

LOCAL ASYMPTOTIC NORMALITY PROPERTY FOR FRACTIONAL GAUSSIAN NOISE UNDER HIGH-FREQUENCY OBSERVATIONS

BY ALEXANDRE BROUSTE

Département de Mathématiques, Université du Maine

AND

BY MASAOKI FUKASAWA

Graduate School of Engineering Science, Osaka University
Graduate School of Social Sciences, Tokyo Metropolitan University

Local Asymptotic Normality (LAN) property for fractional Gaussian noise under high-frequency observations is proved with non-diagonal rate matrices depending on the parameter to be estimated. In contrast to the LAN families in the literature, non-diagonal rate matrices are inevitable. As consequences of the LAN property, a maximum likelihood sequence of estimators is shown to be asymptotically efficient and the likelihood ratio test on the Hurst parameter is shown to be an asymptotically uniformly most powerful unbiased test for two-sided hypotheses.

1. Introduction. The theory of Local Asymptotic Normality (LAN) provides a powerful framework under which we are able to discuss asymptotic optimality of estimators. When the LAN property holds true for a statistical experiment, minimax theorems [13, 21] can be applied and a lower bound for the variance of the estimators can be derived. Beyond the classical IID setting [14], the LAN property (or Local Asymptotic Mixed Normality property) has been proved for various statistical models including linear processes [16], ergodic Markov chains [25], ergodic diffusions [11, 20], diffusions under high-frequency observations [12], diffusions with observational noise [9, 10, 24] and several Lévy process models [4, 18, 19]. The LAN property for fractional Gaussian noise (fGn) was obtained in [5] under the large sample observation scheme. In this setting, Maximum Likelihood (ML) and Whittle sequences of estimators achieve optimality [7, 8, 22].

The statistical experiment of observing fGn under high-frequency scheme has not been well understood in the literature, despite that high-frequency data has attracted much attention recently due to their increasing availability. At high-frequency, scaling effects from the variance and from the self-similarity of the fGn are melting. The singularity of the joint estimation

of diffusion coefficient and Hurst parameter was already noticed in [2, 15]. A weak LAN property with a singular Fisher matrix was obtained in [17]. Due to this singularity, no minimax theorem can be applied and in particular, it has been unclear whether the ML estimator enjoys any kind of optimality property.

In this paper, we prove that the statistical experiment in fact enjoys the LAN property. To discuss the difference from [17], let us be more precise in the definition of the LAN property.

DEFINITION 1.1. *Let $\Theta \subset \mathbb{R}^d$. A family of measures $\{P_\theta^n, \theta \in \Theta\}$ is called locally asymptotically normal (LAN) at $\theta_0 \in \Theta$ if there exist **nondegenerate** $d \times d$ matrices $\varphi_n(\theta_0)$ and $I(\theta_0)$ such that for any $u \in \mathbb{R}^d$, the likelihood ratio*

$$Z_n(u) = \frac{dP_{\theta_0 + \varphi_n(\theta_0)u}^n}{dP_{\theta_0}^n}$$

admits the representation

$$(1) \quad Z_n(u) = \exp\left(\langle u, \zeta_n(\theta_0) \rangle - \frac{1}{2} \langle I(\theta_0)u, u \rangle + r_n(\theta_0, u)\right),$$

where

$$(2) \quad \zeta_n(\theta_0) \rightarrow \mathcal{N}(0, I(\theta_0)), \quad r_n(\theta_0, u) \rightarrow 0$$

in law under $P_{\theta_0}^n$.

This definition of the LAN property is equivalent to the one given in [14]. The matrix $\varphi_n(\theta_0)$ is often called the rate matrix. Remark that the nondegeneracy of $\varphi_n(\theta_0)$ and $I(\theta_0)$ is essential in the following minimax theorem due to Hajek [13] and Le Cam [21], which implies in particular the asymptotic efficiency of the ML sequence of estimators in regular models; see e.g. Theorem II.12.1 and Remark II 12.2 in [14].

THEOREM 1.1. *Let the family of measures $\{P_\theta^n, \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^d$, be LAN at $\theta_0 \in \Theta$ for nondegenerate matrices $\varphi_n(\theta_0)$ and $I(\theta_0)$. Then, for any sequence of estimators $\hat{\theta}_n$,*

$$\liminf_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \sup_{|\varphi_n(\theta_0)^{-1}(\hat{\theta}_n - \theta)| \leq \delta} E_\theta^n \left[l\left(\varphi_n(\theta_0)^{-1}(\hat{\theta}_n - \theta)\right) \right] \geq \int_{\mathbb{R}^d} l(I(\theta_0)^{-1/2}z) \phi(z) dz$$

for any symmetric, nonnegative quasi-convex function l with $\lim_{|z| \rightarrow \infty} e^{-\epsilon|z|^2} l(z) = 0$ for all $\epsilon > 0$, where ϕ is the density of the d -dimensional standard normal distribution.

For the fGn under high-frequency observations, we consider the estimation of the parameters $(H, \sigma) \in \Theta = (0, 1) \times (0, \infty)$, where H is the Hurst parameter and σ is the diffusion coefficient. It was shown in [17] that the family of measures $\{P_{(H, \sigma)}^n, (H, \sigma) \in \Theta\}$ satisfies both conditions (1) and (2) at any $(H, \sigma) \in \Theta$ for

$$\varphi_n(H, \sigma) = \begin{pmatrix} \frac{1}{\sqrt{n} \log \Delta_n} & 0 \\ 0 & \frac{1}{\sqrt{n}} \end{pmatrix} \quad \text{and} \quad I(H, \sigma) = \begin{pmatrix} 2 & 2/\sigma \\ 2/\sigma & 2/\sigma^2 \end{pmatrix},$$

where n is the sample size and Δ_n is the length of sampling intervals. Note that $I(H, \sigma)$ is singular. On the one hand, this result does not imply that the family is LAN in the sense of Definition 1.1. On the second hand, Theorem 1.1 can not be applied with this result.

To the best of our knowledge, for any LAN family in the literature, it is always possible to take a diagonal rate matrix up to a reparametrization. A typical example is the LAN property for the IID setting where

$$\varphi_n(\theta_0) = \frac{1}{\sqrt{n}} \mathbf{I}_d.$$

Here \mathbf{I}_d is the $d \times d$ identity matrix. However, both in Definition 1.1 and Theorem 1.1, $\varphi_n(\theta_0)$ is required only to be nondegenerate. In this paper, we prove the LAN property for the statistical experiment of observing fGn under high-frequency scheme for a certain class of **non-diagonal** matrices $\varphi_n(\theta_0)$ depending on θ_0 . In particular, Theorem 1.1 can be applied. Non-diagonal rate matrices are inevitable because the Fisher matrix is singular.

