Statistical analysis of some models of fractional type process

Chunhao Cai

Thèse de Doctorat

Chunhao CAI

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université du Maine
sous le label de L'Université Nantes Angers Le Mans

École doctorale : *STIM*

Discipline : *26*
Spécialité : *mathématique*
Unité de recherche : *LMM*

Soutenue le *18 avrill 2014*

Analyse Statistique de quelques Modèles de Processus de Type Fractionnaire

JURY

Rapporteurs :	**Julia MISHURA**, Professeur, University of Kyiv, Kyiv, Ukraine **Jean-François COEURJOLLY**, MCF-HDR, Université Pierre Mendes-France, Grenoble,
Examinateurs :	**Yuri A. KUTOYANTS,** Professeur, Université du Maine, Le Mans **Mathieu ROSENBAUM**, Professeur, Université Pierre et Marie Curie, Paris **Serguï DACHIAN**, MCF-HDR, Université Blaise Pascal, Clermont-Ferrand
Directeur de Thèse :	**Marina KLEPTSYNA**, Professeur, Université du Maine, Le Mans
Co-directeur de Thèse :	**Alexandre BROUSTE**, MCF, Université du Maine, Le Mans

L'Université Nantes Angers Le Mans

Remerciements

Je tiens tout d'abord à remercier particulièrement mes directeurs de thèse, Marina Kleptsyna et Alexandre Brouste pour leurs conseils, leurs grandes disponibilités et leurs générosités avec lesquelles ils m'ont fait partager leurs travaux, leurs idées et leurs intuitions. La pertinence de leurs questions et de leurs remarques ont toujours su me motiver et me diriger dans mes recherches. Je leur témoigne enfin ma profonde reconnaissance pour m'avoir fait confiance en me proposant de venir travailler avec eux, et pour leur soutien constant au cours de ces quatres années. Je leur exprime toute mon admiration, tant sur le plan humain que sur le plan professionnel.

Je voudrais témoigner toute ma gratitude à Mishura Julia et Coeurjolly Jean-François qui ont accepté d'être les rapporteurs de ma thèse. Leur lecteur attentive et leurs remarques précieuses m'ont permis d'améliorer ce travaille. Je tiens à remercier Pavel Chigansky et Alain Le Breton pour leurs discussions sur mes travaux. Je remercie très vivement Yury Kutoyants, Rosenbaum Mathieu et Dachian sergueï qui ont accepté de faire partie du jury lors de la soutenance de ma thèse.

Je voudrais remercier tous les membres de l'équipe 'Probabilité et Statistique' de l'université du Maine. Je garderai un excellent souvenir de la bonne ambiance qui règne au sein du laboratoire. Un grand merci à tous mes amis doctorants pour les bons moments que nous avons partagés. Je les remercie tous pour leurs convivialité et leurs aides.

Enfin, je remercie chalereusement toute ma famille: mon père, ma mère qui, magré la distance, m'ont soutenu, encouragé et poussé durant toutes mes années d'étude.

Résumé

Cette thèse porte sur l'analyse statistique de quelques modèles de processus stochastiques gouvernés par des bruits de type fractionnaire, en temps discret ou continu.

Dans le Chapitre 1, nous étudions le problème d'estimation par maximum de vraisemblance (EMV) des paramètres d'un processus autorégressif d'ordre p (AR(p)) dirigé par un bruit gaussien stationnaire, qui peut être à longue mémoire comme le bruit gaussien fractionnaire. Nous donnons une formule explicite pour l'EMV et nous analysons ses propriétés asymptotiques. En fait, dans notre modèle la fonction de covariance du bruit est supposée connue, mais le comportement asymptotique de l'estimateur (vitesse de convergence, information de Fisher) n'en dépend pas.

Le Chapitre 2 est consacré à la détermination de l'entrée optimale (d'un point de vue asymptotique) pour l'estimation du paramètre de dérive dans un processus d'Ornstein-Uhlenbeck fractionnaire partiellement observé mais contrôlé. Nous exposons un principe de séparation qui nous permet d'atteindre cet objectif. Les propriétés asymptotiques de l'EMV sont démontrées en utilisant le programme d'Ibragimov-Khasminskii et le calcul de transformées de Laplace d'une fonctionnelle quadratique du processus.

Dans le Chapitre 3, nous présentons une nouvelle approche pour étudier les propriétés du mouvement brownien fractionnaire mélangé et de modèles connexes, basée sur la théorie du filtrage des processus gaussiens. Les résultats mettent en lumière la structure de semimartingale et mènent à un certain nombre de propriétés d'absolue continuité utiles. Nous établissons l'équivalence des mesures induites par le mouvement brownien fractionnaire mélangé avec une dérive stochastique, et en déduisons l'expression correspondante de la dérivée de Radon-Nikodym. Pour un indice de Hurst $H > 3/4$, nous obtenons une représentation du mouvement brownien fractionnaire mélangé comme processus de type diffusion dans sa filtration naturelle et en déduisons une formule de la dérivée de Radon-Nikodym par rapport à la mesure de Wiener. Pour $H < 1/4$, nous montrons l'équivalence da la mesure avec celle la composante fractionnaire et obtenons une formule pour la densité correspondante. Un domaine d'application potentielle est l'analyse statistique des modèles gouvernés par des bruits fractionnaires mélangés. A titre d'exemple, nous considérons le modele de régression linéaire de base et montrons comment definir l'EMV et étudié son comportement asymptotique.

Abstract

This thesis focuses on the statistical analysis of some models of stochastic processes generated by fractional noise in discrete or continuous time.

In Chapter 1, we study the problem of parameter estimation by maximum likelihood (MLE) for an autoregressive process of order p (AR (p)) generated by a stationary Gaussian noise, which can have long memory as the fractional Gaussian noise. We exhibit an explicit formula for the MLE and we analyze its asymptotic properties. Actually in our model the covariance function of the noise is assumed to be known but the asymptotic behavior of the estimator (rate of convergence, Fisher information) does not depend on it.

Chapter 2 is devoted to the determination of the asymptotical optimal input for the estimation of the drift parameter in a partially observed but controlled fractional Ornstein-Uhlenbeck process. We expose a separation principle that allows us to reach this goal. Large sample asymptotical properties of the MLE are deduced using the Ibragimov-Khasminskii program and Laplace transform computations for quadratic functionals of the process.

In Chapter 3, we present a new approach to study the properties of mixed fractional Brownian motion (fBm) and related models, based on the filtering theory of Gaussian processes. The results shed light on the semimartingale structure and properties lead to a number of useful absolute continuity relations. We establish equivalence of the measures, induced by the mixed fBm with stochastic drifts, and derive the corresponding expression for the Radon-Nikodym derivative. For the Hurst index $H > 3/4$ we obtain a representation of the mixed fBm as a diffusion type process in its own filtration and derive a formula for the Radon-Nikodym derivative with respect to the Wiener measure. For $H < 1/4$, we prove equivalence to the fractional component and obtain a formula for the corresponding derivative. An area of potential applications is statistical analysis of models, driven by mixed fractional noises. As an example we consider only the basic linear regression setting and show how the MLE can be defined and studied in the large sample asymptotic regime.

Contents

Introduction

Cette thèse est consacrée à l'analyse statistique de quelques modèles de processus stochastiques gouvernés par des bruits de type fractionnaire, en temps discret ou continu. De tels modèles permettent de rendre compte de phénomènes de longue mémoire et d'autosimilarité qui ont été observés dans de nombreux champs d'applications : hydrologie [43], météorologie, économie [28], finance mathématique, géophysique et biologie [52].

En **temps discret**, l'analyse des modèles linéaires ou non linéaires gouvernés par des bruits blancs a été abondamment développée dans la littérature. Ainsi, le problème de l'estimation paramétrique de modèles autorégressifs générés par des bruits blancs a été particulièrement étudié pendant des décennies. De nombreuses propriétés asymptotiques (distribution, biais, erreur quadratique) de l'estimateur de maximum de vraisemblance (EMV) ont été exhibées pour les modèles autorégressifs d'ordre 1 (AR(1)), ceci dans tous les cas possibles : stable, instable et explosif (voir, par exemple, [3, 15, 63, 66, 77, 78]). Concernant les modèles autorégressifs d'ordre p (AR(p)) avec des bruits blancs, les résultats sur le comportement asymptotique de l'EMV sont moins exhaustifs même s'il y a encore de nombreuses contributions (voir, par exemple, [3, 20, 42, 44, 51, 62]).

Au cours des trente dernières années, de nombreux articles ont été consacrés à l'analyse statistique des processus AR qui peuvent représenter des phénomènes de mémoire longue. Bien sûr, les modèles pertinents mettent en jeu des structures plus ou moins spécifiques de dépendance dans les perturbations. Il y a plusieurs articles consacrés au problème de l' estimation des paramètres du bruit gaussien fractionnaire et des modèles voisins (voir [2, 23, 27, 30, 68, 81] pour des contributions et les références qui s'y trouvent). Il est à noter que dans les modèles autorégressifs stationnaires perturbés par des bruits fortement dépendants, l'estimateur des moindres carrés n'est généralement pas consistant.

Pour autant que nous sachions, il n'y a pas de contribution à l'estimation par maximum de vraisemblance des coefficients d'un processus AR(p) avec des bruits gaussions stationnaires quelconques, en particulier avec des bruits gaussiens fractionnaires. Même si les conditions générales dans lesquelles l'EMV est consistant et asymptotiquement normal ont été données dans [71], il serait nécessaire, pour appliquer ce résultat, d'étudier les dérivées secondes de la matrice de covariance de l'échantillon d'observation. Pour éviter cette difficulté, certains auteurs ont suivi une autre approche, suggérée par Whittle, [23] qui s'applique pour les séries stationnaires. Mais, même pour un $AR(1)$, dans le cas explosif, il n'est déj à plus possible d'appliquer les théorèmes de [23] et de déduire les propriétés de l'estimateur.

