are verified along the same lines. □

6.2. Proof of Lemma 5.1

In this subsection we will omit θ_0 from the notations for brevity. Covariance function of the gradient process satisfies (5.11), where $\Lambda(\mathrm{i}\lambda)$ introduced in (5.10) is exactly the restriction of $\Lambda(z)$ defined in (6.4) to the imaginary axis. Due to Lemma 6.4,

$$\mathsf{E}\,\partial_i\rho_s(X)\partial_j\rho_t(X) = Q_{ij}(t-s) + R^{(1)}_{ij}(s,t) + R^{(2)}_{ij}(s,t) + R^{(3)}_{ij}(s,t), \tag{6.31}$$

where we defined

$$Q_{ij}(t-s) := \frac{1}{2\pi}\int_{-\infty}^\infty \partial_j\frac{1}{X(-\mathrm{i}\lambda)}\partial_i\frac{1}{X(\mathrm{i}\lambda)}\Lambda(\mathrm{i}\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda \tag{6.32}$$

and

$$R^{(1)}_{ij}(s,t) := -\frac{1}{2\pi}\int_{-\infty}^\infty \partial_j\frac{1}{X(-\mathrm{i}\lambda)}\overline{\partial_i\widehat{R}_s(\mathrm{i}\lambda)}\Lambda(\mathrm{i}\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda, \tag{6.33}$$

$$R^{(2)}_{ij}(s,t) := -\frac{1}{2\pi}\int_{-\infty}^\infty \partial_j\widehat{R}_t(\mathrm{i}\lambda)\overline{\partial_i\frac{1}{X(-\mathrm{i}\lambda)}}\Lambda(\mathrm{i}\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda, \tag{6.34}$$

$$R^{(3)}_{ij}(s,t) := \frac{1}{2\pi}\int_{-\infty}^\infty \partial_j\widehat{R}_t(\mathrm{i}\lambda)\overline{\partial_i\widehat{R}_s(\mathrm{i}\lambda)}\Lambda(\mathrm{i}\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda. \tag{6.35}$$

The first bound in (5.12) is derived in the following lemma.

Lemma 6.7. *There exists $C > 0$ such that*

$$\big|Q_{ij}(t-s)\big| \leqslant C \wedge |t-s|^{-1}\big|\log|t-s|\big|^3, \quad \forall s,t \in \mathbb{R}_+.$$

Proof. Let us estimate the growth of the integrand in (6.32),

$$f(\lambda) := \partial_j \frac{1}{X(-\mathrm{i}\lambda)} \partial_i \frac{1}{X(\mathrm{i}\lambda)} \Lambda(\mathrm{i}\lambda),$$

at the origin and at infinity. In view (6.30) and (6.14),

$$\partial_i \log X(\mathrm{i}\lambda) = \frac{1}{\pi}\int_0^\infty \frac{\partial_\mathrm{i}\alpha(\tau)}{\tau - \mathrm{i}\lambda} d\tau = \begin{cases} O(\log|\lambda|^{-1}), & \lambda \to 0, \\ O(|\lambda|^{1-2H_0}\log|\lambda|), & \lambda \to \pm\infty. \end{cases} \tag{6.36}$$

Combining this estimate with (6.15) gives

$$\partial_i \frac{1}{X(\mathrm{i}\lambda)} = -\frac{\partial_i \log X(\mathrm{i}\lambda)}{X(\mathrm{i}\lambda)} = \begin{cases} O(|\lambda|^{H_0-1/2}\log|\lambda|^{-1}), & \lambda \to 0, \\ O(|\lambda|^{1-2H_0}\log|\lambda|), & \lambda \to \pm\infty. \end{cases} \tag{6.37}$$

Consequently, in view of formula (6.4),

$$f(\lambda) = \begin{cases} O(\log^2|\lambda|^{-1}), & \lambda \to 0, \\ O(|\lambda|^{2-4H_0}\log^2|\lambda|), & \lambda \to \pm\infty, \end{cases}$$

so that $f \in L^1(\mathbb{R})$ and

$$|Q_{ij}(t-s)| \leqslant \|f\|_1.$$

Similarly we can estimate the derivative $f'(\lambda)$ with respect to λ,

$$f'(\lambda) = \begin{cases} O(|\lambda|^{-1}\log^2|\lambda|^{-1}) & \lambda \to 0, \\ O(|\lambda|^{1-4H_0}\log^2|\lambda|), & \lambda \to \pm\infty. \end{cases}$$

Standard bounds for the Fourier integral of such functions [23] imply

$$|Q_{ij}(t-s)| \leqslant \left|\int_{-\infty}^\infty f(\lambda)e^{\mathrm{i}(t-s)\lambda}d\lambda\right| \leqslant c|t-s|^{-1}\big|\log|t-s|\big|^3,$$

for some constant $c > 0$. The claimed estimate follows by combining the two bounds. ☐

The next lemma proves the second bound in (5.12).

Lemma 6.8. *There exist constants $b \in (0, \frac{1}{2})$, $C > 0$ and $T_{\min} > 0$ such that for all $s,t \geqslant T_{\min}$,*

$$\begin{aligned} \big|R^{(1)}_{ij}(s,t)\big| &\leqslant Cs^{-1/2}, \\ \big|R^{(2)}_{ij}(s,t)\big| &\leqslant Ct^{-1/2}, \\ \big|R^{(3)}_{ij}(s,t)\big| &\leqslant C(st)^{-b}. \end{aligned} \tag{6.38}$$

Proof. The expression in (6.23) satisfies the bound

$$\begin{aligned}\big|X(\mathrm{i}\lambda)\partial_i\widehat{R}_t(\mathrm{i}\lambda)\big| \leqslant \Big|\big(p_t(-\mathrm{i}\lambda)+q_t(-\mathrm{i}\lambda)+1\big)\partial_i\log X(\mathrm{i}\lambda)\Big|+\\ \Big|\big(p_t(\mathrm{i}\lambda)-q_t(\mathrm{i}\lambda)\big)\partial_i\log X(\mathrm{i}\lambda)\Big|+2\big|\partial_i p_t(\mathrm{i}\lambda)\big|+2\big|\partial_i q_t(\mathrm{i}\lambda)\big|,\end{aligned} \tag{6.39}$$

where we used the conjugacy $\overline{X(\mathrm{i}\lambda)}=X(-\mathrm{i}\lambda)$. Thus the expression for $R^{(1)}_{ij}(s,t)$ in (6.33) satisfies

$$\begin{aligned}\big|R^{(1)}_{ij}(s,t)\big| \leqslant \int_{-\infty}^{\infty}\Big|\partial_j\frac{1}{X(-\mathrm{i}\lambda)}\overline{\partial_i\widehat{R}_s(\mathrm{i}\lambda)}\Lambda(\mathrm{i}\lambda)\Big|d\lambda = \\ \int_{-\infty}^{\infty}\Big|\partial_j\log X(\mathrm{i}\lambda)\Big|\Big|X(\mathrm{i}\lambda)\partial_i\widehat{R}_s(\mathrm{i}\lambda)\Big|\Big|\frac{\Lambda(\mathrm{i}\lambda)}{X(-\mathrm{i}\lambda)X(\mathrm{i}\lambda)}\Big|d\lambda \leqslant \\ 2\int_{-\infty}^{\infty}f_1(\lambda)\Big(\big|\partial_i p_s(\mathrm{i}\lambda)\big|+\big|\partial_i q_s(\mathrm{i}\lambda)\big|\Big)d\lambda+ \\ 2\int_{-\infty}^{\infty}f_2(\lambda)\Big(\big|p_s(\mathrm{i}\lambda)+\tfrac12\big|+\big|q_s(\mathrm{i}\lambda)+\tfrac12\big|\Big)d\lambda,\end{aligned} \tag{6.40}$$

where we used (6.39) and defined

$$\begin{aligned}f_1(\lambda) &:= \varepsilon\Big|\partial_j\log X(\mathrm{i}\lambda)\Big|,\\ f_2(\lambda) &:= \varepsilon\Big|\partial_j\log X(\mathrm{i}\lambda)\partial_i\log X(\mathrm{i}\lambda)\Big|.\end{aligned}$$

Due to the estimate (6.36),

$$f_1(\lambda)=\begin{cases}O\big(\log|\lambda|^{-1}\big), & \lambda\to 0,\\ O\big(|\lambda|^{1-2H_0}\log|\lambda|\big), & \lambda\to\pm\infty,\end{cases}$$

and

$$f_2(\lambda)=\begin{cases}O\big(\log^2|\lambda|^{-1}\big), & \lambda\to 0,\\ O\big(|\lambda|^{2-4H_0}\log^2|\lambda|\big), & \lambda\to\pm\infty.\end{cases}$$