Basics for the fractional Brownian motion and the fGn are recalled in Section 2. The statistical experiment under high-frequency scheme is described and the LAN property result is proved in Section 3. Efficient rates for the estimation of H and σ are given in Section 4 giving the optimality of the ML sequence of estimators and the asymptotically uniformly most powerful unbiased property for the likelihood ratio test on the Hurst parameter for two-sided hypotheses.

2. Fractional Brownian motion and fractional Gaussian noise. Here we briefly review the basics of the fractional Brownian motion, fractional Gaussian noise and their large sample theory. A centered Gaussian process B^H is called a fractional Brownian motion with Hurst parameter H if

$$E[B_t^H B_s^H] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t - s|^{2H})$$

for all $t, s \in \mathbb{R}$. Such a process exists and is continuous for any $H \in (0, 1)$ by Kolmogorov's extension and continuity theorems. The process enjoys a self-similarity property : for any $\Delta > 0$ and $t \in \mathbb{R}$,

$$B_{t+\Delta}^H - B_t^H \sim \Delta^H B_1^H$$

in law. The spectral density of regular increments, that is, the function f_H characterized by

$$\begin{aligned} E[B_1^H(B_{k+1}^H - B_k^H)] &= \frac{1}{2} (|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\sqrt{-1}k\lambda} f_H(\lambda) d\lambda, \quad k \in \mathbb{Z} \end{aligned}$$

is given by

$$f_H(\lambda) = C_H 2(1 - \cos(\lambda)) \sum_{k \in \mathbb{Z}} |\lambda + 2k\pi|^{-2H-1} \quad \text{with} \quad C_H = \frac{\Gamma(2H+1) \sin(\pi H)}{2\pi}.$$

For a fixed interval $\Delta > 0$, assume we observe the increments

$$X_n = (\sigma B_{\Delta}^H, \sigma B_{2\Delta}^H - \sigma B_{\Delta}^H, \sigma B_{3\Delta}^H - \sigma B_{2\Delta}^H, \dots, \sigma B_{n\Delta}^H - \sigma B_{(n-1)\Delta}^H),$$

where (H, σ) is unknown. The random vector X_n is called fractional Gaussian noise (fGn). Denote by $T_n(H)$ the covariance matrix of

$$\frac{X_n}{\sigma \Delta^H},$$

of which the distribution does not depend on σ and Δ by the self-similarity property. The (i, j) element of $T_n(H)$ is given by

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\sqrt{-1}(i-j)\lambda} f_H(\lambda) d\lambda.$$

Let us suppose $\Delta = 1$ for brevity. The log-likelihood of X_n is then given by

$$\begin{aligned} \ell_n(H, \sigma) &= -\frac{n}{2} \log(2\pi) - n \log \sigma \\ &\quad - \frac{1}{2} \log |T_n(H)| - \frac{1}{2\sigma^2} \langle X_n, T_n(H)^{-1} X_n \rangle. \end{aligned}$$

Let

$$\begin{aligned} A_n &= \sqrt{n} \left\{ \frac{1}{n\sigma^2} \langle X_n, T_n(H)^{-1} X_n \rangle - 1 \right\}, \\ B_n &= \frac{1}{\sqrt{n}} \left\{ \frac{1}{2} \partial_H \log |T_n(H)| + \frac{1}{2\sigma^2} \langle X_n, \partial_H \{T_n(H)^{-1}\} X_n \rangle \right\}. \end{aligned}$$

The score function is then given by

$$\nabla \ell_n = \begin{pmatrix} \partial_H \ell_n \\ \partial_\sigma \ell_n \end{pmatrix} = \begin{pmatrix} -B_n \sqrt{n} \\ A_n \sqrt{n} \sigma^{-1} \end{pmatrix}.$$

Let $P_{(H,\sigma)}^n$ be the measure on \mathbb{R}^n induced by X_n .

THEOREM 2.1. *The family of measures $\{P_{(H,\sigma)}^n; (H, \sigma) \in (0, 1) \times (0, \infty)\}$ is LAN at any $(H, \sigma) \in (0, 1) \times (0, \infty)$ with*

$$\varphi_n(H, \sigma) = \begin{pmatrix} \frac{1}{\sqrt{n}} & 0 \\ 0 & \frac{1}{\sqrt{n}} \end{pmatrix}$$

and

$$I(H, \sigma) = \begin{pmatrix} \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda & \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda \\ \frac{1}{2\pi\sigma} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda & \frac{2}{\sigma^2} \end{pmatrix}.$$

In particular, (A_n, B_n) converges in law to a centered normal distribution with covariance

$$(3) \quad J(H) = \begin{pmatrix} 2 & -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda \\ -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda & \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda \end{pmatrix}.$$

PROOF. See [5]. □

3. The LAN property in high-frequency observation. Let X_n be again the observed fractional Gaussian noise

$$X_n = \left(\sigma B_{\Delta_n}^H, \sigma B_{2\Delta_n}^H - \sigma B_{\Delta_n}^H, \sigma B_{3\Delta_n}^H - \sigma B_{2\Delta_n}^H, \dots, \sigma B_{n\Delta_n}^H - \sigma B_{(n-1)\Delta_n}^H \right).$$

Here we consider high-frequency asymptotics, which means $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$. We suppose $\inf_n n\Delta_n > 0$. The parameters to be estimated are still $H \in (0, 1)$ and $\sigma > 0$. As before, the distribution of

$$\frac{1}{\sigma \Delta_n^H} X_n$$

is stationary centered Gaussian and does not depend on σ and Δ_n by the self-similarity property. Its covariance matrix is $T_n(H)$ defined in the previous section. Denote by $\ell_n(H, \sigma)$ the log-likelihood of X_n

$$\begin{aligned} \ell_n(H, \sigma) &= -\frac{n}{2} \log(2\pi) - nH \log \Delta_n - n \log \sigma \\ &\quad - \frac{1}{2} \log |T_n(H)| - \frac{1}{2\sigma^2 \Delta_n^{2H}} \langle X_n, T_n(H)^{-1} X_n \rangle. \end{aligned}$$

Let

$$A_n = \sqrt{n} \left\{ \frac{1}{n\sigma^2\Delta_n^{2H}} \langle X_n, T_n(H)^{-1} X_n \rangle - 1 \right\},$$

$$B_n = \frac{1}{\sqrt{n}} \left\{ \frac{1}{2} \partial_H \log |T_n(H)| + \frac{1}{2\sigma^2\Delta_n^{2H}} \langle X_n, \partial_H \{T_n(H)^{-1}\} X_n \rangle \right\}.$$

We use the same notation as in the previous section because their distributions are the same. In particular, we have that $(A_n, B_n) \rightarrow (A, B)$ in law under $P_{(H,\sigma)}^n$ for a nondegenerate Gaussian random variable (A, B) whose covariance is given by (3). In this setting, the score function is given by

$$(4) \quad \nabla \ell_n = \begin{pmatrix} \partial_H \ell_n \\ \partial_\sigma \ell_n \end{pmatrix} = \begin{pmatrix} A_n \sqrt{n} \log \Delta_n - B_n \sqrt{n} \\ A_n \sqrt{n} \sigma^{-1} \end{pmatrix}.$$

From this expression, we clearly see that the leading terms of $\partial_H \ell_n$ and $\partial_\sigma \ell_n$ are linearly dependent, which is exactly the reason why we have a singular Fisher matrix when considering diagonal rate matrices as in [17]. Now, we state the main result of the paper.