Dans le **Chapitre 1**, nous étudions le problème d'estimation (par maximum de vraisemblance) des paramètres d'un processus autorégressif d'ordre p (AR(p)) dirigé par un bruit gaussien stationnaire, qui peut être à longue mémoire comme le bruit gaussien fractionnaire.

Nous donnons une formule explicite pour l'EMV et nous analysons ses propriétés asymptotiques. En fait, dans notre modèle la fonction de covariance des perturbations est supposée connue, mais le comportement asymptotique de l'estimateur de coefficient (vitesse de convergence, information de Fisher) n'en dépend pas.

On considère le processus $(X_n,\, n \geq 1)$ définit par

$$X_n = \sum_{i=1}^{p} \vartheta_i X_{n-i} + \xi_n, \quad n \geq 1, \quad X_r = 0, \quad r = 0,\, -1, \dots,\, -(p-1),$$

où $\xi = (\xi_n,\, n \in \mathbb{Z})$ est une suite de variables gaussiennes centrées, stationnaire et régulière, *i.e.*

$$\int_{-\pi}^{\pi} |\ln f_\xi(\lambda)|\, d\lambda < \infty,$$

où $f_\xi(\lambda)$ est la densité spectrale de ξ.

Nous supposons que la covariance $c = (c(m, n),\, m, n \geq 1)$, où

$$\mathbf{E}\xi_m \xi_n = c(m,\, n) = \rho(|n - m|), \quad \rho(0) = 1, \tag{1}$$

est définie positive. Pour une valeur fixe du paramètre $\vartheta = (\vartheta_1,\, \dots,\, \vartheta_p) \in \mathbb{R}^p$, soit $\mathbf{P}_\vartheta^N$ la mesure de probabilité induite par $X^{(N)}$. Soit $\mathcal{L}(\vartheta,\, X^{(N)})$ la fonction de vraisemblance définie par la dérivée de Radon-Nikodym de $\mathbf{P}_\vartheta^N$ par rapport à la mesure de Lebesgue. Notre objectif est d'étudier les propriétés asymptotiques du MLE $\widehat{\vartheta}_n$ de ϑ basé sur l'échantillon d'observation $X^{(n)} = (X_1, \dots, X_n)$ défini par

$$\hat{\vartheta}_N = \sup_{\vartheta \in \mathbb{R}^p} \mathcal{L}(\vartheta,\, X^{(N)}). \tag{2}$$

Dans un premier temps, pour preparer à l'analyse de la consistance (ou de la forte consistance) de $\hat{\vartheta}_N$ et preciser sa distribution limite nous transformons notre modèle d'observation dans un modèle "équivalent" avec des bruits gaussiens indépendants. Cela permet d'écrire explicitement l'EMV et la différence entre $\hat{\vartheta}_N$ et la valeur réelle ϑ apparait comme le produit d'une martingale par l'inverse de son processus croissant. Ensuite, nous pouvons utiliser des calculs de transformées de Laplace pour prouver les propriétés asymptotiques de l'EMV

On note A_0 la matrice de taille $p \times p$ définie par:

$$A_0 = \begin{pmatrix} \vartheta_1 & \vartheta_2 & \cdots & \vartheta_{p-1} & \vartheta_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

Soit $r(\vartheta)$ le rayon spectral de A_0. Définissons le domaine Θ des valeurs du paramètre $\vartheta = (\vartheta_1,\, \dots,\, \vartheta_p) \in \mathbb{R}^p$ par:

$$\Theta = \left\{\vartheta \in \mathbb{R}^p \,|\, r(\vartheta) < 1\right\}.$$

Nous montrons que l'EMV $\hat{\vartheta}_N$ de $\vartheta = (\vartheta_1,\, \vartheta_2,\, \dots,\, \vartheta_p)$ est

- consistant, *i.e.*, pour tout $\vartheta \in \Theta$ et $\nu > 0$,

$$\lim_{N\to\infty} \mathbf{P}_\vartheta^N \left\{ \left\| \hat{\vartheta}_N - \vartheta \right\| > \nu \right\} = 0\,, \tag{3}$$

- asymptotiquement normal

$$\sqrt{N}\left(\hat{\vartheta}_N - \vartheta\right) \overset{law}{\Rightarrow} \mathcal{N}(\mathbf{0}, \mathcal{I}^{-1}(\vartheta)), \tag{4}$$

  où $\mathcal{I}(\vartheta)$ set la solution unique de l'équation de Lyapounov:

$$\mathcal{I}(\vartheta) = A_0 \mathcal{I}(\vartheta) A_0^* + bb^*. \tag{5}$$

  où b est le vecteur de $\mathbb{R}^p$:

$$b = \begin{pmatrix} 1 \\ \mathbf{0}_{(p-1)\times 1} \end{pmatrix}.$$

  Il est intéressant de souligner que la covariance asymptotique $\mathcal{I}^{-1}(\vartheta)$ est en fait la même que dans le cas standard où ξ est un bruit blanc.

- De plus les moments de $\hat{\vartheta}_N$ convergent, *i.e.* pour tout $\vartheta \in \Theta$ et $q > 0$

$$\lim_{N\to\infty} \left| \mathbf{E}_\vartheta \left\| \sqrt{N}\left(\hat{\vartheta}_N - \vartheta\right) \right\|^q - \mathbf{E}\left\|\eta\right\|^q \right| = 0, \tag{6}$$

  où $\| \ \|$ est la norme euclidienne de $\mathbb{R}^p$ et η est un vecteur gaussien centrée de matrice de covariance $\mathcal{I}(\vartheta)^{-1}$.

Dans le cas d'un processus autorégressif d'ordre 1 ($p = 1$), l'EMV $\hat{\vartheta}_N$ est même fortement consistant: pour tout $\vartheta \in \mathbb{R}$,

$$\lim_{N\to\infty} \hat{\vartheta}_N = \vartheta \quad p.s..$$

Le Chapitre 1 se termine par des résultats de simulations illustrant la convergence de l'estimateur. Les simulations sont faites pour un bruit gaussien fractionnaire (fGn), un bruit autorégressif (AR(1)) et un bruit moyenne mobile MA(1).

En **temps continu**, les modèles dirigés par le mouvement brownien ont été abondamment étudiés dans la littérature. Ainsi, les problèmes d'estimation paramétrique pour les processus de diffusion, éventuellement contrôlés ont été abordés dans le cas d'observation complète et d'observation partielle(voir, par exemple, [4, 26, 46, 47, 53, 53, 54, 57]). En particulier, pour un processus d'Ornstein-Uhlenbeck, l'étude statistique couvre l'ensemble de propriétés en horizon de temps fini et des propriétés asymptotiques de l'EMV du paramètre de dérive [41]. De même le cas d'un système contrôlé a été traité dans [57].

En vue de rendre compte de phénomènes de longue mémoire, le mouvement brownien fractionnaire a été à la base de la construction de modèles où il remplace le mouvement brownien ordinaire.

Le mouvement brownien fractionnaire a été introduit par Kolmogorov en 1940 sous le nom de spirale de Wiener pour modéliser la turbulence dans les fluides. Il obtient également sa représentation spectrale. En 1968, Mandelbrot et Van Ness propose une représentation sous la forme de l'intégrale d'un noyau déterministe par rapport à un "mouvement brownien standard bilatéral " et lui donnent son nom actuel (pour plus d'information sur le sujet, voir [50]). En 1969, Molchan et Golosov construisent un mouvement brownien fractionnaire comme intégrale de Wiener d'un noyau plus complexe par rapport à un mouvement brownien standard. Cette dernière représentation est à l'origine de nombreux développements théoriques impliquant le mouvement brownien fractionnaire, notamment parce que les filtrations naturelles du processus de Wiener en jeu et du mouvement brownien fractionnaire coincident.

Le mouvement brownien fractionnaire est un processus aléatoire auto-similaire, sa loi de probabilité demeurant invariante par un changement d'échelle temporelle particulier:

$$(B^H_{at})_{t\in R} \overset{d}{=} a^H B^H_t,\ t \in R,\ a > 0.$$

C'est un processus aux accroissements stationnaires.

L'indice de Hurst H caractérise également la structure de dépendance et la mémoire du processus : les accroissements sur des intervalles disjoints sont corrélés positivement si $H > 1/2$ et négativement si $H < 1/2$. De plus, la décroissance de cette corrélation lorsque ces intervalles s'éloignent est lente pour $H > 1/2$ (longue mémoire) et rapide pour $H < 1/2$ (mémoire courte).

Pour $H = 1/2$, le mouvement brownien fractionnaire est le mouvement brownian standard, les accroissements sont alors indépendants et le processus est sans mémoire.

Dans le cas du mouvement brownien fractionnaire, l'indice de Hurst est également une mesure de la régularité des trajectoires. Plus l'indice de Hurst est grand, plus la trajectoire est régulière et inversement. Le mouvement brownien fractionnaire est l'unique processus gaussien qui à la fois autosimilaire et à longue mémoire. Il est utile dans la suite de préciser que pour $H \neq 1/2$, le mbf n'est pas une semimartingale et le calcul d'Itô usuel n'est pas utilisable. Utiliser ces processus devient donc un challenge intéressant même pour des développements théoriques.

Dans les **Chapitres 2 et 3** nous étudions deux modèles spécifiques: l'un est dirigé par un mouvement brownien fractionnaire et l'autre par le mélange d'un mouvement brownien ordinaire et d'un mouvement brownien fractionnaire.

Dans le **Chapitre 2** nous considérons le problème d'estimation du paramètre de dérive d'un processus d'Ornstein-Uhlenbeck fractionnaire avec contrôle.