Thus $f_1, f_2\in L^2(\mathbb{R})$. By estimate (6.26) with $r=0$,

$$\int_{-\infty}^{\infty}f_1(\lambda)\big|\partial_i p_s(\mathrm{i}\lambda)\big|d\lambda \leqslant \|f_1\|\,\|\partial_i p_s\|\leqslant Cs^{-1/2}.$$

The same estimate is valid for the rest of the integrals in (6.40) and the first bound in (6.38) follows. The second bound is proved similarly. To prove the third bound, note that

$$\begin{aligned}\big|R^{(3)}_{ij}(s,t)\big| \leqslant \int_{-\infty}^{\infty}\Big|X(\mathrm{i}\lambda)\partial_j\widehat{R}_t(\mathrm{i}\lambda)X(\mathrm{i}\lambda)\partial_i\widehat{R}_s(\mathrm{i}\lambda)\Big|d\lambda \leqslant \\ \left(\int_{-\infty}^{\infty}\Big|X(\mathrm{i}\lambda)\partial_j\widehat{R}_t(\mathrm{i}\lambda)\Big|^2d\lambda\right)^{1/2}\left(\int_{-\infty}^{\infty}\Big|X(\mathrm{i}\lambda)\partial_i\widehat{R}_s(\mathrm{i}\lambda)\Big|^2d\lambda\right)^{1/2}.\end{aligned} \tag{6.41}$$

In view of (6.39),

$$\begin{aligned}&\int_{-\infty}^{\infty}\Big|X(\mathrm{i}\lambda)\partial_j\widehat{R}_t(\mathrm{i}\lambda)\Big|^2d\lambda\leqslant\\&4\int_{-\infty}^{\infty}\Big|\big(p_t(-\mathrm{i}\lambda)+q_t(-\mathrm{i}\lambda)+1\big)\partial_i\log X(\mathrm{i}\lambda)\Big|^2d\lambda+\\&4\int_{-\infty}^{\infty}\Big|\big(p_t(\mathrm{i}\lambda)-q_t(\mathrm{i}\lambda)\big)\partial_i\log X(\mathrm{i}\lambda)\Big|^2d\lambda+\\&8\int_{-\infty}^{\infty}\Big|\partial_ip_t(\mathrm{i}\lambda)\Big|^2d\lambda+8\int_{-\infty}^{\infty}\Big|\partial_iq_t(\mathrm{i}\lambda)\Big|^2d\lambda.\end{aligned}\tag{6.42}$$

Due to (6.36), $\big|\partial_i\log X(\mathrm{i}\lambda,\eta)\big|^2\leqslant C|\lambda|^{-r}$ for any $r\in(0,1)$. Hence, with $r>0$ small enough, Lemma 6.6 guarantees that all the integrals in (6.42) are bounded by Ct^{r-1}. Applying the same argument to the second term in (6.41) we conclude that

$$\big|R^{(3)}_{ij}(s,t)\big|\leqslant Cs^{r/2-1/2}t^{r/2-1/2}.$$

This verifies the last bound in (6.38) with $b:=1/2-r/2\in(0,1/2)$. □

Finally, the next lemma verifies formula (5.13).

Lemma 6.9.

$$Q_{ij}(0)=\frac{\varepsilon}{4\pi}\int_{-\infty}^{\infty}\partial_i\log\big(\varepsilon+\widehat{K}_{\theta_0}(\lambda)\big)\partial_j\log\big(\varepsilon+\widehat{K}_{\theta_0}(\lambda)\big)d\lambda.$$

Proof. In view of (6.16) the expression in (6.32) can be written as

$$\begin{aligned}Q_{ij}(0)=&\frac{1}{2\pi}\int_{-\infty}^{\infty}\partial_j\frac{1}{X(-\mathrm{i}\lambda)}\partial_i\frac{1}{X(\mathrm{i}\lambda)}\Lambda(\mathrm{i}\lambda)d\lambda=\\&\frac{\varepsilon}{2\pi}\int_{-\infty}^{\infty}\partial_i\log X(\mathrm{i}\lambda)\partial_j\log X(-\mathrm{i}\lambda)d\lambda.\end{aligned}$$

On the other hand,

$$\begin{aligned}&\frac{\varepsilon}{4\pi}\int_{-\infty}^{\infty}\partial_i\log\Lambda(\mathrm{i}\lambda)\partial_j\log\Lambda(\mathrm{i}\lambda)d\lambda=\\&\frac{\varepsilon}{2\pi}\int_{-\infty}^{\infty}\partial_i\log X(\mathrm{i}\lambda)\partial_j\log X(-\mathrm{i}\lambda)d\lambda+\frac{\varepsilon}{2\pi}\int_{-\infty}^{\infty}\partial_i\log X(\mathrm{i}\lambda)\partial_j\log X(\mathrm{i}\lambda)d\lambda.\end{aligned}$$

Hence the formula in question is true if we show that the latter integral vanishes. In view of (6.14),

$$\begin{aligned}&\int_{-\infty}^{\infty}\partial_i\log X(\mathrm{i}\lambda)\partial_j\log X(\mathrm{i}\lambda)d\lambda=\\&\frac{1}{\pi^2}\int_0^{\infty}\int_0^{\infty}\partial_{\mathrm{i}}\alpha(\tau)\partial_j\alpha(r)\left(\int_{-\infty}^{\infty}\frac{1}{\tau-\mathrm{i}\lambda}\frac{1}{r-\mathrm{i}\lambda}d\lambda\right)d\tau dr=0.\end{aligned}$$

The last equality holds since for any $r,\tau\in\mathbb{R}_+$ the integral in the brackets vanishes, as can be readily checked by the standard contour integration. □

6.3. Proof of Lemma 5.2

This lemma involves only one dimensional distributions of the process $\rho_t(X,\theta)$ and its partial derivatives. On the other hand, unlike in Lemma 5.1, θ may be distinct from θ_0, the true value of the parameter, which determines the distribution of the sample X^T. In this subsection, we will stress this distinction by adding the relevant parameter value to the notations.

We have to show that for all sufficiently small $\delta > 0$ there exist constants $C > 0$ and $T_{\min} > 0$ such that

$$\sup_{\|\theta-\theta_0\|\leqslant\delta} \mathsf{E}\big(\partial_i\partial_j\rho_t(X,\theta)\big)^2 \leqslant C, \quad \forall t \geqslant T_{\min}.$$

Similarly to (6.31)

$$\mathsf{E}\big(\partial_i\partial_j\rho_t(X,\theta)\big)^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Big|\partial_i\partial_j\widehat{g}_t(\mathrm{i}\lambda;\theta)\Big|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant$$
$$\int_{-\infty}^{\infty}\Big|\partial_i\partial_j\frac{1}{X(\mathrm{i}\lambda;\theta)}\Big|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda + \int_{-\infty}^{\infty}\Big|\partial_i\partial_j\widehat{R}_t(\mathrm{i}\lambda;\theta)\Big|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda,$$

where the bound holds due to decomposition (6.22). It remains to prove that both terms in the right hand side are bounded functions of $t \in [T_{\min},\infty)$ for some $T_{\min} > 0$, uniformly over θ in a δ-vicinity of θ_0. This is done in the following two lemmas.

Lemma 6.10. *For all sufficiently small $\delta > 0$, there exists a constant $C > 0$ such that*

$$\sup_{\|\theta-\theta_0\|\leqslant\delta}\int_{-\infty}^{\infty}\Big|\partial_i\partial_j\frac{1}{X(\mathrm{i}\lambda;\theta)}\Big|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant C.$$

Proof. The second order derivatives of $\alpha(\tau,\theta)$ defined in (6.11) are continuous in τ and satisfy, cf. (6.30)

$$\partial_i\partial_j\alpha(\tau,\theta) = \begin{cases} O(1), & \tau \to 0,\\ O(\tau^{1-2H}\log^2\tau), & \tau\to\infty.\end{cases}$$

Consequently

$$\partial_i\partial_j\log X(\mathrm{i}\lambda,\theta) = \frac{1}{\pi}\int_0^\infty \frac{\partial_i\partial_j\alpha(\tau,\theta)}{\tau-\mathrm{i}\lambda}d\tau = \begin{cases} O(\log|\lambda|^{-1}), & \lambda\to 0,\\ O(|\lambda|^{1-2H}\log^2|\lambda|), & \lambda\to\pm\infty,\end{cases}$$

and in view of (6.15) and (6.36),

$$\partial_i\partial_j\frac{1}{X(\mathrm{i}\lambda;\theta)} = \frac{1}{X(\mathrm{i}\lambda;\theta)}\Big(\partial_i\log X(\mathrm{i}\lambda;\theta)\partial_j\log X(\mathrm{i}\lambda;\theta) - \partial_i\partial_j\log X(\mathrm{i}\lambda;\theta)\Big) =$$
$$\begin{cases} |\lambda|^{H-1/2}\log^2|\lambda|^{-1}, & \lambda\to 0,\\ |\lambda|^{1-2H}\log^2|\lambda|, & \lambda\to\infty.\end{cases}$$