THEOREM 3.1. *Consider a sequence of matrices*

$$\varphi_n = \varphi_n(H, \sigma) = \frac{1}{\sqrt{n}} \begin{pmatrix} \alpha_n & \bar{\alpha}_n \\ \beta_n & \bar{\beta}_n \end{pmatrix}$$

with the following properties :

1. $n|\varphi_n| = \alpha_n \bar{\beta}_n - \bar{\alpha}_n \beta_n \neq 0$.
2. $\alpha_n \rightarrow \alpha$ for some $\alpha \in \mathbb{R}$.
3. $\bar{\alpha}_n \rightarrow \bar{\alpha}$ for some $\bar{\alpha} \in \mathbb{R}$.
4. $\gamma_n := \alpha_n \log \Delta_n + \beta_n \sigma^{-1} \rightarrow \gamma$ for some $\gamma \in \mathbb{R}$.
5. $\bar{\gamma}_n := \bar{\alpha}_n \log \Delta_n + \bar{\beta}_n \sigma^{-1} \rightarrow \bar{\gamma}$ for some $\bar{\gamma} \in \mathbb{R}$.
6. $\alpha \bar{\gamma} - \bar{\alpha} \gamma \neq 0$.

Then, the family $\{P_{(H,\sigma)}^n; (H, \sigma) \in (0, 1) \times (0, \infty)\}$ is LAN at any (H, σ) for $\varphi_n(H, \sigma)$ defined previously and

$$(5) \quad I(H, \sigma) = \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} J(H) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix},$$

where $J(H)$ is defined by (3). The convergence of r_n in (2) is uniform in u on compact sets.

PROOF. Let $\theta_0 = (H_0, \sigma_0)$ and $u \in \mathbb{R}^2$. Let $\bar{B}(\theta_0, r) = \{\theta \in (0, 1) \times (0, \infty); |\theta - \theta_0| \leq r\}$ for $r > 0$. By Taylor's formula,

$$\ell_n(\theta_0 + \varphi_n u) - \ell_n(\theta_0) = \langle \varphi_n^* \nabla \ell_n(\theta_0), u \rangle + \frac{1}{2} \langle u, \varphi_n^* \nabla^2 \ell_n(\theta_0) \varphi_n u \rangle + r_n,$$

where

$$|r_n| \leq \frac{1}{6} |\varphi_n u| |u|^2 \max\{|\varphi_n^* \nabla^3 \ell_n(\theta) \varphi_n|; \theta \in \bar{B}(\theta_0, |\varphi_n u|)\}.$$

Step 1). Let us show that $\varphi_n^* \nabla \ell_n(\theta_0) \rightarrow \mathcal{N}(0, I(\theta_0))$. Note that

$$(6) \quad \varphi_n^* \nabla \ell_n = \begin{pmatrix} A_n \gamma_n - B_n \alpha_n \\ A_n \bar{\gamma}_n - B_n \bar{\alpha}_n \end{pmatrix}.$$

Therefore, it converges in law to

$$\begin{pmatrix} A\gamma - B\alpha \\ A\bar{\gamma} - B\bar{\alpha} \end{pmatrix} = \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix},$$

which is Gaussian.

Step 2). Here we compute

$$\varphi_n^* \nabla^2 \ell_n \varphi_n.$$

Let

$$\begin{aligned} C_n &= \frac{1}{n\sigma^2 \Delta_n^{2H}} \langle X_n, T_n(H)^{-1} X_n \rangle, \\ D_n &= \frac{1}{n\sigma^2 \Delta_n^{2H}} \langle X_n, \partial_H \{T_n(H)^{-1}\} X_n \rangle, \\ E_n &= \frac{1}{n} \left\{ \frac{1}{2} \partial_H^2 \log |T_n(H)| + \frac{1}{2\sigma^2 \Delta_n^{2H}} \langle X_n, \partial_H^2 \{T_n(H)^{-1}\} X_n \rangle \right\}. \end{aligned}$$

Note that

$$\partial_H \log |T_n(H)| = -\text{tr}(\partial_H \{T_n(H)^{-1}\} T_n(H))$$

and so,

$$\partial_H^2 \log |T_n(H)| = -\text{tr}(\partial_H^2 T_n(H)^{-1} T_n(H)) + \text{tr}(T_n(H)^{-1} \partial_H T_n(H) T_n(H)^{-1} \partial_H T_n(H)).$$

Then, the (1, 1) element of $\nabla^2 \ell_n \varphi_n$ is

$$\begin{aligned} & \alpha_n (\log \Delta_n \partial_H A_n - \partial_H B_n) + \beta_n \frac{\partial_H A_n}{\sigma} \\ &= \gamma_n \partial_H A_n - \alpha_n \partial_H B_n \\ &= \gamma_n \sqrt{n} (-2C_n \log \Delta_n + D_n) - \alpha_n \sqrt{n} (E_n - D_n \log \Delta_n) \end{aligned}$$

and the (1, 2) element is

$$\begin{aligned} & \alpha_n \frac{\partial_H A_n}{\sigma} + \beta_n \left(-\frac{A_n}{\sigma^2} + \frac{\partial_\sigma A_n}{\sigma} \right) \\ &= -2 \frac{\sqrt{n}}{\sigma} C_n \gamma_n + \frac{\sqrt{n}}{\sigma} D_n \alpha_n - \beta_n \frac{A_n}{\sigma^2}. \end{aligned}$$

It follows then that the (1, 1) element of $\varphi_n^* \nabla^2 \ell_n \varphi_n$ is

$$-2C_n \gamma_n^2 + 2D_n \gamma_n \alpha_n - \alpha_n^2 E_n - \beta_n^2 \frac{A_n}{\sigma^2 \sqrt{n}}.$$

From this, it is clear that the (2, 2) element of $\varphi_n^* \nabla^2 \ell_n \varphi_n$ is

$$-2C_n \bar{\gamma}_n^2 + 2D_n \bar{\gamma}_n \bar{\alpha}_n - \bar{\alpha}_n^2 E_n - \bar{\beta}_n^2 \frac{A_n}{\sigma^2 \sqrt{n}}.$$