Soit $X = (X_t,\, t \geq 0)$ et $Y = (Y_t,\, t \geq 0)$ le processus signal et le processus d'observation respectivement. Nous traitons le cas d'une observation complete du signal

$$dY_t = dX_t = -\vartheta X_t dt + u(t)dt + dV^H_t, \quad t > 0, \tag{7}$$

et le cas d'une observation partielle linéaire du signal dans un bruit additif

$$\begin{cases} dX_t &= -\vartheta X_t dt + u(t)dt + dV^H_t, \\ dY_t &= \mu X_t dt + dW^H_t, \end{cases} \quad t > 0, \tag{8}$$

où $V^H = (V^H_t,\, t \geq 0)$ et $W^H = (W^H_t, t \geq 0)$ sont deux mouvements browniens fractionnaires indépendants, $u = (u(t),\, t \geq 0)$ est une fonction déterministe et les conditions initiales sont fixées, $X_0 = Y_0 = 0$.

Supposons que le paramètre $\vartheta > 0$ est inconnue et doit être estimée compte tenu de la trajectoire observée $Y^T = (Y_t,\, 0 \le t \le T)$. Soit $\mathcal{L}(\vartheta, Y^T)$ la fonction de vraisemblance et nous définissons:

$$\mathcal{J}_T(\vartheta) = \sup_{u \in \mathcal{U}_T} \mathcal{I}_T(\vartheta, u),$$

où l'information de Fisher est:

$$\mathcal{I}_T(\vartheta, u) = -\mathbf{E}_\vartheta \frac{\partial^2}{\partial \vartheta^2} \ln \mathcal{L}(\vartheta, Y^T)$$

et $\mathcal{U}_T$ un espace fonctionnel de contrôle défini par (2.12) et (2.13) page 37.

Notre objectif principal est de trouver l'estimateur $\overline{\vartheta}_T$ du paramètre ϑ qui soit asymptotiquement efficace dans le sens où, pour tout compact

$$\sup_{\vartheta \in \mathbb{K}} \mathcal{J}_T(\vartheta) \mathbf{E}_\vartheta \left(\overline{\vartheta}_T - \vartheta\right)^2 = 1 + o(1)\,, \tag{9}$$

lorsque $T \to \infty$.

Le problème consiste à trouver un contrôle optimal $u_{opt}(t)$ qui maximise l'information de Fisher des équations (7) et (8), puis de déduire les propriétés asymptotiques de l'EMV de $\vartheta > 0$ dans les équations avec contrôle optimal. En adaptant la méthode développée dans [12], on obtient le contrôle optimal

$$u_{opt}(t) = \frac{\kappa}{\sqrt{2\lambda}} t^{|H - \frac{1}{2}|}$$

où

$$\kappa = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(\frac{1}{2} + H\right) \quad \text{et} \quad \lambda = \frac{H\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{2(1-H)\Gamma(\frac{3}{2}-H)}.$$

Comme l'entrée optimale ne dépend pas de ϑ, un candidat possible est le EMV. Nous prouvons que le EMV est efficace au sens de (2.3) et on en déduit ses propriétés asymptotiques:

- consistent uniformément sur les compacts $K \subset \mathbb{R}_*^+$, *i.e.* pour tout $\nu > 0$,

$$\lim_{T \to \infty} \sup_{\vartheta \in \mathbb{K}} \mathbf{P}_\vartheta^T \left\{ \left| \hat{\vartheta}_T - \vartheta \right| > \nu \right\} = 0\,;$$

- asymptotiquement normal (uniformément sur K), *i.e.* pour $T \to \infty$

$$\lim_{T \to \infty} \sup_{\vartheta \in \mathbb{K}} \left| \mathbf{E}_\vartheta f\left(\sqrt{T}\left(\hat{\vartheta}_T - \vartheta\right)\right) - \mathbf{E} f(\xi) \right| = 0 \quad \forall f \in \mathcal{C}_b$$

où $\xi \sim \mathcal{N}(0, \mathcal{I}^{-1}(\vartheta))$;

- convergence des moments (uniformément sur K), *i.e.* pour $p > 0$,

$$\lim_{T \to \infty} \sup_{\vartheta \in \mathbb{K}} \left| \mathbf{E}_\vartheta \left| \sqrt{T}\left(\hat{\vartheta}_T - \vartheta\right) \right|^p - \mathbf{E} \left| \xi \right|^p \right| = 0.$$

avec une information de Fisher explicite donnée par

$$\lim_{T\to+\infty} \frac{\mathcal{J}_T(\vartheta)}{T} = \mathcal{I}(\vartheta),$$

$$\mathcal{I}(\vartheta) = \begin{cases} \dfrac{1}{2\vartheta} + \dfrac{1}{\vartheta^2} & \text{(cas d'observation directe)} \\ \dfrac{1}{2\vartheta} - \dfrac{2\vartheta}{\alpha(\alpha+\vartheta)} + \dfrac{\vartheta^2}{2\alpha^3} + \dfrac{\mu^2}{\alpha^2\vartheta^2} & \text{(cas d'observation partielle)} \end{cases}$$

et $\alpha = \sqrt{\mu^2 + \vartheta^2}$.

La vérification des conditions d'application du programme de Ibragimov-Khasminskii [31, Théorème I.10.1] est basée sur le calcul de la transformée de Laplace du terme quadratique du rapport de vraisemblance.

Dans le **Chapitre 3** nous étudions un modèle dirigé par le mélange d'un mouvement brownien ordinaire et d'un mouvement brownien fractionnaire, c'est-à-dire

$$X_t = B_t + B_t^H,\ t \in [0,T], \quad T > 0, \tag{10}$$

où $B = (B_t, t \geq 0)$ est un mouvement brownian et $B^H = (B_t^H, t \geq 0)$ est un mouvement brownian fractionnaire de l'indice de Hurst $H \in (0,1)$ indépendant de B.

L' intérêt pour le processus (10) a été déclenché par l 'article de P.Cheridito [16], où l'auteur a découvert un curieux changement dans les propriétés de X apparaissant pour $H = \frac{3}{4}$. Il s'avère que X est une semimartingale dans sa propre filtration si et seulement si $H = \frac{1}{2}$ ou $H \in \left(\frac{3}{4}, 1\right]$ et, en outre, dans ce dernier cas, la mesure de probabilité μ^X, induite par X sur l'espace mesurable des fonctions continues $C([0,T])$ est équivalente à la mesure de Wiener μ^W. Comme le processus B^H n'est pas lui-même une semimartingale, à moins que $H = \frac{1}{2}$ ou $H = 1$, cette affirmation signifie que B^H peut être régularisé en une semimartingale par addition d'une perturbation brownienne indépendante. Dans [16] ce fait est examiné en mathématique financière pour l'évaluation des options de et des opportunités d'arbitrage sur les marchés (voir aussi [17]). Une revue exhaustive des dévelopements de cette thématique liés à la finance peut être trouvée dans [7] . En plus d'être d'intérêt pour la communauté financière, le résultat dans [16] a également conduit à un certain nombre de généralisations élégantes et de preuves alternatives (voir, par exemple, [6, 74, 75]).

Nous présentons une nouvelle approche pour étudier les propriétés du mouvement brownien fractionnaire mélangé et des modèles connexes, basée sur la théorie du filtrage des processus gaussiens. Dans ce chapitre, nous procédons à l'analyse stochastique du processus X avec un $H \in (0, 1]$ et de modèles plus généraux additifs, gouvernés par X. Pour $H > \frac{3}{4}$ et $H < \frac{1}{4}$, nous obtenons des représentations de X comme processus de diffusion et de type de diffusion fractionnaires respectivement, et dérivons les formules correspondantes pour les dérivées de Radon-Nikodym par rapport aux mesures de Wiener standard et fractionnaire (Théorème 3.3). En particulier , cela suggère une nouvelle preuve directe du théorème de la régularisation déj à mentionné de Cheridito [16] et de sa généralisation par H. van Zanten [75].

Nous insistons sur le rôle de la *martingale fondamentale*, qui engendre la même filtration que X et par rapport à laquelle X peut être représenté comme une intégrale stochastique et vice versa. Pour le mouvement brownien fractionnaire mélangé avec dérive cette notion se généralise naturellement à la *semimartingale fondamentale* et conduit à un changement de mesure de type de Girsanov.

Un domaine d'applications potentielles est l'analyse statistique des modèles gouvernés par des bruits fractionnaires mélangés. A titre d'example nous considérons le cas de la régression linéaire de base et montrons comment l'estimateur du maximum de vraisemblance (EMV) peut être défini et étudié dans le régime asymptotique.

Nous présentons d'abord une *analyse stochastique du mélange d'un mouvement brownien et d'un mouvement brownien fractionnaire.*

Soient $X = (X_t, 0 \le t \le T)$ le processus de mélange, $\mathcal{F}^X = (\mathcal{F}_t^X, 0 \le t \le T)$ et $\mathcal{F} = (\mathcal{F}_t, 0 \le t \le)$, les filtrations engendrées par X et $(B,\ B^H)$ respectivement.

On considère la $\mathcal{F}^X-$martingale $M = (M_t, 0 \le t \le T)$ définie par

$$M_t = \mathbf{E}\big(B_t | \mathcal{F}_t^X\big), \quad t \in [0, T].$$

Remarquablement, M encode la plupart des caractéristiques essentielles du processus X, ce qui rend sa structure particulier transparente.

M admet la représentation

$$M_t = \int_0^t g(s,t) dX_s, \quad \langle M \rangle_t = \int_0^t g(s,t) ds,$$

où le noyau $g(s,t)$ satisfait l'équation intégro-différentielle :

$$g(s,t) + H \frac{d}{ds} \int_0^t g(r,t) |s-r|^{2H-1} \mathrm{sign}(s-r) dr = 1, \quad 0 < s < t \le T. \tag{11}$$

L'équation (11) peut être réécrite comme une équation intégrale avec un noyau faiblement singulier, dont la formule précise est déterminé par la valeur de H.

Nous montrons, dans un premier temps, que le processus X admet la représentation

$$X_t = \int_0^t G(s,t) dM_s, \quad t \in [0, T],$$

où

$$G(s,t) := 1 - \frac{d}{d\langle M \rangle_s} \int_0^t g(\tau, s) d\tau, \quad 0 \le s \le t \le T. \tag{12}$$

De plus les filtrations naturelles de X et M coïncident.