This estimate and (6.4) imply

$$\left|\partial_i\partial_j \frac{1}{X(\mathrm{i}\lambda;\theta)}\right|^2 \Lambda(\mathrm{i}\lambda;\theta_0) = \begin{cases} |\lambda|^{-2\delta}\log^4|\lambda|^{-1}, & \lambda\to 0,\\ |\lambda|^{2-4H}\log^4|\lambda|, & \lambda\to\infty. \end{cases} \tag{6.43}$$

This function is integrable on $\mathbb{R}$ for all sufficiently small $\delta>0$ which verifies the claim. □

Lemma 6.11. *For all sufficiently small $\delta>0$, there exist positive constants C, $T_{\min}$ and c such that*

$$\sup_{\|\theta-\theta_0\|\leqslant\delta}\int_{-\infty}^{\infty}\left|\partial_i\partial_j \widehat{R}_t(\mathrm{i}\lambda;\theta)\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant Ct^{-c}, \quad \forall t\geqslant T_{\min}.$$

Proof. In view of formula (6.23), it suffices to show that for all sufficiently small $\delta>0$, there exist positive constants C, $T_{\min}$ and c such that

$$\begin{aligned} I_1(t) &:= \int_{-\infty}^{\infty}\left|\partial_i\partial_j\frac{1}{X(\mathrm{i}\lambda,\theta)}\big(p_t(\mathrm{i}\lambda,\theta)+\tfrac12\big)\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant Ct^{-c},\\ I_2(t) &:= \int_{-\infty}^{\infty}\left|\partial_j\frac{1}{X(\mathrm{i}\lambda,\theta)}\partial_i p_t(\mathrm{i}\lambda,\theta)\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant Ct^{-c},\\ I_3(t) &:= \int_{-\infty}^{\infty}\left|\frac{1}{X(\mathrm{i}\lambda,\theta)}\partial_i\partial_j p_t(\mathrm{i}\lambda,\theta)\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)d\lambda \leqslant Ct^{-c}, \end{aligned} \tag{6.44}$$

for all θ such that $\|\theta-\theta_0\|\leqslant\delta$ and all $t\geqslant T_{\min}$. The same bounds are obviously true for $q_t(\mathrm{i}\lambda,\theta)$ and its derivatives as well.

Take an $r>0$ small enough so that the assertion of Lemma 6.6 holds. Then for any sufficiently small $\delta>0$, the estimate (6.43) implies that

$$\left|\partial_i\partial_j\frac{1}{X(\mathrm{i}\lambda;\theta)}\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)\leqslant C_1|\lambda|^{-r}$$

for some constant $C_1>0$, and the first bound in (6.44) holds with $c=1-r$ by Lemma 6.6. The second bound holds by the same argument since, in view of (6.37) and (6.4),

$$\left|\partial_j\frac{1}{X(\mathrm{i}\lambda,\theta)}\right|^2\Lambda(\mathrm{i}\lambda;\theta_0) = \begin{cases} O(|\lambda|^{-2\delta}\log^2|\lambda|^{-1}), & \lambda\to 0,\\ O(|\lambda|^{2-4H}\log^2|\lambda|), & \lambda\to\pm\infty. \end{cases}$$

To prove the third bound, note that by (6.16) and (6.4),

$$\left|\frac{1}{X(\mathrm{i}\lambda,\theta)}\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)-\varepsilon = \varepsilon\Big(\frac{\Lambda(\mathrm{i}\lambda;\theta_0)}{\Lambda(\mathrm{i}\lambda;\theta)}-1\Big) = \begin{cases} O(|\lambda|^{-2\delta}), & \lambda\to 0,\\ O(|\lambda|^{1-2H_0+2\delta}), & \lambda\to\pm\infty. \end{cases}$$

Thus

$$\begin{aligned} I_3(t) \leqslant \varepsilon&\int_{-\infty}^{\infty}\left|\partial_i\partial_j p_t(\mathrm{i}\lambda,\theta)\right|^2 d\lambda + \\ &\int_{-\infty}^{\infty}\left|\partial_i\partial_j p_t(\mathrm{i}\lambda,\theta)\right|^2\left|\left|\frac{1}{X(\mathrm{i}\lambda,\theta)}\right|^2\Lambda(\mathrm{i}\lambda;\theta_0)-\varepsilon\right|d\lambda \leqslant Ct^{-1}+Ct^{1-r}, \end{aligned}$$

where the last bound is true due to Lemma 6.6. □

7. Proofs of Theorems 3.2 and 3.4

As explained in Sect. 5.2, the LAN property in the small noise setting is derived from the large time asymptotics. It will be convenient to change some notations in order to emphasize the more relevant variables. In particular, we will indicate the dependence of solution to (5.3) on ε by the subscript and keep in mind its dependence on θ, omitting it from the notations. Thus the equation (5.3) reads

$$\varepsilon g_\varepsilon(t,s) + \int_0^t \sigma^2 c_H |s-r|^{2H-2} g_\varepsilon(t,r)dr = \sigma^2 c_H s^{2H-2}, \quad 0<s<t, \tag{7.1}$$

where we defined $c_H = H(2H-1)$.

7.1. The key lemmas

The following lemma reveals a useful relation between derivatives of $g_\varepsilon(t,s)$ with respect to the parameter and time variables.

Lemma 7.1. *The solution to* (7.1) *with* $\varepsilon = 1$ *satisfies*

$$t\frac{\partial}{\partial t} g_1(t,s) + s\frac{\partial}{\partial s} g_1(t,s) + g_1(t,s) = (2H-1)\sigma^2 \frac{\partial}{\partial \sigma^2} g_1(t,s), \quad 0<s<t.$$

Proof. The function $g_1(t,s)$ diverges to ∞ as $s\to 0$, which makes a useful differentiation formula from [40] inapplicable, cf. (7.4) below. To avoid this difficulty, define the function $h(s,t) = sg_1(t,s)$, then

$$s\frac{\partial}{\partial s} g_1(t,s) + g_1(t,s) = \frac{\partial}{\partial s}\big(sg_1(t,s)\big).$$

Multiplying the equation

$$g_1(t,s) + \int_0^t \sigma^2 c_H |s-r|^{2H-2} g_1(t,r)dr = \sigma^2 c_H s^{2H-2}, \quad 0<s<t, \tag{7.2}$$

by s and rearranging terms gives

$$h(s,t) + \int_0^t \sigma^2 c_H |s-r|^{2H-2} h(r,t)dr = \\ \sigma^2 c_H \left(\int_0^t |s-r|^{2H-2}(r-s) g_1(t,r)dr + s^{2H-1}\right). \tag{7.3}$$

The expression in the brackets in the right hand side is differentiable in s with the derivative

$$\frac{\partial}{\partial s}\left(s^{2H-1} + \int_0^t |s-r|^{2H-2}(r-s) g_1(t,r)dr\right) = \\ \frac{\partial}{\partial s}\left(s^{2H-1} - \int_0^s (s-r)^{2H-1} g_1(t,r)dr + \int_s^t (r-s)^{2H-1} g_1(t,r)dr\right) =$$

$$(2H-1)\left(s^{2H-2}-\int_0^s (s-r)^{2H-2}g_1(t,r)dr-\int_s^t (r-s)^{2H-2}g_1(t,r)dr\right)=$$
$$(2H-1)\left(s^{2H-2}-\int_0^s |s-r|^{2H-2}g_1(t,r)dr\right)=\frac{2H-1}{c_H\sigma^2}g_1(t,s).$$

Since the solution $h(s,t)$ is differentiable at any $s\in(0,t)$ (see [40])

$$\begin{aligned}&\frac{\partial}{\partial s}\int_0^t h(r,t)|s-r|^{2H-2}dr=\\&\int_0^t |s-r|^{2H-2}\frac{\partial}{\partial r}h(r,t)dr+h(0,t)s^{2H-2}-h(t,t)(t-s)^{2H-2},\end{aligned}\tag{7.4}$$

where $h(0,t)=0$ by (7.3). Thus the right hand side of (7.3) is differentiable and in view of the above formulas

$$\begin{aligned}\frac{\partial}{\partial s}h(s,t)+\int_0^t \sigma^2 c_H|s-r|^{2H-2}\frac{\partial}{\partial r}h(r,t)dr=&\\(2H-1)g_1(t,s)+\sigma^2 c_H t g_1(t,t)(t-s)^{2H-2}.&\end{aligned}\tag{7.5}$$

Arguing differentiability of $g_1(t,s)$ with respect to t as in [6, Lemma 3.5(i)] and taking the derivative of (7.2) we get

$$\frac{\partial}{\partial t}g_1(t,s)+\int_0^t \sigma^2 c_H|s-r|^{2H-2}\frac{\partial}{\partial t}g_1(t,r)dr=-\sigma^2 c_H g_1(t,t)(t-s)^{2H-2}.$$