Also it is already not difficult to see that the (1, 2) element of $\varphi_n^* \nabla^2 \ell_n \varphi_n$ is

$$-2C_n \gamma_n \bar{\gamma}_n + D_n (\bar{\gamma}_n \alpha_n + \gamma_n \bar{\alpha}_n) - \alpha_n \bar{\alpha}_n E_n - \beta_n \bar{\beta}_n \frac{A_n}{\sigma^2 \sqrt{n}}.$$

It is clear that $nC_n \sim \chi^2(n)$ and so, $C_n \rightarrow 1$. We already know $A_n \rightarrow A$. By the conditions of φ_n , $\beta_n = O(|\log \Delta_n|)$ and $\bar{\beta}_n = O(|\log \Delta_n|)$. By the same argument as the proof of Lemma 3.5 of [5], we have

$$D_n \rightarrow -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) d\lambda = E[AB]$$

and

$$E_n \rightarrow \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 d\lambda = E[B^2].$$

Therefore,

$$\begin{aligned} & (7) \\ & \varphi_n^* \nabla^2 \ell_n \varphi_n \\ & \rightarrow - \begin{pmatrix} 2\gamma^2 - 2\gamma\alpha E[AB] + \alpha^2 E[B^2] & 2\gamma\bar{\gamma} - (\gamma\bar{\alpha} + \bar{\gamma}\alpha) E[AB] + \alpha\bar{\alpha} E[B^2] \\ 2\gamma\bar{\gamma} - (\gamma\bar{\alpha} + \bar{\gamma}\alpha) E[AB] + \alpha\bar{\alpha} E[B^2] & 2\bar{\gamma}^2 - 2\bar{\gamma}\bar{\alpha} E[AB] + \bar{\alpha}^2 E[B^2] \end{pmatrix}. \end{aligned}$$

This coincides with $-I(H, \sigma)$ because $E[A^2] = 2$.

Step 3) It remains to show $r_n \rightarrow 0$. From the computation in Step 2, we

deduce that the tensor $-\varphi_n^* \nabla^3 \ell_n \varphi_n$ consists of the vectors

$$\begin{aligned} & 2\nabla C_n \gamma_n^2 - 2\gamma_n \alpha_n \nabla D_n + \alpha_n^2 \nabla E_n + \beta_n^2 \frac{\nabla A_n}{\sigma^2 \sqrt{n}}, \\ & 2\nabla C_n \gamma_n \bar{\gamma}_n - (\gamma_n \bar{\alpha}_n + \bar{\gamma}_n \alpha_n) \nabla D_n + \alpha_n \bar{\alpha}_n \nabla E_n + \beta_n \bar{\beta}_n \frac{\nabla A_n}{\sigma^2 \sqrt{n}}, \\ & 2\nabla C_n \bar{\gamma}_n^2 - 2\bar{\gamma}_n \bar{\alpha}_n \nabla D_n + \bar{\alpha}_n^2 \nabla E_n + \bar{\beta}_n^2 \frac{\nabla A_n}{\sigma^2 \sqrt{n}}, \end{aligned}$$

By the same argument as the proof of Lemma 3.7 of [5], we have that there exists $\epsilon > 0$ such that $\nabla C_n, \nabla D_n$ and ∇E_n are of $O_p(n^{1/2-\epsilon} |\log \Delta_n|)$ uniformly in $\theta \in \bar{B}(\theta_0, \epsilon)$. On the other hand,

$$\begin{aligned} \alpha_n^2 + \bar{\alpha}_n^2 &\rightarrow \alpha^2 + \bar{\alpha}^2, \\ \frac{\beta_n^2 + \bar{\beta}_n^2}{1 + \sigma^2 |\log \Delta_n|^2} &\rightarrow \alpha^2 + \bar{\alpha}^2, \end{aligned}$$

which implies that $|\varphi_n u| = O(\sqrt{|\log \Delta_n|/n})$. Since $\inf_n n \Delta_n > 0$ by the assumption, we conclude that $r_n \rightarrow 0$ uniformly in u on compact sets. \square

Several examples of rate matrices can be elicited:

1. A positive definite symmetric matrix for rate φ_n is

$$\begin{aligned} \varphi_n &= \frac{1}{\sqrt{n}} \begin{pmatrix} -\frac{2}{\log \Delta_n} & \sigma \\ \sigma & -\sigma^2 \log \Delta_n \end{pmatrix} \\ &= \frac{1}{\sqrt{n}} \begin{pmatrix} \cos \psi_n & \sin \psi_n \\ -\sin \psi_n & \cos \psi_n \end{pmatrix} \begin{pmatrix} \lambda_n^+ & 0 \\ 0 & \lambda_n^- \end{pmatrix} \begin{pmatrix} \cos \psi_n & -\sin \psi_n \\ \sin \psi_n & \cos \psi_n \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} \lambda_n^\pm &= -\frac{1}{2} \left(\frac{2}{\log \Delta_n} + \sigma^2 \log \Delta_n \pm \sqrt{\frac{4}{(\log \Delta_n)^2} + \sigma^4 (\log \Delta_n)^2} \right), \\ \tan \psi_n &= -\frac{1}{2\sigma} \left(\frac{2}{\log \Delta_n} - \sigma^2 \log \Delta_n - \sqrt{\frac{4}{(\log \Delta_n)^2} + \sigma^4 (\log \Delta_n)^2} \right), \end{aligned}$$

for which $\alpha = 0$, $\bar{\alpha} = \sigma$, $\gamma = -1$ and $\bar{\gamma} = 0$. There seems no simpler positive definite symmetric matrix satisfying the conditions of Theorem 3.1.

2. Let

$$\varphi_n = \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{\sigma \log \Delta_n}{1 + \sigma^2 (\log \Delta_n)^2} & -1 \\ \frac{1}{1 + \sigma^2 (\log \Delta_n)^2} & \sigma \log \Delta_n \end{pmatrix},$$

for which the conditions of Theorem 3.1 are satisfied with $\alpha = \bar{\gamma} = 0$, $\bar{\alpha} = -1$ and $\gamma = \sigma^{-1}$. Note that

$$\varphi_n^{-1} = \sqrt{n} \begin{pmatrix} \lambda_n^+ & 0 \\ 0 & \lambda_n^- \end{pmatrix} \begin{pmatrix} \cos \psi_n & \sin \psi_n \\ -\sin \psi_n & \cos \psi_n \end{pmatrix},$$

where

$$\lambda_n^+ = \sqrt{1 + \sigma^2 (\log \Delta_n)^2}, \quad \lambda_n^- = \frac{1}{\sqrt{1 + \sigma^2 (\log \Delta_n)^2}}$$

and

$$\cos \psi_n = \frac{\sigma \log \Delta_n}{\sqrt{1 + \sigma^2 (\log \Delta_n)^2}}, \quad \sin \psi_n = \frac{1}{\sqrt{1 + \sigma^2 (\log \Delta_n)^2}}.$$

The above expression suggests that after the rotation of angle ψ_n depending on Δ_n in a local parameter space, we have a familiar diagonal norming with, however, unfamiliar rates $\sqrt{n}\lambda_n^\pm$. Note also that by setting formally $\Delta_n = 1$, we recover Theorem 2.1.