Dans un second temps, nous montrons que X *est un processus de diffusion* pour $H \in (\frac{3}{4}, 1)$, solution de l'équation différentielle stochastique

$$X_t = W_t - \int_0^t \varphi_s(X) ds, \quad t \in [0, T],$$

où

$$W_t = \int_0^t \frac{1}{g(s,s)} dM_s$$

est un $\mathcal{F}^X$–mouvement brownien et $\varphi_t(X) = \int_0^t R(s,t) dX_s$,

$$R(s,t) := \frac{\dot{g}(s,t)}{g(t,t)}, \quad \dot{g}(s,t) := \frac{\partial}{\partial t} g(s,t), \quad s \ne t.$$

De plus, les mesures μ^X et μ^W sont équivalentes et

$$\frac{d\mu^X}{d\mu^W}(X) = \exp\left\{-\int_0^T \varphi_t(X)dX_t - \frac{1}{2}\int_0^T \varphi_t^2(X)dt\right\}.$$

L'analogue pour μ^X et μ^{B^H} quand $H < \frac{1}{4}$ est également formulée.

Nous poursuivrons par une *analyse stochastique du mélange d'un mouvement brownian et d'un mouvement brownian fractionnaire avec une dérive.*

Considérons un processus $Y = (Y_t)$ défini par

$$Y_t = \int_0^t f(s)ds + X_t, \quad t \in [0, T], \tag{13}$$

où $f = \big(f(t)\big)$ est un processus aux trajectoires continues et tel que $\mathbf{E}\int_0^T \big|f(t)\big|dt < \infty$, adapté à une filtration $\mathcal{G} = (\mathcal{G}_t)$, par rapport à laquelle M est une martingale.

Alors Y admet la représentation:

$$Y_t = \int_0^t G(s,t)dZ_s \tag{14}$$

avec G, définie dans (12), où le processus $Z = (Z_t)$

$$Z_t = \int_0^t g(s,t)dY_s, \quad t \in [0, T]$$

est une $\mathcal{G}$-semimartingale dont la décomposition de Doob - Meyer est

$$Z_t = M_t + \int_0^t \Phi(s)d\langle M\rangle_s, \tag{15}$$

où

$$\Phi(t) = \frac{d}{d\langle M\rangle_t}\int_0^t g(s,t)f(s)ds. \tag{16}$$

En particulier, $\mathcal{F}_t^Y = \mathcal{F}_t^Z$, P-p.s. pour tout $t \in [0, T]$ et, si

$$\mathbf{E}\exp\left\{-\int_0^T \Phi(t)dM_t - \frac{1}{2}\int_0^T \Phi^2(t)d\langle M\rangle_t\right\} = 1,$$

les mesures μ^X et μ^Y sont équivalentes et la densité de Radon-Nikodym correspondante est donnée par

$$\frac{d\mu^Y}{d\mu^X}(Y) = \exp\left\{\int_0^T \hat{\Phi}(t)dZ_t - \frac{1}{2}\int_0^T \hat{\Phi}^2(t)d\langle M\rangle_t\right\}, \tag{17}$$

où $\hat{\Phi}(t) = \mathbf{E}\big(\Phi(t)|\mathcal{F}_t^Y\big)$.

Nous complétons le chapitre par l'*étude du problème d'estimation du paramètre* de dérive ϑ lorsqu'on observe la trajectoire $(Y_t, 0 \le t \le T)$ dans le modèle de régression

$$Y_t = \vartheta t + B_t + B_t^H, \quad 0 \le t \le T. \tag{18}$$

Dans le modèle (18), l'EMV de ϑ est donné

$$\widehat{\vartheta}_T = \frac{\int_0^T g(s,T)dY_s}{\int_0^T g(s,T)ds}.$$

Pour $H \in (0,1)$, nous montrons que l'EMV $\widehat{\vartheta}_T$ est fortement consistant et que de plus:

- pour $H > \frac{1}{2}$,

$$\lim_{T\to\infty} T^{2-2H}\mathbf{E}(\widehat{\vartheta}_T - \vartheta)^2 = \frac{2H\Gamma(H+\frac{1}{2})\Gamma(3-2H)}{\Gamma(\frac{3}{2}-H)},$$

  où $\Gamma(\cdot)$ est la fonction Gamma.

- pour $H < \frac{1}{2}$

$$\lim_{T\to\infty} T\mathbf{E}\big(\widehat{\vartheta}_T - \vartheta\big)^2 = 1.$$

Chapter 1

Asymptotic properties of the MLE for the autoregressive process coefficients under stationary noise

1 Statement of the problem

1.1 Introduction

The problem of parametric estimation in classical autoregressive (AR) models generated by white noises has been studied for decades. In particular, for such autoregressive models of order 1 (AR(1)) consistency and many other asymptotic properties (distribution, bias, quadratic error) of the Maximum Likelihood Estimator (MLE) have been completely analyzed in all possible cases: stable, unstable and explosive (see, *e.g.*, [3, 15, 63, 66, 77, 78]). Concerning autoregressive models of order p (AR(p)) with white noises, the results about the asymptotic behavior of the MLE are less exhaustive but there are still many contributions in the literature (see, *e.g.*, [3, 20, 42, 44, 51, 62]).

In the past thirty years numerous papers have been devoted to the statistical analysis of AR processes which may represent long memory phenomenons as encountered in various fields as econometrics [28], hydrology [43] or biology [52]. Of course the relevant models exit from the white noise frame and they involve more or less specific structures of dependence in the perturbations. There are several papers devoted to the estimation problem of the parameters of fractional Gaussian noises and fractionally defferenced models (see, *e.g.*, [2, 23, 27, 30, 68, 81] for contributions and other references). It worth mentioning that in a stationary autoregressive models perturbed by strongly dependent noises the Least Square estimator is generally not consistent.

As far as we know, there is no contribution in the ML estimation of the coefficients of an AR(p) processes with depending noises, particularly with the fractional Gaussian noises. General conditions under which the MLE is consistent and asymptotically normal for stationary sequences have been given in [71]. In order to apply this result, it would be necessary to study the second derivatives of the covariance matrix of the observation sample $(X_1, \ldots, X_N)$. To avoid this difficulty, some authors followed an other approach suggested by Whittle [23] (which is not MLE) for stationary sequences. But even in autoregressive models of order 1 as soon as $|\vartheta| > 1$, the process is not stationary anymore and it is not possible to apply theorems in [23] to deduce estimator properties.

In this part, we deal with an AR(p) generated by an arbitrary regular stationary Gaussian noise. We exhibit an explicit formula for the MLE of the parameter and we analyze its asymptotic properties. Actually in our model the covariance function of the perturbation is know but the asymptotic behavior of the coefficient estimator (the rate of convergence, the Fisher information) does not depend on the structure of the noises covariance.

1.2 Statement of the problem

We consider an AR(p) process $(X_n,\, n \geq 1)$ defined by the recursion

$$X_n = \sum_{i=1}^{p} \vartheta_i X_{n-i} + \xi_n, \quad n \geq 1, \quad X_r = 0, \quad r = 0,\, -1, \ldots,\, -(p-1), \tag{1.1}$$

where $\xi = (\xi_n,\, n \in \mathbb{Z})$ is a centered regular stationary Gaussian sequence, *i.e.*

$$\int_{-\pi}^{\pi} |\ln f_\xi(\lambda)|\, d\lambda < \infty, \tag{1.2}$$

where $f_\xi(\lambda)$ is the spectral density of ξ. We suppose that the covariance $c = (c(m,n),\, m,n \geq 1)$, where

$$\mathbf{E}\xi_m\xi_n = c(m,\, n) = \rho(|n-m|), \quad \rho(0) = 1, \tag{1.3}$$

is positive defined.

For a fixed value of the parameter $\vartheta = (\vartheta_1,\, \ldots,\, \vartheta_p) \in \mathbb{R}^p$, let $\mathbf{P}_\vartheta^N$ denote the probability measure induced by $X^{(N)}$. Let $\mathcal{L}(\vartheta,\, X^{(N)})$ be the likelihood function defined by the Radon-Nikodym derivative of $\mathbf{P}_\vartheta^N$ with respect to the Lebesgue measure. Our goal is to study the large sample asymptotical properties of the Maximum Likelihood Estimator (MLE) $\widehat{\vartheta}_N$ of ϑ based on the observation sample $X^{(N)} = (X_1, \ldots, X_N)$:

$$\hat{\vartheta}_N = \sup_{\vartheta \in \mathbb{R}^p} \mathcal{L}(\vartheta,\, X^{(N)}). \tag{1.4}$$

At first, preparing for the analysis of the consistency (or strong consistency) of $\hat{\vartheta}_N$ and its limit distribution we transform our observation model into an "equivalent" model with independent Gaussian noises. This allows to write explicitly the MLE and actually, the difference between $\hat{\vartheta}_N$ and the real value ϑ appears as the product of a martingale by the inverse of its bracket process. Then we can use Laplace transforms computations to prove the asymptotical properties of the MLE.

The chapter is organized as follows. Section 2 contains theoretical results and simulations. Sections 3 and 4 are devoted to preliminaries and auxiliary results. The proofs of the main results are presented in Section 5.