Multiplying this equation by t and adding the result to (7.5) gives

$$\begin{aligned}&\Big(t\frac{\partial}{\partial t}g_1(t,s)+\frac{\partial}{\partial s}h(s,t)\Big)+\\&\int_0^t \sigma^2 c_H|s-r|^{2H-2}\Big(t\frac{\partial}{\partial t}g_1(t,r)+\frac{\partial}{\partial r}h(r,t)\Big)dr=(2H-1)g_1(t,s).\end{aligned}\tag{7.6}$$

On the other hand, differentiating (7.2) with respect to σ^2 shows that

$$\begin{aligned}\frac{\partial}{\partial\sigma^2}g_1(t,s)+\int_0^t \sigma^2 c_H|s-r|^{2H-2}\frac{\partial}{\partial\sigma^2}g_1(t,r)dr+&\\\int_0^t c_H|s-r|^{2H-2}g_1(t,r)dr=c_Hs^{2H-2},&\end{aligned}$$

or equivalently,

$$\frac{\partial}{\partial\sigma^2}g_1(t,s)+\int_0^t \sigma^2 c_H|s-r|^{2H-2}\frac{\partial}{\partial\sigma^2}g_1(t,r)dr=\frac{1}{\sigma^2}g_1(t,s).$$

Comparing this equation to (7.6) we conclude that

$$t\frac{\partial}{\partial t}g_1(t,s)+\frac{\partial}{\partial s}h(s,t)=(2H-1)\sigma^2\frac{\partial}{\partial\sigma^2}g_1(t,s)$$

by uniqnuness of the solution. □

The solution to (7.1) satisfies the following pivotal scaling property with respect to ε.

Lemma 7.2. *Let* $\gamma = 1/(2H-1)$ *and define*

$$M(\varepsilon,\theta) = \begin{pmatrix} 1 & -2\sigma^2 \log \varepsilon^{-\gamma} \\ 0 & 1 \end{pmatrix}, \quad \nu(\varepsilon,\theta) = \begin{pmatrix} 4\sigma^2 \log^2 \varepsilon^{-\gamma} & -2\log \varepsilon^{-\gamma} \\ -2\log \varepsilon^{-\gamma} & 0 \end{pmatrix}.$$

Then for any $\varepsilon > 0$ *and* $t > s > 0$*,*

$$g_\varepsilon(t,s) = \varepsilon^{-\gamma} g_1\big(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\big), \tag{7.7}$$

$$\nabla g_\varepsilon(t,s) = \varepsilon^{-\gamma} \nabla g_1\big(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\big) M(\varepsilon,\theta)^\top, \tag{7.8}$$

$$\begin{aligned} \nabla^2 g_\varepsilon(t,s) &= \varepsilon^{-\gamma}\nu(\varepsilon,\theta)\frac{\partial}{\partial\sigma^2} g_1\big(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\big) \\ &\quad + \varepsilon^{-\gamma} M(\varepsilon,\theta)\nabla^2 g_1\big(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\big) M(\varepsilon,\theta)^\top. \end{aligned} \tag{7.9}$$

Proof. Identity (7.7) is obtained by scaling all the variables in equation (7.2) by $\varepsilon^{-\gamma}$. To verify the identities for derivatives it will be convenient to use the short notations

$$\begin{aligned} g_1'(t,s) &:= \frac{\partial}{\partial H} g_1(t,s), \\ g_1^\bullet(t,s) &:= \frac{\partial}{\partial \sigma^2} g_1(t,s), \\ g_1''(t,s) &:= \frac{\partial^2}{\partial H^2} g_1(t,s), \\ g_1'^\bullet(t,s) &:= \frac{\partial}{\partial H}\frac{\partial}{\partial \sigma^2} g_1(t,s), \\ g_1^{\bullet\bullet}(t,s) &:= \frac{\partial^2}{\partial {\sigma^2}^2} g_1(t,s), \end{aligned}$$

and define the variables $u := s\varepsilon^{-\gamma}$ and $v := t\varepsilon^{-\gamma}$. Then

$$\frac{\partial}{\partial\sigma^2} g_\varepsilon(t,s) = \varepsilon^{-\gamma} g_1^\bullet\big(v,u\big)$$

and, in view of Lemma 7.1,

$$\begin{aligned} &\frac{\partial}{\partial H} g_\varepsilon(t,s) = \varepsilon^{-\gamma} g_1'(v,u) + \\ &\frac{\partial\gamma}{\partial H}\Big(\varepsilon^{-\gamma}\log\varepsilon^{-1} g_1(v,u) + \varepsilon^{-\gamma}\frac{\partial u}{\partial\gamma}\frac{\partial}{\partial u} g_1(v,u) + \varepsilon^{-\gamma}\frac{\partial v}{\partial\gamma}\frac{\partial}{\partial v} g_1(v,u)\Big) = \\ &\varepsilon^{-\gamma} g_1'(v,u) - 2\gamma^2\varepsilon^{-\gamma}\log\varepsilon^{-1}\Big(g_1(v,u) + u\frac{\partial}{\partial u} g_1(v,u) + v\frac{\partial}{\partial v} g_1(v,u)\Big) = \\ &\varepsilon^{-\gamma}\Big(g_1'(v,u) - 2\log\varepsilon^{-\gamma}\sigma^2 g_1^\bullet(v,u)\Big), \end{aligned}$$

which verifies (7.8). Taking another derivative with respect to H we get

$$\frac{\partial^2}{\partial H^2} g_\varepsilon(s,t) = -2\gamma^2\varepsilon^{-\gamma}\log\varepsilon^{-1}\Big(g_1'(v,u) - 2\log\varepsilon^{-\gamma}\sigma^2 g_1^\bullet(v,u)\Big) +$$

$$\varepsilon^{-\gamma}\frac{\partial}{\partial H}\Big(g_1'(v,u) - 2\log \varepsilon^{-\gamma}\sigma^2 g_1^{\bullet}(v,u)\Big).$$

Here

$$\frac{\partial}{\partial H}g_1'(v,u) = g_1''(v,u) - 2\gamma^2\log\varepsilon^{-1}\Big(u\frac{\partial}{\partial u}g_1'(v,u) + v\frac{\partial}{\partial v}g_1'(v,u)\Big).$$

By Lemma 7.1

$$v\frac{\partial}{\partial v}g_1'(v,u) + u\frac{\partial}{\partial u}g_1'(v,u) + g_1'(v,u) = 2\sigma^2 g_1^{\bullet}(v,u) + \frac{\sigma^2}{\gamma}g_1^{\bullet\prime}(v,u)$$

and hence

$$\frac{\partial}{\partial H}g_1'(v,u) = g_1''(v,u) - 2\gamma^2\log\varepsilon^{-1}\Big(2\sigma^2 g_1^{\bullet}(v,u) + \frac{\sigma^2}{\gamma}g_1^{\bullet\prime}(v,u) - g_1'(v,u)\Big).$$

Similarly,

$$\begin{aligned}
&\frac{\partial}{\partial H}\Big(\gamma g_1^{\bullet}(v,u)\Big) = -2\gamma^2 g_1^{\bullet}(v,u) + \gamma\frac{\partial}{\partial H}g_1^{\bullet}(v,u) = \\
&-2\gamma^2 g_1^{\bullet}(v,u) + \gamma\Big(g_1^{\bullet\prime}(v,u) - 2\gamma^2\log\varepsilon^{-1}\Big(u\frac{\partial}{\partial u}g_1^{\bullet}(v,u) + v\frac{\partial}{\partial v}g_1^{\bullet}(v,u)\Big)\Big).
\end{aligned}$$

By Lemma 7.1

$$v\frac{\partial}{\partial v}g_1^{\bullet}(v,u) + u\frac{\partial}{\partial u}g_1^{\bullet}(v,u) + g_1^{\bullet}(v,u) = \frac{1}{\gamma}g_1^{\bullet}(v,u) + \frac{\sigma^2}{\gamma}g_1^{\bullet\bullet}(v,u),$$

and hence

$$\begin{aligned}
\frac{\partial}{\partial H}\Big(\gamma g_1^{\bullet}(v,u)\Big) = &-2\gamma^2 g_1^{\bullet}(v,u) + \gamma g_1^{\bullet\prime}(v,u) \\
&-2\gamma^2\log\varepsilon^{-\gamma}\Big(\frac{1}{\gamma}g_1^{\bullet}(v,u) + \frac{\sigma^2}{\gamma}g_1^{\bullet\bullet}(u,v) - g_1^{\bullet}(v,u)\Big).
\end{aligned}$$

Plugging these equations we get

$$\begin{aligned}
\frac{\partial^2}{\partial H^2}g_\varepsilon(t,s) = &\,\varepsilon^{-\gamma}\Big(g_1''(v,u) - 4\log\varepsilon^{-\gamma}\sigma^2 g_1^{\bullet\prime}(v,u) + 4\sigma^4\log^2\varepsilon^{-\gamma}g_1^{\bullet\bullet}(v,u)\Big) \\
&+ 4\varepsilon^{-\gamma}\log^2\varepsilon^{-\gamma}\sigma^2 g_1^{\bullet}(v,u).
\end{aligned}$$