3. Another example is that

$$\varphi_n = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 \\ \sigma(\gamma - \log \Delta_n) & \sigma(\bar{\gamma} - \log \Delta_n) \end{pmatrix}$$

for $(\gamma, \bar{\gamma})$ with $\gamma \neq \bar{\gamma}$, for which $\alpha = \bar{\alpha} = 1$.

4. Two other examples of rate matrices will be used in the following section, namely

$$(8) \quad \varphi_n = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ -\sigma \log \Delta_n & 1 \end{pmatrix},$$

which gives $\alpha = 1$, $\bar{\alpha} = 0$, $\gamma = 0$ and $\bar{\gamma} = \sigma^{-1}$ and

$$(9) \quad \varphi_n = \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{1}{\log \Delta_n} & 1 \\ 0 & -\sigma \log \Delta_n \end{pmatrix},$$

which gives $\alpha = 0$, $\bar{\alpha} = 1$, $\gamma = 1$ and $\bar{\gamma} = 0$.

Remark that all the examples are non-diagonal rate matrix depending on the parameter σ .

4. Efficient rates of estimation and optimality of ML estimators.

4.1. *The efficient estimation rate for H .* As the rate matrix, let us take (8), which gives $\alpha = 1$, $\bar{\alpha} = 0$, $\gamma = 0$ and $\bar{\gamma} = \sigma^{-1}$. It is worth mentioning that the rate matrix is not diagonal and depends on the parameter σ . By Theorem II.11.2 of [14], the LAN property implies that for any regular estimator $(\hat{H}_n, \hat{\sigma}_n)$,

$$\liminf_{n \rightarrow \infty} E_{(H, \sigma)}^n \left[l \left(\varphi_n^{-1} \begin{pmatrix} \hat{H}_n - H \\ \hat{\sigma}_n - \sigma \end{pmatrix} \right) \right] \geq E \left[l \left(\begin{pmatrix} E[B^2] & -E[AB]/\sigma \\ -E[AB]/\sigma & 2/\sigma^2 \end{pmatrix}^{-1/2} N \right) \right]$$

for any loss function l satisfying the condition given in Theorem 1.1, where $N \sim \mathcal{N}(0, I_2)$. Note that the matrix in the right hand side coincides with $I(H, \sigma)$ in Theorem 2.1. Since

$$\varphi_n^{-1} = \sqrt{n} \begin{pmatrix} 1 & 0 \\ \sigma \log \Delta_n & 1 \end{pmatrix},$$

we obtain the asymptotic lower bound

$$\liminf_{n \rightarrow \infty} n E_{(H, \sigma)} [(\hat{H}_n - H)^2] \geq \frac{2}{2E[B^2] - E[AB]^2}$$

by taking $l(x, y) = x^2$. This means that the efficient rate of estimation for H is \sqrt{n} when both H and σ are unknown. Note that when σ is known, the efficient rate for H is $\sqrt{n} \log \Delta_n$, which follows from, say, [17]. Although the singular LAN property of [17] seems to admit of a $\sqrt{n} \log \Delta_n$ -consistent estimator even when σ is unknown, the above lower bound implies that such an estimator, if any, cannot be regular. A possible relation to super-efficiency phenomena is remained as a topic for future research.

4.2. *The efficient estimation rate for σ .* As the rate matrix, let us take (9), which gives $\alpha = 0$, $\bar{\alpha} = 1$, $\gamma = 1$ and $\bar{\gamma} = 0$. It is worth repeating that the rate matrix is not diagonal and depends on the parameter σ . Again by Theorem II.11.2 of [14], the LAN property implies that for any regular estimator $(\hat{H}_n, \hat{\sigma}_n)$,

$$\liminf_{n \rightarrow \infty} E_{(H, \sigma)}^n \left[l \left(\varphi_n^{-1} \begin{pmatrix} \hat{H}_n - H \\ \hat{\sigma}_n - \sigma \end{pmatrix} \right) \right] \geq E \left[l \left(\begin{pmatrix} 2 & -E[AB] \\ -E[AB] & E[B^2] \end{pmatrix}^{-1/2} N \right) \right]$$

for any loss function l satisfying the condition given in Theorem 1.1, where $N \sim \mathcal{N}(0, I_2)$. Since

$$\varphi_n^{-1} = -\frac{\sqrt{n}}{\sigma} \begin{pmatrix} -\sigma \log \Delta_n & -1 \\ 0 & \frac{1}{\log \Delta_n} \end{pmatrix},$$

we obtain the asymptotic lower bound

$$\liminf_{n \rightarrow \infty} \frac{n}{\sigma^2 |\log \Delta_n|^2} E_{(H, \sigma)} [(\hat{\sigma}_n - \sigma)^2] \geq \frac{2}{2E[B^2] - E[AB]^2}$$

by taking $l(x, y) = y^2$. This means that the efficient rate of estimation for σ is $\sqrt{n}/|\log \Delta_n|$ when both H and σ are unknown. Note that when H is known, the efficient rate for σ is \sqrt{n} , which follows from, say, [17]. Here applies a similar remark to the end of the previous subsection.

4.3. *ML estimators.* We prove, in this section, that a ML sequence of estimators $(\hat{H}_n, \hat{\sigma}_n)$ is asymptotically efficient in that the minimax bound in Theorem 1.1 is attained. By (4), it is given as a root of $A_n = B_n = 0$ if any. Therefore,

$$\hat{\sigma}_n^2 = \frac{1}{n\Delta_n^{2\hat{H}_n}} \langle X_n, T_n(\hat{H}_n)^{-1} X_n \rangle, \quad \partial_H \log |T_n(\hat{H}_n)| + n \frac{\langle X_n, \partial_H \{T_n(\hat{H}_n)^{-1}\} X_n \rangle}{\langle X_n, T_n(\hat{H}_n)^{-1} X_n \rangle} = 0.$$

The latter equation implies that, by the self-similarity property, the limit law of \hat{H}_n is the same as in the large sample case treated in Section 2. In particular $\hat{H}_n - H = O_p(n^{-1/2})$; see e.g., Dahlhaus [7]. On the other hand, the limit law of $\hat{\sigma}_n$ is different because

$$\frac{\hat{\sigma}_n - \sigma}{\sigma} \approx \log \hat{\sigma}_n - \log \sigma \approx -(\hat{H}_n - H) \log \Delta_n$$

as is easily seen. In particular, $\hat{\sigma}_n - \sigma = O_p(|\log \Delta_n| n^{-1/2})$ and the joint law of

$$\left(\sqrt{n}(\hat{H}_n - H), \frac{\sqrt{n}}{\log \Delta_n} (\hat{\sigma}_n - \sigma) \right)$$

is degenerate to a one-dimensional distribution.