2 Results and illustrations

2.1 Results

We define the $p \times p$ companion matrix A_0 and the vector $b \in \mathbb{R}^p$ as follows:

$$A_0 = \begin{pmatrix} \vartheta_1 & \vartheta_2 & \cdots & \vartheta_{p-1} & \vartheta_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \mathbf{0}_{(p-1)\times 1} \end{pmatrix}. \tag{1.5}$$

Let $r(\vartheta)$ be the spectral radius of A_0. The following results hold:

Theorem 1.1. *Let $p \geq 1$ and the parameter set be:*

$$\Theta = \{\vartheta \in \mathbb{R}^p \,|\, r(\vartheta) < 1\}. \tag{1.6}$$

The MLE $\hat{\vartheta}_N$ is consistent, i.e., for any $\vartheta \in \Theta$ and $\nu > 0$,

$$\lim_{N\to\infty} \mathbf{P}_\vartheta^N \left\{ \left\| \hat{\vartheta}_N - \vartheta \right\| > \nu \right\} = 0, \tag{1.7}$$

and asymptotically normal

$$\sqrt{N}\left(\hat{\vartheta}_N - \vartheta\right) \overset{law}{\Rightarrow} \mathcal{N}(\mathbf{0}, \mathcal{I}^{-1}(\vartheta)), \tag{1.8}$$

where $\mathcal{I}(\vartheta)$ is the unique solution of the Lyapounov equation:

$$\mathcal{I}(\vartheta) = A_0 \mathcal{I}(\vartheta) A_0^* + bb^*, \tag{1.9}$$

for A_0 and b defined in (1.5).

Moreover we have the convergence of the moments: for any $\vartheta \in \Theta$ and $q > 0$

$$\lim_{N\to\infty} \left| \mathbf{E}_\vartheta \left\| \sqrt{N}\left(\hat{\vartheta}_N - \vartheta\right) \right\|^q - \mathbf{E} \left\|\eta\right\|^q \right| = 0, \tag{1.10}$$

where $\| \, \|$ denotes the Euclidian norm on $\mathbb{R}^p$ and η is a zero mean Gaussian random vector with covariance matrix $\mathcal{I}(\vartheta)^{-1}$.

Remark 1. *It is worth to emphasize that the asymptotic covariance $\mathcal{I}^{-1}(\vartheta)$ is actually the same as in the standard case where (ξ_n) is a white noise.*

In the case $p = 1$ we can strengthen the assertions of Theorem 1.1. In particular, the strong consistency and uniform convergence on compacts of the moments hold.

Theorem 1.2. *Let $p = 1$ and the parameter set be $\Theta = \mathbb{R}$. The MLE $\hat{\vartheta}_N$ is strongly consistent, i.e. for any $\vartheta \in \Theta$*

$$\lim_{N\to\infty} \hat{\vartheta}_N = \vartheta \qquad a.s.. \tag{1.11}$$

Moreover, $\hat{\vartheta}_N$ is uniformly consistent and satisfies the uniform convergence of the moments on compacts $\mathbb{K} \subset (-1, 1)$, i.e. for any $\nu > 0$:

$$\lim_{N\to\infty} \sup_{\vartheta\in\mathbb{K}} \mathbf{P}_\vartheta^N \left\{ \left| \hat{\vartheta}_N - \vartheta \right| > \nu \right\} = 0, \tag{1.12}$$

and for any $q > 0$:,

$$\lim_{N\to\infty} \sup_{\vartheta\in\mathbb{K}} \left| \mathbf{E}_\vartheta \left| \sqrt{N} \left(\hat{\vartheta}_N - \vartheta \right) \right|^q - \mathbf{E} \left|\eta\right|^q \right| = 0, \tag{1.13}$$

where $\eta \sim \mathcal{N}(0,\, 1-\vartheta^2)$.

Remark 2. *It is worth mentioning that condition* (1.2) *can be rewritten in terms of the covariance function ρ : $\rho(n) \sim n^{-\alpha}$, $\alpha > 0$.*

2.2 Simulations

In this section we present for $p = 1$ three illustrations of the behavior of the MLE corresponding to noises which are MA(1), AR(1) and fGn.

Moving average noise MA(1) Here we consider MA(1) noises where

$$\xi_{n+1} = \frac{1}{\sqrt{1+\alpha^2}}(\varepsilon_{n+1} + \alpha\varepsilon_n), \quad n \geq 1,$$

where $(\varepsilon_n, n \geq 1)$ is a sequence of i.i.d. zero-mean standard Gaussian variables. Then the covariance function is given by

$$\rho(|n-m|) = \mathbb{1}_{\{|n-m|=0\}} + \frac{\alpha}{1+\alpha^2}\mathbb{1}_{\{|n-m|=1\}}.$$

Condition (1.2) is fulfilled for $|\alpha| < 1$.

Autoregressive noise (AR(1)) Here we consider stationary autoregressive AR(1) noises where

$$\xi_{n+1} = \sqrt{1-\alpha^2}\varepsilon_{n+1} + \alpha\xi_n, \quad n \geq 1,$$

where $(\varepsilon_n, n \geq 1)$ is a sequence of i.i.d. zero-mean standard Gaussian variables. Then the covariance function is

$$\rho(|n-m|) = \alpha^{|n-m|}.$$

Condition (1.2) is fulfilled for $|\alpha| < 1$.

Fractional Gaussian noise fGn Here the covariance function of (ξ_n) is

$$\rho(|m-n|) = \frac{1}{2}\left(|m-n+1|^{2H} - 2|m-n|^{2H} + |m-n-1|^{2H}\right),$$

for a known Hurst exponent $H \in (0,1)$. For simulation of the fGn we use Wood and Chan method (see [80]). The explicit formula for the spectral density of fGn sequence has been exhibited in [70]. Condition (1.2) is fulfilled for any $H \in (0,1)$.

On Figure 1.1 we can see that in conformity with Theorem 1.2, in the three cases the MLE is asymptotically normal with the same limiting variance as in the classical i.i.d. case.

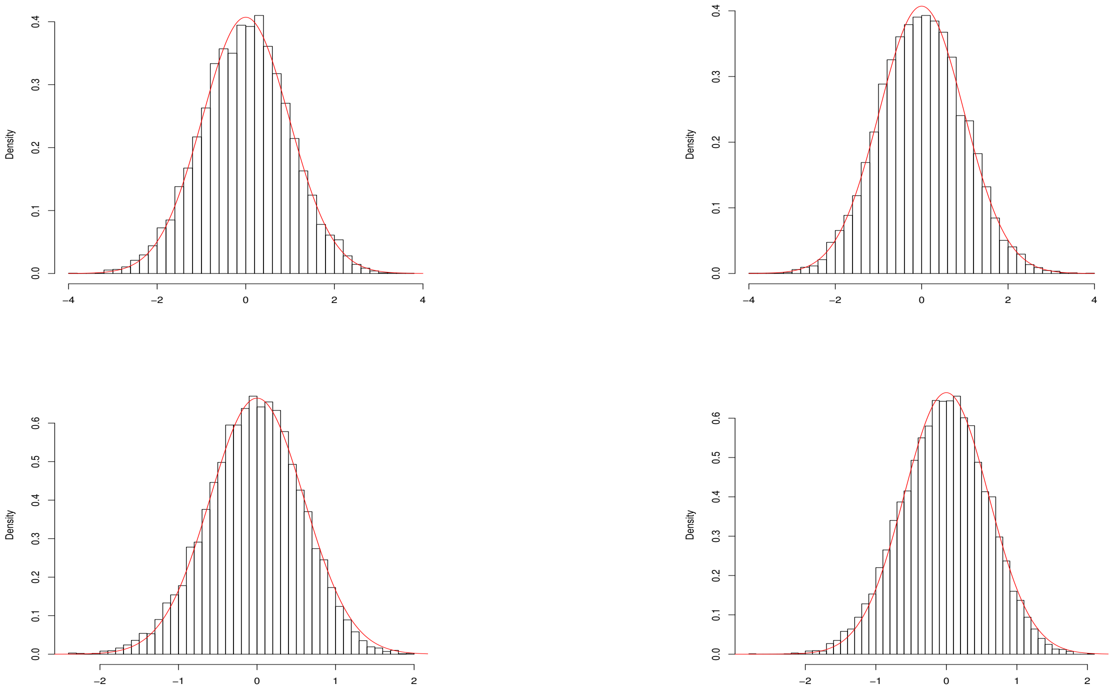

Figure 1.1: Asymptotical normality $N = 2000$ for the MLE in different cases by Monte-Carlo simulation of $M = 10000$ independent replications for AR(1) noises (top left) and MA noises (top right), both for $\alpha = 0.4$ and $\vartheta = 0.2$, and fGn noises for $H = 0.2$ (bottom left) and for $H = 0.8$ (bottom right) both for $\vartheta = 0.8$.

3 Preliminaries

3.1 Stationary Gaussian sequences

We begin with some well known properties of a stationary scalar Gaussian sequence $\xi = (\xi_n)_{n\geq 1}$. We denote by $(\sigma_n\varepsilon_n)_{n\geq 1}$ *the innovation type sequence* of ξ defined by

$$\sigma_1\varepsilon_1 = \xi_1, \quad \sigma_n\varepsilon_n = \xi_n - \mathbf{E}(\xi_n \,|\, \xi_1, \ldots, \xi_{n-1}), \quad n \geq 2,$$

where $\varepsilon_n \sim \mathcal{N}(0,1)$, $n \geq 1$ are independent. It follows from the Theorem of Normal Correlation ([49], Theorem 13.1) that there exists a deterministic kernel denoted by $k(n,m)$, $n \geq 1$, $m \leq n$, such that

$$\sigma_n\varepsilon_n = \sum_{m=1}^{n} k(n,m)\xi_m, \quad k(n,n) = 1. \tag{1.14}$$

In the sequel, for $n \geq 1$, we denote by β_{n-1} the partial correlation coefficient

$$-k(n,1) = \beta_{n-1},\ n \geq 1. \tag{1.15}$$

The following relations between $k(\cdot,\cdot)$, the covariance function $\rho(\cdot)$ defined by (1.3), the sequence of partial correlation coefficients $(\beta_n)_{n\geq 1}$ and the variances of innovations $(\sigma_n^2)_{n\geq 1}$ hold (see Levinson-Durbin algorithm [21])

$$\sigma_n^2 = \prod_{m=1}^{n-1}(1-\beta_m^2), \quad n \geq 2, \quad \sigma_1 = 1, \tag{1.16}$$

$$\sum_{m=1}^{n} k(n,\, m)\rho(m) = \beta_n \sigma_n^2, \tag{1.17}$$

$$k(n+1,\, n+1-m) = k(n,\, n-m) - \beta_n k(n,\, m). \tag{1.18}$$

Since we assume the positive definiteness of the covariance $c(\cdot,\, \cdot)$, there also exists an inverse deterministic kernel $K = (K(n,\, m),\, n \geq 1,\, m \leq n)$ such that

$$\xi_n = \sum_{m=1}^{n} K(n,m)\sigma_m \varepsilon_m. \tag{1.19}$$

Remark 3. *Actually, kernels k and K are nothing but the ingredients of the Choleski decomposition of covariance and inverse of covariance matrices. Namely,*

$$\Gamma_n^{-1} = k_n D_n^{-1} k_n^* \quad \textit{and} \quad \Gamma_n = K_n^* D_n K_n,$$

where $\Gamma_n = ((\rho(|i-j|)))$, k_n and K_n are $n \times n$ lower triangular matrices with ones as diagonal entries and $k(i,j)$ and $K(i,j)$ as subdiagonal entries respectively and D_n is an $n \times n$ diagonal matrix with σ_i^2 as diagonal entries. Here $*$ denotes the transposition.