The other two second order derivatives are

$$\begin{aligned}
\frac{\partial^2}{(\partial\sigma^2)^2}g_\varepsilon(t,s) =&\,\varepsilon^{-\gamma}g_1^{\bullet\bullet}\big(t\varepsilon^{-\gamma}, s\varepsilon^{-\gamma}\big), \\
\frac{\partial^2}{\partial\sigma^2\partial H}g_\varepsilon(t,s) =&\,\varepsilon^{-\gamma}\Big(g_1^{\prime\bullet}(v,u) - 2\log\varepsilon^{-\gamma}\sigma^2 g_1^{\bullet\bullet}(v,u) - 2\log\varepsilon^{-\gamma}g_1^{\bullet}(v,u)\Big).
\end{aligned}$$

In matrix notation this gives (7.9). □

7.2. *Proof of Theorem 3.2*

In view of Lemma 7.2,

$$\nabla\rho^\varepsilon_t(X^\varepsilon,\theta_0)=\int_0^t \nabla g_\varepsilon(t,t-s)dX^\varepsilon_s=$$
$$\int_0^t \nabla g_\varepsilon(t,t-s)\sigma_0 dB^{H_0}_s+\sqrt{\varepsilon}\int_0^t \nabla g_\varepsilon(t,t-s)dB_s=$$
$$\int_0^t \varepsilon^{-\gamma_0}\nabla g_1\big(t\varepsilon^{-\gamma_0},(t-s)\varepsilon^{-\gamma_0}\big)M(\varepsilon,\theta_0)^\top\sigma_0 dB^{H_0}_s+$$
$$\sqrt{\varepsilon}\int_0^t \varepsilon^{-\gamma_0}\nabla g_1\big(t\varepsilon^{-\gamma_0},(t-s)\varepsilon^{-\gamma_0}\big)M(\varepsilon,\theta_0)^\top dB_s\overset{d}{=}$$
$$\varepsilon^{-\gamma_0}\varepsilon^{\gamma_0 H_0}\left(\int_0^{t\varepsilon^{-\gamma_0}}\nabla g_1\big(t\varepsilon^{-\gamma_0},t\varepsilon^{-\gamma_0}-s\big)\sigma_0 dB^{H_0}_s\right)M(\varepsilon,\theta_0)^\top+$$
$$\varepsilon^{1/2-\gamma_0}\varepsilon^{\gamma_0/2}\left(\int_0^{t\varepsilon^{-\gamma_0}}\nabla g_1\big(t\varepsilon^{-\gamma_0},t\varepsilon^{-\gamma_0}-s\big)dB_s\right)M(\varepsilon,\theta_0)^\top=$$
$$\varepsilon^{(1-\gamma_0)/2}\nabla\rho^1_{t\varepsilon^{-\gamma_0}}(X^1,\theta_0)M(\varepsilon,\theta_0)^\top,$$

where the equality in distribution holds by the self-similarity of the fBm. This equality holds simultaneously for all $t\in\mathbb{R}_+$ and hence the two processes coincide in distribution. Consequently

$$\frac{1}{\varepsilon}\phi(\varepsilon)^\top\int_0^T \nabla^\top\rho^\varepsilon_t(X^\varepsilon,\theta_0)\nabla\rho^\varepsilon_t(X^\varepsilon,\theta_0)dt\phi(\varepsilon)\overset{d}{=}$$
$$\varepsilon^{-\gamma_0}\phi(\varepsilon)^\top M(\varepsilon,\theta_0)\bigg(\int_0^T \nabla^\top\rho^1_{t\varepsilon^{-\gamma_0}}(X^1,\theta_0)\nabla\rho^1_{t\varepsilon^{-\gamma_0}}(X^1,\theta_0)dt\bigg)M(\varepsilon,\theta_0)^\top\phi(\varepsilon)=$$
$$T\varepsilon^{-\gamma_0}\phi(\varepsilon)^\top M(\varepsilon,\theta_0)\bigg(\frac{1}{T\varepsilon^{-\gamma_0}}\int_0^{T\varepsilon^{-\gamma_0}}\nabla^\top\rho^1_t(X^1,\theta_0)\nabla\rho^1_t(X^1,\theta_0)dt\bigg)M(\varepsilon,\theta_0)^\top\phi(\varepsilon).$$

In view of (5.8)

$$\frac{1}{T\varepsilon^{-\gamma_0}}\int_0^{T\varepsilon^{-\gamma_0}}\nabla^\top\rho^1_t(X^1,\theta_0)\nabla\rho^1_t(X^1,\theta_0)dt\xrightarrow[\varepsilon\to 0]{L_2(\Omega)}I(\theta_0;1),$$

and hence (5.15) holds if $\phi(\varepsilon)=\phi(\varepsilon,\theta_0)$ satisfies (3.4). It remains to show that (5.16) holds for the same choice of $\phi(\varepsilon,\theta_0)$. To this end,

$$\mathsf{E}\frac{1}{\varepsilon}\int_0^T\left(\int_0^1\int_0^\tau u^\top\phi(\varepsilon)^\top\nabla^2\rho^\varepsilon_t(X^\varepsilon,\theta_0+s\phi(\varepsilon)u)\phi(\varepsilon)udsd\tau\right)^2dt\leqslant$$
$$\frac{1}{\varepsilon}\int_0^1\int_0^T\mathsf{E}\Big(u^\top\phi(\varepsilon)^\top\nabla^2\rho^\varepsilon_t(X^\varepsilon,\theta_0+s\phi(\varepsilon)u)\phi(\varepsilon)u\Big)^2dtds.$$

Under the condition (3.4), $\|\phi(\varepsilon,\theta_0)\| = O\big(\varepsilon^{\gamma_0/2}\log\varepsilon^{-1}\big)$ and it suffices to check that

$$\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\int_0^T \mathsf{E}\Big\|\nabla^2\rho_t^\varepsilon(X^\varepsilon,\theta)\Big\|^2 dt \xrightarrow[\varepsilon\to 0]{} 0,$$

uniformly over $\{\theta : \|\theta_0-\theta\|\leqslant\delta\}$ for all $\delta>0$ small enough. Recall that

$$\nabla^2\rho_t^\varepsilon(X^\varepsilon,\theta) = \int_0^t \nabla^2 g_\varepsilon(t,t-s)\sigma_0 dB_s^{H_0} + \sqrt{\varepsilon}\int_0^t \nabla^2 g_\varepsilon(t,t-s)dB_s.$$

In view of Lemma 7.2

$$\begin{aligned}
&\int_0^t \nabla^2 g_\varepsilon(t,t-s)dB_s^{H_0} = \\
&\varepsilon^{-\gamma}M(\varepsilon,\theta)\int_0^t \nabla^2 g_1\big(t\varepsilon^{-\gamma},(t-s)\varepsilon^{-\gamma}\big)dB_s^{H_0}M(\varepsilon,\theta)^\top + \\
&\varepsilon^{-\gamma}\nu(\varepsilon,\theta)\int_0^t \frac{\partial}{\partial\sigma^2}g_1\big(t\varepsilon^{-\gamma},(t-s)\varepsilon^{-\gamma}\big)dB_s^{H_0} \overset{d}{=} \\
&\varepsilon^{\gamma H_0-\gamma}M(\varepsilon,\theta)\int_0^{t\varepsilon^{-\gamma}} \nabla^2 g_1\big(t\varepsilon^{-\gamma},t\varepsilon^{-\gamma}-s\big)dB_s^{H_0}M(\varepsilon,\theta)^\top + \\
&\varepsilon^{\gamma H_0-\gamma}\nu(\varepsilon,\theta)\int_0^{t\varepsilon^{-\gamma}} \frac{\partial}{\partial\sigma^2}g_1\big(t\varepsilon^{-\gamma},t\varepsilon^{-\gamma}-s\big)dB_s^{H_0} =: J_1(\varepsilon,t\varepsilon^{-\gamma}) + J_2(\varepsilon,t\varepsilon^{-\gamma}).
\end{aligned}$$