THEOREM 4.1. *Let $\theta = (H, \sigma)^* \in (0, 1) \times (0, \infty)$ and $\varphi_n = \varphi_n(H, \sigma)$ be a sequence of matrices satisfying the conditions of Theorem 3.1. Let $\hat{\theta}_n = (\hat{H}_n, \hat{\sigma}_n)^*$ be such a sequence that*

$$\varphi_n^* \nabla \ell_n(\hat{\theta}_n) = o_p(1), \quad \hat{\theta}_n - \theta = O_p(|\log \Delta_n|^k n^{-1/2})$$

for some $k \in \mathbb{N}$. Then,

$$\varphi_n^{-1}(\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(0, I(H, \sigma)^{-1})$$

in law under $P_{(H,\sigma)}^n$ and

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{|\varphi_n^{-1}(\tau - \theta)| \leq \delta} E_{\tau}^n[l(\varphi_n^{-1}(\hat{\theta}_n - \tau))] = \int_{\mathbb{R}^2} l(I(H, \sigma)^{-1/2}z) \phi(z) dz$$

for any bounded continuous function l on \mathbb{R}^2 , where $I(H, \sigma)$ is defined by (5).

PROOF. By Taylor's formula,

$$\varphi_n^* \nabla \ell_n(\hat{\theta}_n) = \varphi_n^* \nabla \ell_n(\theta) + \varphi_n^* \nabla^2 \ell_n(\theta)(\hat{\theta}_n - \theta) + r_n,$$

where

$$|r_n| \leq \frac{1}{2} |\hat{\theta}_n - \theta| |\varphi_n^{-1}(\hat{\theta}_n - \theta)| \max\{|\varphi_n^* \nabla^3 \ell_n(\theta) \varphi_n|; \theta \in \bar{B}(\theta_0, |\hat{\theta}_n - \theta|)\}.$$

It is not difficult to see $\alpha_n \bar{\beta}_n - \bar{\alpha}_n \beta_n \rightarrow \sigma(\alpha \bar{\gamma} - \bar{\alpha} \gamma)$ and so, $\|\varphi_n^{-1}\| = O(|\log \Delta_n| n^{1/2})$. Therefore, by the same argument as Step 3 in the proof of Theorem 3.1, we can show that $r_n = o_p(1)$. Then, by (6) and (7), we have

$$(10) \quad \varphi_n^{-1}(\hat{\theta}_n - \theta) \approx -\{\varphi_n^* \nabla^2 \ell_n \varphi_n\}^{-1} \varphi_n^* \nabla \ell_n \approx I(H, \sigma)^{-1} \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} \begin{pmatrix} A_n \\ B_n \end{pmatrix}.$$

From this expression the first assertion of the theorem follows. Further, since

$$(11) \quad \log \frac{dP_{\theta + \varphi_n u}^n}{dP_{\theta}^n} \approx \langle \varphi_n^* \nabla \ell_n, u \rangle - \frac{1}{2} \langle u, I(H, \sigma) u \rangle \\ \approx \left\langle \begin{pmatrix} A_n \\ B_n \end{pmatrix}, \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix} u \right\rangle - \frac{1}{2} \langle u, I(H, \sigma) u \rangle$$

uniformly in u on compact sets as we have seen in the proof of Theorem 3.1, the joint law of

$$\left(\varphi_n^{-1}(\hat{\theta}_n - \theta), \log \frac{dP_{\theta + \varphi_n u_n}^n}{dP_{\theta}^n} \right)$$

under P_{θ}^n converges to a normal distribution with covariance

$$E \left[I(H, \sigma)^{-1} \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} (A, B) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix} u \right] = u$$

for any sequence u_n with $\lim_{n \rightarrow \infty} u_n = u$. Then, by Le Cam's third lemma,

$$\varphi_n^{-1}(\hat{\theta}_n - \theta) \rightarrow \mathcal{N}(u, I(H, \sigma)^{-1})$$

in law under $P_{\theta + \varphi_n u_n}^n$, from which the second assertion follows. \square

4.4. *Hypothesis testing.* By the standard argument, as a corollary of Theorems 3.1 and 4.1, we have a limit theorem for the likelihood ratio test statistic Λ_n for the hypothesis testing $\theta = \theta_0$ v.s. $\theta \neq \theta_0$, that is,

$$-2 \log \Lambda_n \approx -\langle \varphi_n^{-1}(\hat{\theta}_n - \theta_0), \varphi_n^* \nabla^2 \ell_n \varphi_n^{-1}(\hat{\theta}_n - \theta_0) \rangle \rightarrow \chi(2)$$

under the null hypothesis, where $\hat{\theta}_n$ is the ML estimator. On the other hand, for a composite null hypothesis, the usual argument fails due to the singularity of the asymptotic Fisher matrix. Still, we have the following result.

THEOREM 4.2. *Consider the hypothesis testing $H = H_0$ v.s. $H \neq H_0$ with σ unknown. Let Λ_n be the likelihood ratio statistic, that is,*

$$-\log \Lambda_n = \ell_n(\hat{\theta}_n) - \ell_n((H_0, \tilde{\sigma}_n(H_0))^*),$$

where $\hat{\theta}_n$ is the ML estimator of $\theta = (H, \sigma)^*$ and $\tilde{\sigma}_n(H_0)$ is the ML estimator of σ under the null hypothesis $H = H_0$. Then,

$$-2 \log \Lambda_n \rightarrow \chi(1)$$

in law under the null hypothesis.