Remark 4. *It is worth mentioning that condition* (1.2) *implies that*

$$\sum_{n\geq 1} \beta_n^2 < \infty. \tag{1.20}$$

Indeed, for every regular stationary Gaussian sequence $\xi = (\xi_n,\, n \in \mathbb{Z})$, there exists a sequence of i.i.d $\mathcal{N}(0,1)$ random variables $(\widetilde{\varepsilon}_n,\, n \in \mathbb{Z})$ and a sequence of real numbers a_k, $k \geq 0$ with $a_0 \neq 0$ such that:

$$\xi_n = \sum_{k=0}^{\infty} a_k \widetilde{\varepsilon}_{n-k},$$

and for all $n \in \mathbb{Z}$ the σ-algebra generated by $(\xi_k)_{-\infty<k\leq n}$ coincides with the σ-algebra generated by $(\widetilde{\varepsilon}_k)_{-\infty<k\leq n}$.

Note that the variance σ_n^2 of the innovations is also the one step predicting error and the following equalities hold thanks to the stationarity of ξ:

$$\lim_{n\to\infty} \prod_{m=1}^{n-1}(1-\beta_m^2) = \lim_{n\to\infty} \sigma_n^2$$

$$= \lim_{n\to\infty} \mathbf{E}\left(\xi_n - \mathbf{E}(\xi_n|\xi_1, \cdots \xi_{n-1})\right)^2 = \lim_{n\to\infty} \mathbf{E}\left(\xi_0 - \mathbf{E}(\xi_0|\xi_{-1}, \cdots \xi_{-n+1})\right)^2$$

$$= \mathbf{E}\left(\xi_0 - \mathbf{E}(\xi_0|\xi_s, s \leq -1)\right)^2 = \mathbf{E}\left(\xi_0 - \mathbf{E}(\xi_0|\varepsilon_s, s \leq -1)\right)^2 = a_0^2 > 0$$

which implies (1.20).

3.2 Model Transformation

As usual, for the first step we extend the dimension of the observations in order to work with a first order autoregression in $\mathbb{R}^p$. Namely, let Y_n, $n \geq 1$, be $Y_n = (X_n, X_{n-1}, \ldots, X_{n-(p-1)})^*$ then $Y = (Y_n, n \geq 1)$ satisfies the first order autoregressive equation:

$$Y_n = A_0 Y_{n-1} + b\xi_n, \quad n \geq 1, \quad Y_0 = \mathbf{0}_{p\times 1}, \tag{1.21}$$

where A_0 and b are defined in (1.5). For the second step we take an appropriate linear transformation of Y in order to have i.i.d. noises in the corresponding observations. For this goal let us introduce the process $Z = (Z_n, n \geq 1)$ such that

$$Z_n = \sum_{m=1}^{n} k(n, m) Y_m, \quad n \geq 1, \tag{1.22}$$

where $k = (k(n, m), n \geq 1, m \leq n)$ is the kernel appearing in (1.14). Since we have also

$$Y_n = \sum_{m=1}^{n} K(n, m) Z_m, \tag{1.23}$$

where $K = (K(n, m), n \geq 1, m \leq n)$ is the inverse kernel of k (see (1.19)), the filtration of Z coincides with the filtration of Y (and also the filtration of X). Actually, it was shown in [13] that Z can be considered as the first component of a $2p$ dimensional $AR(1)$ process $\zeta = (\zeta_n, n \geq 1)$ governed by i.i.d. noises. More precisely, the process $\zeta = (\zeta_n, n \geq 1)$ defined by :

$$\zeta_n = \begin{pmatrix} Z_n \\ \sum_{r=1}^{n-1} \beta_r Z_r \end{pmatrix},$$

is a 2p-dimensional Markovian process which satisfies the following equation:

$$\zeta_n = \mathbf{A}_{n-1}\zeta_{n-1} + \ell\sigma_n\varepsilon_n, \quad n \geq 1, \quad \zeta_0 = \mathbf{0}_{2p\times 1}, \tag{1.24}$$

where

$$\mathbf{A}_n = \begin{pmatrix} A_0 & \beta_n A_0 \\ \beta_n \mathbf{Id}_{p\times p} & \mathbf{Id}_{p\times p} \end{pmatrix}, \quad \ell = \begin{pmatrix} 1 \\ \mathbf{0}_{(2p-1)\times 1} \end{pmatrix}, \tag{1.25}$$

and $(\varepsilon_n, n \geq 1)$ are i.i.d. zero mean standard Gaussian variables. Now the initial estimation problem is replaced by the problem of estimation of the unknown parameter ϑ from the observations of $\zeta = (\zeta_n, n \geq 1)$.

3.3 Maximum Likelihood Estimator

It follows directly from equation (1.24) that the log-likelihood function is nothing but:

$$\ln \mathcal{L}(\vartheta, X^{(N)}) = -\frac{1}{2}\sum_{n=1}^{N}\left(\frac{\ell^*(\zeta_n - \mathbf{A}_{n-1}\zeta_{n-1})}{\sigma_n}\right)^2 - \frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{n=1}^{N}\ln\sigma_n^2$$

and that the maximum likelihood estimator $\hat{\vartheta}_N$ is:

$$\hat{\vartheta}_N = \left(\sum_{n=1}^{N} \frac{a_{n-1}^* \zeta_{n-1} \zeta_{n-1}^* a_{n-1}}{\sigma_n^2} \right)^{-1} \cdot \left(\sum_{n=1}^{N} \frac{a_{n-1}^* \zeta_{n-1} \ell^* \zeta_n}{\sigma_n^2} \right). \tag{1.26}$$

Then we can write

$$\hat{\vartheta}_N - \vartheta = (\langle M \rangle_N)^{-1} \cdot M_N, \tag{1.27}$$

where

$$M_N = \sum_{n=1}^{N} \frac{a_{n-1}^* \zeta_{n-1}}{\sigma_n} \varepsilon_n\,, \qquad \langle M \rangle_N = \sum_{n=1}^{N} \frac{a_{n-1}^* \zeta_{n-1} \zeta_{n-1}^* a_{n-1}}{\sigma_n^2}, \tag{1.28}$$

with

$$a_n = \begin{pmatrix} \mathbf{Id}_{p\times p} \\ \beta_n \mathbf{Id}_{p\times p} \end{pmatrix}. \tag{1.29}$$

Note that $(M_n,\, n \geq 1)$ is a martingale and $(\langle M \rangle_n,\, n \geq 1)$ is its bracket process.

Remark 5. *It is worth mentioning that in the classical i.i.d. case, i.e., when* $\beta_n = 0,\, n \geq 1,\, M_N$ *and* $\langle M \rangle_N$ *in equations* (1.27)-(1.28) *reduce to:*

$$M_N = \sum_{n=1}^{N} Y_{n-1} \varepsilon_n\,, \qquad \langle M \rangle_N = \sum_{n=1}^{N} Y_{n-1} Y_{n-1}^*.$$

Of course, under the condition $r(\vartheta) < 1$ *due to the law of the large numbers and the central limit theorem for martingales the following convergences hold:*

$$\mathbf{P}_\vartheta - \lim_{N\to\infty} \frac{1}{N} \langle M \rangle_N = \mathcal{I}(\vartheta),\ \frac{1}{\sqrt{N}} M_N \overset{law}{\Rightarrow} \mathcal{N}\left(\mathbf{0},\, \mathcal{I}(\vartheta)\right), \tag{1.30}$$

where $\mathcal{I}(\vartheta)$ *is the unique solution of the Lyapounov equation* (1.9). *This implies immediately the consistency and the asymptotic normality of the MLE.*

4 Auxiliary results

Actually, the proof of Theorems 1.1- 1.2 is crucially based on the asymptotic study for N tending to infinity of the Laplace transform:

$$L_N^\vartheta(\mu) = \mathbf{E}_\vartheta \exp\left(-\frac{\mu}{2} \alpha^* \langle M \rangle_N \alpha \right), \tag{1.31}$$

for arbitrary $\alpha \in \mathbb{R}^p$ and a positive real number μ, where $\langle M \rangle_N$ is defined by (1.28). It can be rewritten as

$$L_N^\vartheta(\mu) = \mathbf{E}_\vartheta \exp\left(-\frac{\mu}{2} \sum_{n=1}^{N-1} \zeta_n^* \mathcal{M}_n \zeta_n \right), \tag{1.32}$$

where $\mathcal{M}_n = \dfrac{1}{\sigma_{n+1}^2} a_n \alpha \alpha^* a_n^*$, a_n is defined by (1.29) and ζ satisfies the equation (1.24). In the sequel we will suppose that all the eigenvalues of A_0 are simple and different from 0. Actually, it is not a real restriction, since the general case can be studied by using small perturbations arguments.