The first term satisfies

$$\begin{aligned}
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\int_0^T \mathsf{E}\Big\|J_1(\varepsilon,t\varepsilon^{-\gamma})\Big\|^2 dt = \\
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}T\frac{1}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|J_1(\varepsilon,t)\Big\|^2 dt \leqslant \\
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\varepsilon^{2\gamma H_0-2\gamma}\big\|M(\varepsilon,\theta)\big\|^4\frac{T}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|\int_0^t \nabla^2 g_1\big(t,t-s\big)dB_s^{H_0}\Big\|^2 dt \leqslant \\
&\varepsilon^{2(\gamma_0-\gamma)+2(H_0-H)\gamma+\gamma}\log^8\varepsilon^{-1}TC \xrightarrow[\varepsilon\to 0]{} 0
\end{aligned}$$

where the last bound is due to Lemma 5.2 and the convergence holds uniformly over $\{\theta : \|\theta-\theta_0\|\leqslant\delta\}$ for all sufficiently small $\delta>0$. Similarly,

$$\begin{aligned}
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\int_0^T \mathsf{E}\Big\|J_2(\varepsilon,t\varepsilon^{-\gamma})\Big\|^2 dt = \\
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}T\frac{1}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|J_2(\varepsilon,t)\Big\|^2 dt = \\
&\varepsilon^{2\gamma_0-1}\varepsilon^{2\gamma H_0-2\gamma}\log^4\varepsilon^{-1}\|\nu(\varepsilon,\theta)\|^2\frac{T}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|\int_0^t \frac{\partial}{\partial\sigma^2}g_1\big(t,t-s\big)dB_s^{H_0}\Big\|^2 dt \leqslant \\
&\varepsilon^{2(\gamma_0-\gamma)+2(H_0-H)\gamma+\gamma}\log^8\varepsilon^{-1}TC \xrightarrow[\varepsilon\to 0]{} 0,
\end{aligned}$$

where the last inequality is due to Lemma 5.1. Analogously,

$$
\begin{aligned}
&\sqrt{\varepsilon}\int_0^t \nabla^2 g_\varepsilon(t,t-s)dB_s = \\
&\sqrt{\varepsilon}\varepsilon^{-\gamma} M(\varepsilon,\theta)\Big(\int_0^t \nabla^2 g_1\big(t\varepsilon^{-\gamma},(t-s)\varepsilon^{-\gamma}\big)dB_s\Big)M(\varepsilon,\theta)^\top + \\
&\sqrt{\varepsilon}\varepsilon^{-\gamma}\nu(\varepsilon,\theta)\int_0^t \frac{\partial}{\partial\sigma^2} g_1\big(t\varepsilon^{-\gamma},(t-s)\varepsilon^{-\gamma}\big)dB_s \overset{d}{=} \\
&\varepsilon^{(1-\gamma)/2} M(\varepsilon,\theta)\Big(\int_0^{t\varepsilon^{-\gamma}} \nabla^2 g_1\big(t\varepsilon^{-\gamma},t\varepsilon^{-\gamma}-s\big)dB_s\Big)M(\varepsilon,\theta)^\top + \\
&\varepsilon^{(1-\gamma)/2}\nu(\varepsilon,\theta)\int_0^{t\varepsilon^{-\gamma}} \frac{\partial}{\partial\sigma^2} g_1\big(t\varepsilon^{-\gamma},t\varepsilon^{-\gamma}-s\big)dB_s =: J_1(\varepsilon,t\varepsilon^{-\gamma}) + J_2(\varepsilon,t\varepsilon^{-\gamma}).
\end{aligned}
$$

The first term vanishes asymptotically as $\varepsilon \to 0$,

$$
\begin{aligned}
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\int_0^T \mathsf{E}\Big\|J_1(\varepsilon,t\varepsilon^{-\gamma})\Big\|^2 dt = \\
&\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1} T\frac{1}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|J_1(\varepsilon,t)\Big\|^2 dt \leqslant \\
&\varepsilon^{2\gamma_0-1}\varepsilon^{1-\gamma}\log^4\varepsilon^{-1}\|M(\varepsilon,\theta)\|^4 T\frac{1}{T\varepsilon^{-\gamma}}\int_0^{T\varepsilon^{-\gamma}} \mathsf{E}\Big\|\int_0^t \nabla^2 g_1\big(t,t-s\big)dB_s\Big\|^2 dt \leqslant \\
&\varepsilon^{\gamma_0+(\gamma_0-\gamma)}\log^4\varepsilon^{-1}\|M(\varepsilon,\theta)\|^4 TC \xrightarrow[\varepsilon\to 0]{} 0,
\end{aligned}
$$

where the last inequality holds by Lemma 5.2. Similarly,

$$
\varepsilon^{2\gamma_0-1}\log^4\varepsilon^{-1}\int_0^T \mathsf{E}\Big\|J_2(\varepsilon,t\varepsilon^{-\gamma})\Big\|^2 dt \xrightarrow[\varepsilon\to 0]{} 0.
$$

This verifies (5.16) and completes the proof.

7.3. Proof of Corollary 3.3

For brevity denote the matrices in (3.4) by $M := M(\varepsilon,\theta_0)$ and $I := I(\theta_0;1)$. Consider the Cholesky decomposition

$$
MIM^\top = LL^\top,
$$

where L is the unique lower triangular matrix with positive diagonal entries. A simple calculation shows that

$$
L = \begin{pmatrix} m(\varepsilon)\sqrt{I_{22}} & 0 \\ \sqrt{I_{22}} & \dfrac{1}{m(\varepsilon)}\sqrt{I_{11}-I_{12}^2/I_{22}} \end{pmatrix}\big(1+o(1)\big), \quad \varepsilon\to 0
$$

where $m(\varepsilon) = 2\sigma_0^2 \log \varepsilon^{-1/(2H_0-1)}$. Hence the matrix

$$\phi(\varepsilon, \theta_0)^\top = \varepsilon^{1/(4H_0-2)} \frac{1}{\sqrt{T}} L^{-1}$$

satisfies condition (3.4). The assertion (2) of Corollary 3.3 is obtained by applying Theorem 2.2 to loss functions constant in the first variable. The assertion (1) is proved similarly, using the upper triangular Cholesky decomposition.

7.4. Proof of Theorem 3.4

For a fixed σ_0^2, the LAN property of the one dimensional family $\big(\mathsf{P}^\varepsilon_{(H,\sigma_0^2)}\big)_{H\in(3/4,1)}$ is obtained by considering the likelihood ratio (5.14) with diagonal $\phi(\varepsilon, \theta_0)$ and u restricted to the line $\{u_1 e_1 : u_1 \in \mathbb{R}\}$ where $e_1 = (1,0)^\top$. For the vectors from this subspace, the limit (5.15) holds if, cf. (3.4),

$$\varepsilon^{-1/(2H_0-1)} e_1^\top \phi(\varepsilon,\theta_0)^\top M(\varepsilon,\theta_0) T I(\theta_0;1) M(\varepsilon,\theta_0)^\top \phi(\varepsilon,\theta_0) e_1 \xrightarrow[\varepsilon\to 0]{} 1.$$

For diagonal $\phi(\varepsilon, \theta_0)$, this convergence is true if

$$\phi_{11}(\varepsilon,\theta_0) = \varepsilon^{1/(4H_0-2)} \frac{1}{2\sigma_0^2 \log \varepsilon^{-1/(2H_0-1)}} \frac{1}{\sqrt{T I_{22}(\theta_0;1)}}$$

which is the scaling claimed in (3.5). The property (5.16) continues to hold as before. This proves assertion (1) of Theorem 3.4. Assertion (2) is proved analogously, by restricting u to the subspace $\{u = u_2 e_2 : u_2 \in \mathbb{R}\}$ with $e_2 = (0,1)^\top$.

Appendix A: Rate optimal estimation in the small noise regime

In the small noise regime the optimal rates of Corollary 3.3 and Theorem 3.4 can be achieved by a modification of the estimator suggested in [18]. Let us briefly sketch the idea. Take any mother wavelet function ψ with compact support and two vanishing moments. Define its translates and dilations

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), \quad j \in \mathbb{N},\ k \in \mathbb{Z}.$$

Consider the wavelet coefficients of σB^H

$$d_{j,k} = \int_{\mathbb{R}} \psi_{j,k}(t) \sigma dB^H_t \tag{A.1}$$

and define the energy of the j-th resolution level

$$Q_j = \sum_{k=0}^{2^{j-1}-1} d_{j,k}^2.$$

Standard calculations show that these random variables satisfy

$$Q_j = \frac{\sigma^2}{2} c_H(\psi) 2^{j(2-2H)} + O_{\mathsf{P}}\big(2^{-j(4H-3)/2}\big), \quad j \to \infty, \tag{A.2}$$

where

$$c_H(\psi) = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(u)\psi(v) H(2H-1)|u-v|^{2H-2} du dv.$$

Consequently,

$$\frac{Q_{j+1}}{Q_j} = 2^{2-2H} + O_{\mathsf{P}}(2^{-j/2}), \quad j \to \infty. \tag{A.3}$$

Natural estimators for the wavelet coefficients are obtained by replacing the fBm in (A.1) with its noisy observation

$$\widetilde{d}_{j,k} := \int_{\mathbb{R}} \psi_{j,k}(t) dX_t. \tag{A.4}$$

Since $\mathsf{E}\widetilde{d}^2_{j,k} = \mathsf{E}d^2_{j,k} + \varepsilon\|\psi\|^2$ it makes sense to estimate $d^2_{j,k}$ by

$$\widehat{d^2_{j,k}} = (\widetilde{d}_{j,k})^2 - \varepsilon\|\psi\|^2$$

and, accordingly,

$$\widehat{Q}_j = \sum_{k=0}^{2^{j-1}-1} \widehat{d^2_{j,k}}.$$

In view of (A.3), the method of moments suggests the estimators

$$\widehat{H}_j = 1 - \frac{1}{2}\log_2 \frac{\widehat{Q}_{j+1}}{\widehat{Q}_j}.$$