PROOF. Let φ_n be a sequence of matrices satisfying the conditions of Theorem 3.1. Then, by (10) and $\langle \nabla \ell_n(\hat{\theta}_n), (\hat{\theta}_n - \theta) \rangle = \langle \varphi_n^* \nabla \ell_n(\hat{\theta}_n), \varphi_n^{-1}(\hat{\theta}_n - \theta) \rangle \approx 0$,

$$\begin{aligned} 2(\ell_n(\hat{\theta}_n) - \ell_n(\theta)) &\approx \langle \varphi_n^{-1}(\hat{\theta}_n - \theta), I(H, \sigma) \varphi_n^{-1}(\hat{\theta}_n - \theta) \rangle \\ &\rightarrow \left\langle \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix}^{-1} J(H)^{-1} \begin{pmatrix} A \\ B \end{pmatrix}, I(H, \sigma) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix}^{-1} J(H)^{-1} \begin{pmatrix} A \\ B \end{pmatrix} \right\rangle \\ &= (A, B) J(H)^{-1} \begin{pmatrix} A \\ B \end{pmatrix} \end{aligned}$$

and $(A, B)^* \sim J(H)^{1/2} N_2$, $N_2 \sim \mathcal{N}(0, I_2)$. On the other hand,

$$\frac{1}{n} \partial_\sigma^2 \ell_n = -\frac{1}{\sigma^2 \sqrt{n}} (A_n + 2 \sqrt{n} C_n) \rightarrow -\frac{2}{\sigma^2}, \quad \sqrt{n}(\tilde{\sigma}_n(H)^2 - \sigma^2) = \sigma^2 A_n$$

and so,

$$\begin{aligned} 2(\ell_n(H, \tilde{\sigma}_n(H))^* - \ell_n(\theta)) &\approx \frac{2}{\sigma^2} |\sqrt{n}(\tilde{\sigma}_n(H) - \sigma)|^2 \\ &\rightarrow \frac{1}{2} A^2. \end{aligned}$$

Therefore,

$$\begin{aligned}
(12) \quad -2 \log \Lambda_n &= 2(\ell_n(\hat{\theta}_n) - \ell_n(\theta)) - 2(\ell_n(H, \tilde{\sigma}_n(H))^* - \ell_n(\theta)) \\
&\rightarrow (A, B) \left\{ J(H)^{-1} - \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\} \begin{pmatrix} A \\ B \end{pmatrix} \\
&= N_2^* \left\{ I_2 - \frac{1}{2} J(H)^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} J(H)^{1/2} \right\} N_2 \sim \chi(1).
\end{aligned}$$

Here we have used that the (1, 1) element of $J(H)$ is 2 and so,

$$(13) \quad \left| I_2 - \frac{1}{2} J(H)^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} J(H)^{1/2} \right| = 0, \quad \text{tr} \left(I_2 - \frac{1}{2} J(H)^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} J(H)^{1/2} \right) = 1.$$

□

In order to argue asymptotic optimality of tests, it is conventional to compare their asymptotic powers under local alternatives. For the LAN family with non-diagonal rate matrices, it is not trivial at a glance which rate matrix we should take to define local alternatives. For our hypotheses $H = H_0$ v.s. $H \neq H_0$, it would be natural to take (8) with unknown true value $\sigma = \sigma_0$ under which

$$\begin{pmatrix} H \\ \sigma \end{pmatrix} = \begin{pmatrix} H_0 \\ \sigma_0 \end{pmatrix} + \varphi_n \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \sqrt{n} \begin{pmatrix} H - H_0 \\ (H - H_0)\sigma_0 \log \Delta_n + \sigma - \sigma_0 \end{pmatrix}$$

so that the local null space is $\mathcal{N} := \{(u_1, u_2)^* | u_1 = 0\}$. For other φ_n than (8), the corresponding local null space depends on n and so, seems difficult to treat. For $\epsilon \in (0, 1)$, we say a test ψ_n is an asymptotically unbiased level ϵ test at σ_0 if

$$\limsup_{n \rightarrow \infty} E_{\theta_0 + \varphi_n u}[\psi_n] \leq \min \left\{ \epsilon, \liminf_{n \rightarrow \infty} E_{\theta_0 + \varphi_n \bar{u}}[\psi_n] \right\}$$

for every $u \in \mathcal{N}$ and $\bar{u} \in \mathcal{N}^c$, where $\theta_0 = (H_0, \sigma_0)^*$. We say a test ψ_n is an asymptotically uniformly most powerful unbiased level ϵ test at σ_0 if it is an asymptotically unbiased level ϵ test at σ_0 and

$$\liminf_{n \rightarrow \infty} E_{\theta_0 + \varphi_n \bar{u}}[\psi_n] \geq \limsup_{n \rightarrow \infty} E_{\theta_0 + \varphi_n \bar{u}}[\psi'_n]$$

for any $\bar{u} \in \mathcal{N}^c$ and any other asymptotically unbiased level ϵ test ψ'_n at σ_0 .

THEOREM 4.3. *Let Λ_n be the likelihood ratio statistic of Theorem 4.2. The test*

$$\psi_n = 1_{\{-2 \log \Lambda_n \geq z_{\epsilon/2}^2\}}, \quad 1 - \Phi(z_{\epsilon/2}) = \epsilon/2$$

is an asymptotically uniformly most powerful unbiased level ϵ test at σ_0 for any $\sigma_0 > 0$.

PROOF. By Theorem 2 of Choi et al. [3],

$$(14) \quad \limsup_{n \rightarrow \infty} E_{\theta_0 + \varphi_n \bar{u}}[\psi'_n] \leq P(|N_1 + \sqrt{\kappa} \bar{u}_1| \geq z_{\epsilon/2})$$

for any $\bar{u} = (\bar{u}_1, \bar{u}_2)^* \in \mathcal{N}^c$ and any asymptotically unbiased level ϵ test ψ'_n at σ_0 , where $N_1 \sim \mathcal{N}(0, 1)$ and κ is the effective information

$$\kappa = \frac{|I(H_0, \sigma_0)|}{I(H_0, \sigma_0)_{2,2}} = \frac{\sigma^2}{2} \left| \begin{pmatrix} E[B^2] & -E[AB]/\sigma \\ -E[AB]/\sigma & 2/\sigma^2 \end{pmatrix} \right| = E[B^2] - \frac{1}{2}E[AB]^2.$$

Here we have used that $\alpha = 1$, $\bar{\alpha} = 0$, $\gamma = 0$ and $\bar{\gamma} = \sigma^{-1}$ for (8). Actually Choi et al. [3] assumed the LAN property with a diagonal rate matrix; however their proof remains valid for non-diagonal cases as long as the local null space is of the form $\mathcal{N} = \{(u_1, u_2)^* | u_1 = 0\}$. On the other hand, by Theorem 4.2, the test ψ_n is of asymptotic level ϵ . Further by (11), Le Cam's third lemma implies that $(A_n, B_n)^* \rightarrow \mathcal{N}(\mu(\bar{u}), J(H_0))$ under $P_{\theta_0 + \varphi_n \bar{u}}$, where

$$\mu(\bar{u}) = J(H_0) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix} \bar{u} = J(H_0) \begin{pmatrix} \bar{u}_2 \sigma^{-1} \\ -\bar{u}_1 \end{pmatrix}$$

for any $\bar{u} = (\bar{u}_1, \bar{u}_2)^* \in \mathcal{N}^c$. Therefore, by the same computation as in (12), the limit law of $-2 \log \Lambda_n$ under $P_{\theta_0 + \varphi_n \bar{u}}$ is

$$\begin{aligned} & (\mu(\bar{u}) + J(H_0)^{1/2} N_2)^* \left\{ J(H_0)^{-1} - \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\} (\mu(\bar{u}) + J(H_0)^{1/2} N_2) \\ & = \kappa \bar{u}_1^2 + 2 \sqrt{\kappa} \bar{u}_1 N + N^2 = |N + \sqrt{\kappa} \bar{u}_1|^2, \end{aligned}$$

where $N_2 \sim \mathcal{N}(0, I_2)$, N is the first element of QN_2 and Q is an orthogonal matrix with

$$Q^* \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q = I_2 - \frac{1}{2} J(H)^{1/2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} J(H)^{1/2}.$$