Lemma 1.1. *The Laplace transform $L_N^\vartheta(\mu)$ can be written explicitly in the following form:*

$$L_N^\vartheta(\mu) = \left(\left(\prod_{n=1}^{N-1} \det \mathbf{A}_n\right) \det \Psi_N^1\right)^{-\frac{1}{2}}, \tag{1.33}$$

where $\mathbf{A}_n$ is defined by equation (1.25) *and*

$$\sigma_N^2 \Psi_N^1 = \Psi_0 \mathbf{J} \prod_{n=1}^{N-1} (\mathcal{A}_\mu \otimes A_1^n + \mathbf{Id}_{2p\times 2p} \otimes A_2^n) \mathbf{J}^* \Psi_0^*. \tag{1.34}$$

Here $\otimes$ is the Kronecker product, $\Psi_0 = (\mathbf{Id}_{2p\times 2p} \quad \mathbf{0}_{2p\times 2p})$,

$$\mathcal{A}_\mu = \begin{pmatrix} A_0^{-1} & A_0^{-1} bb^* \\ \mu\alpha\alpha^* & A_0^* + \mu\alpha\alpha^* A_0^{-1} bb^* \end{pmatrix} \tag{1.35}$$

and 2×2 matrices A_1^n, A_2^n are defined by

$$A_1^n = \begin{pmatrix} 1 & 0 \\ -\beta_n & 0 \end{pmatrix}, \qquad A_2^n = \begin{pmatrix} 0 & -\beta_n \\ 0 & 1 \end{pmatrix}. \tag{1.36}$$

Proof. With the Theorem 1 in Appendix and the property $\mathbf{E}\zeta_n = \mathbf{0}$, we know that

$$L_N^\vartheta(\mu) = \prod_{n=1}^{N-1} \det (\mathbf{Id} + \mu\gamma(n)\mathcal{M}_n)^{-\frac{1}{2}}$$

where $(\gamma(n), n \geq 1)$ is the one step prediction error for the observation

$$Y_n = \mu\mathcal{M}_n\zeta_n + \sqrt{\mu}\mathcal{M}_n^{\frac{1}{2}}\tilde{\varepsilon}_n, \quad n \geq 1.$$

It is known that this error follows a Ricatti equation:

$$\gamma(n) = \mathbf{A}_{n-1}(\mathbf{Id} + \mu\gamma(n-1)\mathcal{M}_{n-1})^{-1}\gamma(n-1)\mathbf{A}_{n-1}^* + \sigma_n^2 \ell\ell^*$$

which can be linearized by $\gamma(n) = (\Psi_n^1)^{-1}\Psi_n^2$ where

$$\begin{cases} \Psi_n^1 = \Psi_{n+1}^1 \mathbf{A}_n - \mu\Psi_n^2\mathcal{M}_n, & n \geq 1, \\ \Psi_{n+1}^2 = \Psi_{n+1}^1 \sigma_{n+1}^2 \ell\ell^* + \Psi_n^2 \mathbf{A}_n^*, & n \geq 1, \end{cases}$$

with $\Psi_0^1 = \mathbf{Id}_{2p\times 2p}$ and $\Psi_0^2 = \mathbf{0}_{2p\times 2p})$.. Moreover, we have

$$\det (\mathbf{Id} + \mu\gamma(n)\mathcal{M}_n) = \frac{\det \Psi_{n+1}^1}{\det \Psi_n^1} \det \mathbf{A}_n.$$

Finally, we have

$$L_N^\vartheta(\mu) = \left(\left(\prod_{n=1}^{N-1} \det \mathbf{A}_n\right) \det \Psi_N^1\right)^{-\frac{1}{2}}.$$

We define $\Psi_n = (\Psi_n^1, \Psi_n^2)$, we can rewrite

$$\begin{cases} \Psi_n^1 = \Psi_{n-1}^1 \mathbf{A}_{n-1}^{-1} + \mu\Psi_{n-1}^2 \mathcal{M}_{n-1}\mathbf{A}_{n-1}^{-1}, & n \geq 1, \\ \Psi_n^2 = \Psi_{n-1}^1 \mathbf{A}_{n-1}^{-1} \ell\ell^*\sigma_n^2 + \Psi_{n-1}^2 \left(\mu\mathcal{M}_{n-1}\mathbf{A}_{n-1}^{-1}\ell\ell^*\sigma_n^2 + \mathbf{A}_{n-1}^*\right), & n \geq 1. \end{cases}$$

Now let us denote by $\widetilde{\Psi}_n^1 = \sigma_n^2 \Psi_n^1$ and $\widetilde{\Psi}_n^2 = \Psi_n^2 \begin{pmatrix} \mathbf{Id}_{p\times p} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & -\mathbf{Id}_{p\times p} \end{pmatrix}$. Then $\widetilde{\Psi}_n = (\widetilde{\Psi}_n^1 \quad \widetilde{\Psi}_n^2)$ satisfies for $n \geq 1$ the following equation

$$\widetilde{\Psi}_n = \widetilde{\Psi}_{n-1} \begin{pmatrix} A_0^{-1} & -\beta_{n-1}\mathbf{Id}_{p\times p} & A_0^{-1}bb^* & \mathbf{0}_{p\times p} \\ -\beta_{n-1}A_0^{-1} & \mathbf{Id}_{p\times p} & -\beta_{n-1}A_0^{-1}bb^* & \mathbf{0}_{p\times p} \\ \mu\alpha\alpha^* A_0^{-1} & \mathbf{0}_{p\times p} & \mu\alpha\alpha^* A_0^{-1}bb^* + A_0^* & -\beta_{n-1}\mathbf{Id}_{p\times p} \\ -\beta_{n-1}(\mu\alpha\alpha^* A_0^{-1}) & \mathbf{0}_{p\times p} & -\beta_{n-1}(\mu\alpha\alpha^* A_0^{-1}bb^* + A_0^*) & \mathbf{Id}_{p\times p} \end{pmatrix}.$$

Let π be the following permutation of $\{1, \cdots, 4p\}$:

$$\pi(i) = \begin{cases} k+1, & i = 2k+1 \\ p+r, & i = 2r \\ 2p+k+1, & i = 2p+2k+1 \\ 3p+r, & i = 2r+2p \end{cases} \tag{1.37}$$

where $k = 0, \cdots, (p-1)$ and $r = 1, \cdots, p$. Denote by $\mathbf{J}$ the correspond permutation matrix

$$\mathbf{J}_{ij} = \delta_{i\,\pi(j)}, \quad i,j = 1, \cdots, 4p.$$

Then $\varphi_n = \widetilde{\Psi}_n \mathbf{J}$ satisfies the following equation:

$$\varphi_n = \varphi_{n-1} \left(\mathcal{A}_\mu \otimes A_1^{n-1} + \mathbf{Id}_{2p\times 2p} \otimes A_2^{n-1} \right), \tag{1.38}$$

which implies that

$$\varphi_{N} = \Psi_0 \mathbf{J} \prod_{n=1}^{N-1} (\mathcal{A}_\mu \otimes A_1^n + \mathbf{Id}_{2p\times 2p} \otimes A_2^n),$$

and consequently that $\sigma_N^2 \Psi_N^1$ satisfies equality (1.34).

□

Preparing for the asymptotic study we state the following result:

Lemma 1.2. *Let $(\beta_n)_{n\geq 1}$ be a sequence of real numbers satisfying the condition* (1.20). *For a fixed real number a let us define a sequence of 2×2 matrices $(S_N(a))_{N\geq 1}$ such that:*

$$S_N(a) = \prod_{n=1}^{N} \begin{pmatrix} a & -\beta_n \\ -a\beta_n & 1 \end{pmatrix} = \prod_{n=1}^{N} (aA_1^n + A_2^n), \tag{1.39}$$

where A_1^n and A_1^n are defined by equation (1.36). *Then*

1. *if $|a| < 1$, $\sup_{N\geq 1} \|S_N(a)\| < \infty$,*

2. *if $|a| > 1$, $\sup_{N\geq 1} \|(S_N(a))^{-1}\| < \infty$,*

3. *if a is sufficiently small, $\inf_{N\geq 1} trace((S_N^{-1}(\frac{1}{a}))S_N(a)) > 0$.*

Proof. The proof of assertions 1 and 2 follows directly from the estimates:

$$\|aA_1^n + A_2^n\| \leq 1 + \beta_n^2 \left(\frac{1+3a^2}{1-a^2}\right), \qquad \text{when} \quad |a| < 1,$$

$$\|(aA_1^n + A_2^n)^{-1}\| \leq 1 + \beta_n^2 \left(\frac{1+a^2}{a^2-1}\right), \qquad \text{when} \quad |a| > 1.$$

The proof of assertion 3 follows from the equality

$$G_N(a) = \frac{1}{1-\beta_N^2} \begin{pmatrix} a & -\beta_N \\ -a\beta_N & 1 \end{pmatrix} G_{N-1}(a) \begin{pmatrix} a & a\beta_N \\ \beta_N & 1 \end{pmatrix}$$

for $G_n(a) = S_n^{-1}(\frac{1}{a}))S_n(a)$. Hence $\mathrm{trace}(G_N(0)) = \frac{1}{\sigma_{N+1}^2}$ and condition (1.20) implies that $\lim_{N\to\infty} \sigma_N^2 = \prod_{n=1}^{\infty} \left(1+\beta_n^2\right) < \infty$ which achieves the proof. □

Actually, in the asymptotic study we work with a small value of μ. Note that for a small μ, matrix $\mathcal{A}_\mu$ defined by (1.35) can be represented as: $\mathcal{A}_\mu = \mathcal{A}_0 + \mu H$, where

$$\mathcal{A}_0 = \begin{pmatrix} A_0^{-1} & A_0^{-1}bb^* \\ \mathbf{0}_{p\times p} & A_0^* \end{pmatrix} \quad H = \begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \alpha\alpha^* & \alpha\alpha^* A_0^{-1}bb^* \end{pmatrix}. \tag{1.40}$$

Representation (1.40) implies that if the spectral radius $r(\vartheta) < 1$ then there are p eigenvalues of $\mathcal{A}_\mu$ such that $|\lambda_i(\mu)| > 1$ (in particular $\lambda_i(0)$, $i = 1, \cdots, p$ are the eigenvalues of A_0^{-1}) and p eigenvalues of $\mathcal{A}_\mu$ such that $|\lambda_j(\mu)| < 1$, $j = p+1, \cdots, 2p$.