The bias of these estimators decreases with j whereas their variance increases. In view of the residual in (A.3) and the optimal rate, known from Corollary 3.3, it is reasonable to suggest that the optimal choice of j should be such that

$$2^{-j/2} = \varepsilon^{1/(4H-2)}. \tag{A.5}$$

This choice is only an "oracle" since it requires H to be known. To mimic this choice of j, asymptotics (A.2) can be used again. To this end (A.5) can be rewritten as $2^{j(2-2H)} = 2^j\varepsilon$, which, in view of (A.2), suggests the selector

$$J^*_\varepsilon = \max\big\{\underline{J} \leqslant j \leqslant J_\varepsilon : \widehat{Q}_j \geqslant 2^j\varepsilon\big\},$$

where $J_\varepsilon = [2\log_2 \varepsilon^{-1}]$ and $\underline{J}$ is an arbitrary nonessential constant. It can be shown that with high probability J^*_ε will be close to $\frac{1}{2H-1}\log_2 \varepsilon^{-1} < J_\varepsilon$ and the ultimate estimator is set to be

$$\widehat{H}(\varepsilon) := \widehat{H}_{J^*_\varepsilon}. \tag{A.6}$$

Proposition A.1. *The estimation error* $\varepsilon^{-1/(4H-2)}(\widehat{H}(\varepsilon)-H)$ *is bounded in* P_θ*-probability, uniformly over compacts in* Θ*, as* $\varepsilon\to 0$.

Proof. (Adaptation of [18].) □

Remark A.2. In the context of Theorem 3.2, this result implies rate optimality for a particular class of loss functions of the form

$$\ell_M(u) = (|u|-M)^+ \wedge 1 \leqslant \mathbf{1}_{\{|u|\geqslant M\}},$$

since the above estimator satisfies

$$\varlimsup_{\varepsilon\to 0}\sup_{\theta\in K}\mathsf{E}_\theta \ell_M\Big(\varepsilon^{-1/(4H-2)}(\widehat{H}(\varepsilon)-H)\Big) < 1$$

for any compact $K\subset\Theta$ and all M large enough.

Estimation of σ^2 can be based on (A.2) as well. The method of [18] implies that the estimator

$$\widehat{\sigma}^2(\varepsilon) = \frac{2}{c_{\widehat{H}(\varepsilon)\wedge 1}(\psi)}\frac{\widehat{Q}_{J^*_\varepsilon}}{2^{J^*_\varepsilon(2-2\widehat{H}(\varepsilon))}}, \tag{A.7}$$

where $\widehat{H}(\varepsilon)$ is defined in (A.6), is rate optimal.

Proposition A.3. *The estimation error*

$$\varepsilon^{-1/(4H-2)}\frac{1}{\log\varepsilon^{-1}}\big(\widehat{\sigma}^2(\varepsilon)-\sigma^2\big)$$

is bounded in P_θ*-probability, uniformly over compacts in* Θ*, as* $\varepsilon\to 0$.

Similarly, the estimators

$$\widetilde{\sigma}^2(\varepsilon) = \frac{2}{c_H(\psi)}\frac{\widehat{Q}_{j_\varepsilon}(\varepsilon)}{2^{j_\varepsilon(2-2H)}}$$

with $j_\varepsilon = \Big[\frac{1}{2H-1}\log\varepsilon^{-1}\Big]$ and

$$\widetilde{H}(\varepsilon) = 1-\frac{1}{2J^*_\varepsilon}\log_2\left(\frac{2}{\sigma^2 c_{\widehat{H}(\varepsilon)\wedge 1}(\psi)}Q_{J^*_\varepsilon}\right)$$

are rate optimal for the corresponding parameter, when the other parameter is known.

A.1. A numerical illustration

Below is a numerical illustration of the estimators $\widehat{H}(\varepsilon)$ from (A.6) and $\widehat{\sigma}(\varepsilon)$ from (A.7). We used db2 Daubechies wavelet function ψ with two vanishing moments and approximated the stochastic integrals in (A.4) by the Riemann-Stieltjes sums on the uniform grid of 2^N points with $N=23$. Figure 1 depicts the empirical distributions of the estimation errors around the true values $H=0.8$ and $\sigma=1$ in $M=10{,}000$ Monte-Carlo trials for a decreasing sequence of values

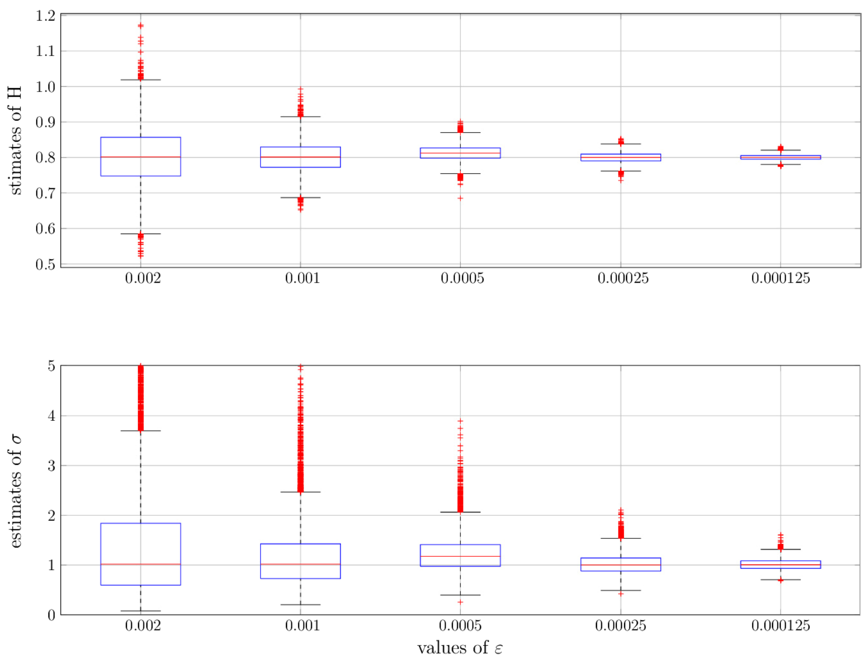

FIG 1. *Estimation errors for the estimators* (A.6) *and* (A.7) *in* $M = 10{,}000$ *Monte Carlo trials. The true values of parameters are* $H = 0.8$ *and* $\sigma = 1$.

of ε. The errors of the estimator $\widehat{H}(\varepsilon)$ appear to be evenly dispersed around the true value of H and become more concentrated as ε decreases. The empirical error distribution of the estimator $\widehat{\sigma}(\varepsilon)$ also shrinks when ε decreases but is visibly skewed towards positive errors. The large outliers correspond to those estimates of H which are close to 1. This effect is due to the function $c_H(\psi)$ in the denominator (A.7) which vanishes at $H = 1$. It becomes practically insignificant as the estimator of H gets more accurate and the denominator departs from the near zero values.

Next we approximated the Root Mean Squared (RMS) error of the estimator (A.6) by averaging over $M = 10{,}000$ Monte-Carlo experiments (so that the statistical approximation error is of order 0.01). The result is depicted in Fig. 2 in the log-log scale. The obtained plot appears to be well fitted to a straight line whose slope is predicted by Proposition A.1 to be equal to $1/(4H-2)$ asymptotically as $\varepsilon \to 0$. Remarkably the actual slope turns out to be very close to this prediction in spite of the imprecisions introduced by the Monte-Carlo averaging, discrete approximation of the stochastic integrals and finitely small values of ε.

Funding

Pavel Chigansky was supported by ISF 1383/18 grant. Marina Kleptsyna was supported by ANR EFFI grant.