Here we have used (13), that

$$J(H) - \frac{1}{2} J(H) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} J(H) = \begin{pmatrix} 0 & 0 \\ 0 & \kappa \end{pmatrix}$$

and that

$$x^2 = (x, y) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

for any $(x, y)^* \in \mathbb{R}^2$. This means that the upper bound of (14) is attained by ψ_n and also that ψ_n is asymptotically unbiased. \square

5. Concluding remarks. We have shown that the statistical experiment of the fGn under high-frequency observations enjoys the LAN property. The idea is the use of non-diagonal rate matrices. The self-similarity property is the key for this analysis. Therefore, an extension to the case where σ is time-dependent is nontrivial. On the other hand, extensions to more general self-similar stationary Gaussian noises or stable Lévy noises under high-frequency observations are straightforward. For the latter, the singularity of the Fisher matrix was discussed in [1] and a weak LAN property was shown in [23] which exactly corresponds to the result of [17] for the fGn.

REFERENCES

- [1] AÏT-SAHALIA, Y. AND JACOD, J. (2008) *Fisher's information for discretely sample Lévy processes*, *Econometrica*, 76(4), 727–761.
- [2] BERZIN, C. & LÉON, J. (2008) *Estimation in models driven by fractional Brownian motion*, *Annales de l'Institut Henri Poincaré – Probabilités et Statistiques*, 44, 191–213
- [3] CHOI, S., HALL, W.J. AND SCHICK, A. (1996) *Asymptotically uniformly most powerful tests in parametric and semiparametric models*, *Annals of Statistics*, 24(2), 841–861.
- [4] CLÉMENT, E. & GLOTER, A. (2016) *Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable Lévy processes*, *Stochastic Process. Appl.* 125 (2015), no. 6, 2316–2352.
- [5] COHEN, S., GAMBOA, F., LACAUX, C. & LOUBES, J.-M. (2013) *LAN property for some fractional type Brownian motion*, *ALEA : Latin American Journal of Probability and Mathematical Statistics*, 10(1), 91–106.
- [6] COEURJOLLY, J.-F. & ISTAS, J. (2001) *Cramer-Rao bounds for fractional Brownian motions*, *Statistics & Probability Letters*, 53, 435–447.
- [7] DAHLHAUS, R. (1989) *Efficient parameter estimation for self-similar processes*, *Annals of Statistics*, 17(4), 1749–1766.
- [8] DAHLHAUS, R. (2006) *Correction Efficient parameter estimation for self-similar processes*, *Annals of Statistics*, 34(2), 1045–1047.
- [9] GLOTER A. AND JACOD, J. (2001) *Diffusions with measurement errors I. Local asymptotic normality*, *ESAIM Probability and Statistics*, 5, 225–242.
- [10] GLOTER A. AND JACOD, J. (2001) *Diffusions with measurement errors II. Optimal estimators*, *ESAIM Probability and Statistics*, 5, 243–260.
- [11] GOBET, E. (2001) *Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach*, *Bernoulli*, 7(6), 899–912.
- [12] GOBET, E. (2002) *LAN property for ergodic diffusions with discrete observations*, *Annales de l'Institut Henri Poincaré*, 38(5), 711–737.
- [13] HÁJEK, J. (1972) *Local asymptotic minimax and admissibility in estimation*, *Proc. 6th Berkeley Symp. Math. Statist. Prob.*, 1, 175–194.
- [14] IBRAGIMOV, I.A. & HAS'MINSKI, R.Z. (1981). *Statistical estimation: asymptotic theory*, Springer-Verlag.
- [15] ISTAS, J. & LANG, G. (1997) *Quadratic variations and estimation of the local Hölder index of a Gaussian process*, *Annales de l'Institut Henri Poincaré – Probabilités et Statistiques*, 33(4), 407–436.

- [16] KAKIZAWA, Y. & TANIGUCHI, M (2000) *Asymptotic theory of statistical inference for time series*, Springer-Verlag, New York.
- [17] KAWAI, R. (2013) *Fisher Information for Fractional Brownian Motion Under High-Frequency Discrete Sampling*, *Communications in Statistics - Theory and Methods*, 42(9), 1628–1636.
- [18] KAWAI, R. (2013) *Local asymptotic normality property for Ornstein-Uhlenbeck processes with jumps under discrete sampling*, *J. Theoret. Probab.*, 26(4), 932–967.
- [19] KAWAI, R. & MASUDA, H. (2013) *Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling*, *ESAIM Probab. Stat.* 17, 13–32.
- [20] KUTOYANTS, Y. (2004) *Statistical inference for ergodic diffusion processes*, Springer-Verlag, London.
- [21] LE CAM, L. (1972) *Limits of experiments*, *Proc. 6th Berkeley Symp. Math. Statist. Prob.*, 1, 245–261.
- [22] LIEBERMAN, O., ROSEMARIN, R. & ROUSSEAU, J. (2012) *Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes*, *Econometric theory*, 28(02), 457–470.
- [23] MASUDA, H. (2009) *Joint estimation of discretely observed stable Lévy processes with symmetric Lévy density*, *J. Japan Statist. Soc.*, 39(1), 49–75.
- [24] REISS, M. (2011) *Asymptotic equivalence for inference on the volatility from noisy observations*, *Annals of Statistics*, 39, 772–802.
- [25] ROUSSAS, G. (1972) *Contiguity of probability measures*. Cambridge University Press.
- [26] VAN DER VAART, A. (1998) *Asymptotic Statistics*, Cambridge University Press.

AVENUE OLIVIER MESSIAEN, 72085 LE MANS CEDEX 9, FRANCE. E-MAIL: Alexandre.Brouste@univ-lemans.fr
1-3 MACHIKANNEYAMA, TOYONAKA, OSAKA, 560-8531, JAPAN. E-MAIL: fukasawa@sigmath.es.osaka-u.ac.jp