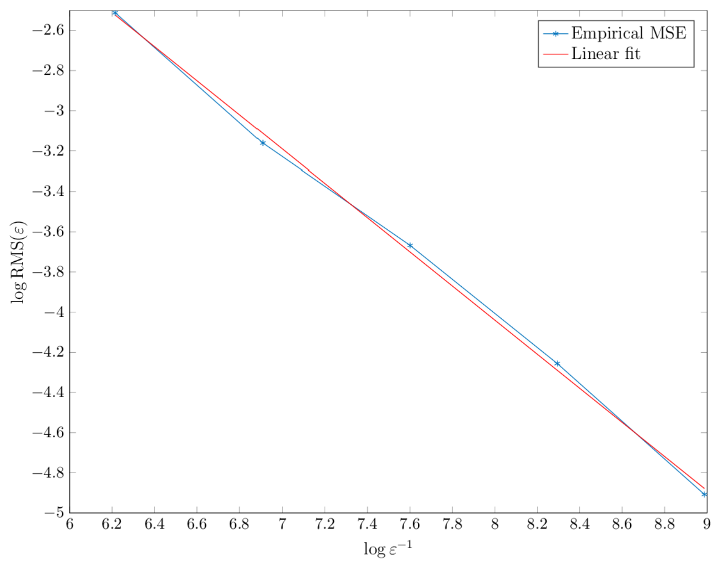

Fig 2. *The log-log plot of the empirical RMS error of the estimator* (A.6) *as a function of* ε*. The linear fit line has slope of* 0.8492 *which equals* $1/(4H-2)$ *for* $H = 0.7944$*. This fits the true value* $H = 0.8$ *with only* 0.7% *deviation.*

References

[1] Achard, S. and Coeurjolly, J.-F. (2010). Discrete variations of the fractional Brownian motion in the presence of outliers and an additive noise. *Stat. Surv.* **4** 117–147. MR2658892 (2011k:60131)

[2] Barton, R. J. and Poor, H. V. (1988). Signal detection in fractional Gaussian noise. *IEEE Trans. Inform. Theory* **34** 943–959. MR982805 (90e:94007)

[3] Baykut, S., Akgül, T. and Ergintav, S. (2007). Estimation of spectral exponent parameter of process in additive white background noise. *EURASIP Journal on Advances in Signal Processing* **2007** 063219. MR2366434

[4] Bender, C., Sottinen, T. and Valkeila, E. (2011). Fractional processes as models in stochastic finance. In *Advanced mathematical methods for finance* 75–103. Springer, Heidelberg. MR2792076

[5] Brouste, A. and Fukasawa, M. (2018). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist.* **46** 2045–2061. MR3845010

[6] Cai, C., Chigansky, P. and Kleptsyna, M. (2016). Mixed Gaussian processes: a filtering approach. *Ann. Probab.* **44** 3032–3075. MR3531685

[7] Cheridito, P. (2001). Mixed fractional Brownian motion. *Bernoulli* **7** 913–934. MR1873835 (2002k:60163)

[8] Chigansky, P. and Kleptsyna, M. (2018). Exact asymptotics in eigen-

problems for fractional Brownian covariance operators. *Stochastic Process. Appl.* **128** 2007–2059. MR3797652

[9] Chong, C., Delerue, T. and Mies, F. (2022). Rate-optimal estimation of mixed semimartingales. arXiv:2207.10464.

[10] Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Stat. Inference Stoch. Process.* **4** 199–227. MR1856174

[11] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766. MR1026311 (91f:62142)

[12] Dahlhaus, R. (2006). Correction: "Efficient parameter estimation for self-similar processes" [Ann. Statist. **17** (1989), no. 4, 1749–1766]. *Ann. Statist.* **34** 1045–1047. MR2283403 (2008a:62074)

[13] Dozzi, M., Mishura, Y. and Shevchenko, G. (2015). Asymptotic behavior of mixed power variations and statistical estimation in mixed models. *Stat. Inference Stoch. Process.* **18** 151–175. MR3348583

[14] Dufitinema, J., Shokrollahi, F., Sottinen, T. and Viitasaari, L. (2021). Long-range dependent completely correlated mixed fractional Brownian motion. arXiv:2104.04992.

[15] Dzhaparidze, K. O. and Yaglom, A. M. (1983). Spectrum parameter estimation in time series analysis. In *Developments in statistics, Vol. 4. Develop. in Statist.* **4** 1–96. Academic Press, New York. MR702676

[16] Embrechts, P. and Maejima, M. (2002). *Selfsimilar processes. Princeton Series in Applied Mathematics*. Princeton University Press, Princeton, NJ. MR1920153 (2004c:60003)

[17] Gakhov, F. D. (1990). *Boundary value problems*. Dover Publications, Inc., New York. Translated from the Russian. Reprint of the 1966 translation. MR1106850

[18] Gloter, A. and Hoffmann, M. (2007). Estimation of the Hurst parameter from discrete noisy data. *Ann. Statist.* **35** 1947–1974. MR2363959

[19] Gohberg, I. C. and Fel'dman, I. A. (1974). *Convolution equations and projection methods for their solution. Translations of Mathematical Monographs, Vol. 41*. American Mathematical Society, Providence, R.I. Translated from the Russian by F. M. Goldware. MR0355675

[20] Hitsuda, M. (1968). Representation of Gaussian processes equivalent to Wiener process. *Osaka J. Math.* **5** 299–312. MR0243614 (39 ##4935)

[21] Ibragimov, I. A. and Has'minskii, R. Z. (1981). *Statistical estimation. Asymptotic theory. Applications of Mathematics* **16**. Springer. MR0620321

[22] Inan, S., Akgül, T., Seyis, C., Saatçilar, R., Baykut, S., Ergintav, S. and Baş, M. (2008). Geochemical monitoring in the Marmara region (NW Turkey): A search for precursors of seismic activity. *Journal of Geophysical Research: Solid Earth* **113** B03401.

[23] Iosevich, A. and Liflyand, E. (2014). *Decay of the Fourier transform*. Birkhäuser/Springer, Basel. Analytic and geometric aspects. MR3308120

[24] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. H. Poincaré Probab. Statist.* **33** 407–436. MR1465796

[25] Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*, second ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Springer-Verlag, Berlin. MR1943877

[26] Kawai, R. (2013). Fisher information for fractional Brownian motion under high-frequency discrete sampling. *Comm. Statist. Theory Methods* **42** 1628–1636. MR3041490

[27] Kent, J. T. and Wood, A. T. A. (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc. Ser. B* **59** 679–699. MR1452033

[28] Kutoyants, Y. A. (2004). *Statistical inference for ergodic diffusion processes*. *Springer Series in Statistics*. Springer-Verlag London, Ltd., London. MR2144185 (2006b:62005)

[29] Liu, G. and Jing, B.-Y. (2018). On Estimation of Hurst Parameter Under Noisy Observations. *Journal of Business & Economic Statistics* **36** 483–492. MR3828975

[30] Liu, G., Zhang, L. and Wang, M. (2018). Estimation of the Hurst parameter in the simultaneous presence of jumps and noise. *Statistics* **52** 1156–1192. MR3834399

[31] Mao, A., Harrison, C. G. A. and Dixon, T. H. (1999). Noise in GPS coordinate time series. *Journal of Geophysical Research: Solid Earth* **104** 2797–2816.

[32] Nazarov, A. I. (2021). Spectral asymptotics for a class of integro-differential equations arising in the theory of fractional Gaussian processes. *Commun. Contemp. Math.* **23** Paper No. 2050049, 25. MR4289918

[33] Pipiras, V. and Taqqu, M. S. (2001). Are classes of deterministic integrands for fractional Brownian motion on an interval complete? *Bernoulli* **7** 873–897. MR1873833 (2002h:60176)

[34] Pipiras, V. and Taqqu, M. S. (2017). *Long-range dependence and self-similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics, [45]*. Cambridge University Press, Cambridge. MR3729426

[35] Rosenbaum, M. (2008). Estimation of the volatility persistence in a discretely observed diffusion model. *Stochastic Process. Appl.* **118** 1434–1462. MR2427046

[36] Sabel, T. and Schmidt-Hieber, J. (2014). Asymptotically efficient estimation of a scale parameter in Gaussian time series and closed-form expressions for the Fisher information. *Bernoulli* **20** 747–774. MR3178517

[37] Shepp, L. A. (1966). Radon-Nikodým derivatives of Gaussian measures. *Ann. Math. Statist.* **37** 321–354. MR0190999 (32 ##8408)

[38] Szymanski, G. (2022). Optimal estimation of the rough Hurst parameter in additive noise. arXiv:2205.13035.

[39] Ukai, S. (1971). Asymptotic distribution of eigenvalues of the kernel in the Kirkwood-Riseman integral equation. *J. Mathematical Phys.* **12** 83–92. MR0275084 (43 ##842)

[40] Vainikko, G. and Pedas, A. (1980/81). The properties of solutions of weakly singular integral equations. *J. Austral. Math. Soc. Ser. B* **22**

419–430. MR626933 (82i:45014)

[41] Walker, A. M. (1964). Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time-series. *J. Austral. Math. Soc.* **4** 363–384. MR0171345

[42] Whittle, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2** 423–434. MR60797

[43] Whittle, P. (1962). Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist.* **39** 105–129. MR162345

[44] Wornell, G. W. and Oppenheim, A. V. (1992). Estimation of fractal signals from noisy measurements using wavelets. *IEEE Transactions on Signal Processing* **40** 611–623.

[45] Xu, C. (2017). Detecting periodic oscillations in astronomy data: revisiting wavelet analysis with coloured and white noise. *Monthly Notices of the Royal Astronomical Society* **466** 3827–3833.