Lemma 1.3. *Suppose that $r(\vartheta) < 1$. Let us take $\mu = \frac{1}{N}$ and denote by $\overline{L}_N^\vartheta(\mu)$:*

$$\overline{L}_N^\vartheta(\mu) = \prod_{i=1}^{p} \left(\frac{\lambda_i(\mu)}{\lambda_i(0)}\right)^{-\frac{N-1}{2}}. \tag{1.41}$$

Then, under condition (1.2),

$$\lim_{N\to\infty} \frac{L_N^\vartheta(\mu)}{\overline{L}_N^\vartheta(\mu)} = 1. \tag{1.42}$$

Proof. Thanks to the definition (1.25) of $\mathbf{A}_n$ the equality

$$\prod_{n=1}^{N-1} \det \mathbf{A}_n = \prod_{n=1}^{N-1} \left[(1-\beta_n^2)^p \frac{1}{\prod_{i=1}^p \lambda_i(0)}\right] = \frac{(\sigma_N^2)^p}{\prod_{i=1}^p \lambda_i(0)^{N-1}}$$

holds. Then due to equation (1.33) to prove (1.42) it is sufficient to check that

$$\lim_{N\to\infty} \frac{\det \sigma_N^2 \Psi_N^1}{(\sigma_N^2)^p \prod_{i=1}^p [\lambda_i(\mu)]^{N-1}} = 1. \tag{1.43}$$

Diagonalizing the matrix $\mathcal{A}_\mu$, *i.e.*, representing $\mathcal{A}_\mu$ as $\mathcal{A}_\mu = G_\mu D(\lambda_i(\mu)) G_\mu^{-1}$ with a diagonal matrix $D(\lambda_i(\mu))$, we have also

$$\mathcal{A}_\mu \otimes A_1^n + \mathbf{Id}_{2p\times 2p} \otimes A_2^n$$

$$= (G_\mu \otimes \mathbf{Id}_{2p\times 2p})(D(\lambda_i(\mu)) \otimes A_1^n + \mathbf{Id}_{2p\times 2p} \otimes A_2^n)(G_\mu^{-1} \otimes \mathbf{Id}_{2p\times 2p}).$$

This equation means that representation (1.34) can be rewritten as:

$$\sigma_N^2 \Psi_N^1 = \Psi_0 \mathbf{J}(G_\mu \otimes \mathbf{Id}_{2p\times 2p}) D(S_{N-1}(\lambda_i(\mu)))(G_\mu^{-1} \otimes \mathbf{Id}_{2p\times 2p})\mathbf{J}^* \Psi_0^*, \tag{1.44}$$

where $D(S_{N-1}(\lambda_i(\mu)))$ is a block diagonal matrix with the block entries $S_{N-1}(\lambda_i(\mu))$, $i \leq 2p$ defined by equation (1.39). Since G_0 is a lower triangular matrix, it follows from (1.44) that

$$\sigma_N^2 \Psi_N^1 = P_\mu \mathcal{D}_1(S_{N-1}(\lambda_i(\mu)))Q_\mu + R_\mu \mathcal{D}_2(S_{N-1}(\lambda_j(\mu)))T_\mu,$$

where

$$\lim_{\mu\to 0} P_\mu Q_\mu = \mathbf{Id}_{2p\times 2p}, \quad \lim_{\mu \to 0} R_\mu = \mathbf{0}_{2p\times 2p},$$

and the block diagonal matrix $\mathcal{D}_1(S_{N-1}(\lambda_i))$ (respectively $\mathcal{D}_2(S_{N-1}(\lambda_j))$) is such that $|\lambda_i(\mu)| > 1$ (respectively $|\lambda_j(\mu)| < 1$).

Since $\det \mathcal{D}_1(S_{N-1}(\lambda_i(\mu))) = (\sigma_N^2)^p \prod_{i=1}^p [\lambda_i(\mu)]^{N-1}$ then, by Lemma 1.2 we get

$$\lim_{N\to\infty} \frac{\det \sigma_N^2 \Psi_N^1}{\det \mathcal{D}_1(S_{N-1}(\lambda_i(\mu)))} = 1,$$

□

The following statement plays a crucial role in the proofs.

Lemma 1.4. *Supposing that $r(\vartheta) < 1$. Then under condition (1.2), for any $\alpha \in \mathbb{R}^p$,*

$$\lim_{N\to\infty} L_N^\vartheta\left(\frac{1}{N}\right) = \exp\left(-\frac{1}{2}\alpha^* \mathcal{I}(\vartheta)\alpha\right) \tag{1.45}$$

where $\mathcal{I}(\vartheta)$ is the unique solution of Lyapunov equation (1.9).

Proof. Following from Lemma 1.3, the limit of $L_N^\vartheta(\frac{1}{N})$ is equal to the limit $\overline{L}_N^\vartheta\left(\frac{1}{N}\right)$ under the condition (1.2) . With Taylor's development, we have

$$\lim_{N\to\infty} \overline{L}_N^\vartheta\left(\frac{1}{N}\right) = \exp\left(-\frac{1}{2}\sum_{i=1}^p \frac{\lambda_i'(0)}{\lambda_i(0)}\right)$$

where $\lambda_i'(0)$ denotes the derivative of the function λ_i with respect to 0. Now we only need to prove

$$\sum_{i=1}^p \left(\frac{\lambda_i'(0)}{\lambda_i(0)}\right) = \alpha^* \mathcal{I}(\vartheta)\alpha$$

where $\mathcal{I}(\vartheta)$ is the unique solution of the Lyapunov equation (1.9).

Let us recall that

$$\mathcal{A}_\mu = \mathcal{A}_0 + \mu H$$

where

$$\mathcal{A}_0 = \begin{pmatrix} A_0^{-1} & A_0^{-1}bb^* \\ \mathbf{0}_{p\times p} & A_0^* \end{pmatrix} \quad H = \begin{pmatrix} \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p} \\ \alpha\alpha^* & \alpha\alpha^* A_0^{-1} bb^* \end{pmatrix}.$$

We define the determinant polynomial $\mathbf{P}(\lambda, \mu) = \det(\mathcal{A}_0 - \lambda \mathbf{Id} + \mu H) = 0$, then

$$\mathbf{P}_{\mu}^{'} + \mathbf{P}_{\lambda}^{'} \lambda_{\mu}^{'} = 0$$

where $f_{\mu}^{'}$ denotes the partial differential of function f with respect to μ. We can get that

$$\lambda_{\mu}^{'} = -\frac{\mathbf{P}_{\mu}^{'}(\lambda, \mu)}{\mathbf{P}_{\lambda}^{'}(\lambda, \mu)}.$$

In fact

$$\mathbf{P}_{\mu}^{'}(\lambda, 0) = \det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}\left((\mathcal{A}_0 - \lambda \mathbf{Id})^{-1} H\right)$$

and

$$\mathbf{P}_{\lambda}^{'}(\lambda, 0) = -\det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}\left((\mathcal{A}_0 - \lambda \mathbf{Id})^{-1}\right).$$

With some computation we will see that

$$(\mathcal{A}_0 - \lambda \mathbf{Id})^{-1} = \begin{pmatrix} (A_0^{-1} - \lambda \mathbf{Id})^{-1} & \mathcal{Q} \\ \mathbf{0}_{p\times p} & (A_0^{*} - \lambda \mathbf{Id})^{-1} \end{pmatrix}$$

where

$$\mathcal{Q} = -(\mathbf{Id} - \lambda A_0)^{-1} b b^* (A_0^* - \lambda \mathbf{Id})^{-1}.$$

So when $\lambda = \lambda_i(0) > 1$

$$\det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}(\mathcal{A}_0 - \lambda \mathbf{Id}) = \det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}(A_0^{-1} - \lambda \mathbf{Id})^{-1}$$

and

$$\operatorname{trace}\left((\mathcal{A}_0 - \lambda \mathbf{Id})^{-1} H\right) = \operatorname{trace}\left[\mathcal{Q}\alpha\alpha^* A_0^{-1} + (A_0^* - \lambda \mathbf{Id})^{-1} \alpha\alpha^* A_0^{-1} b b^*\right]$$

that is to say when $\lambda = \lambda_i(0)$

$$\begin{aligned} \det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}\left((\mathcal{A}_0 - \lambda \mathbf{Id}) H\right) &= \det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}\left(A_0^{-1} \mathcal{Q}\alpha\alpha^*\right) \\ &= \alpha^* \left[\det(\mathcal{A}_0 - \lambda \mathbf{Id}) A_0^{-1} \mathcal{Q}\right] \alpha \end{aligned}$$

which denotes that

$$\begin{aligned} \mathcal{I}(\vartheta) &= \sum_{i=1}^{p} \frac{1}{\lambda_i(0)} \left(\frac{\det(\mathcal{A}_0 - \lambda \mathbf{Id}) A_0^{-1} \mathcal{Q}}{\det(\mathcal{A}_0 - \lambda \mathbf{Id}) \cdot \operatorname{trace}(A_0^{-1} - \lambda \mathbf{Id})^{-1}} \right)_{\lambda = \lambda_i(0)} \\ &= \sum_{i=1}^{p} \frac{1}{\lambda_i(0)} \left(\frac{\prod_{j=1}^{p} (\lambda - \lambda_j(0)) (A_0^{-1} \mathcal{Q})}{\prod_{j=1}^{p} (\lambda - \lambda_j(0)) \cdot \operatorname{trace}(A_0^{-1} - \lambda \mathbf{Id})^{-1}} \right)_{\lambda = \lambda_i(0)}. \end{aligned}$$

Let

$$A_0^{-1} = G^{-1} \mathcal{D}\left(\lambda_\ell(0)\right) G,\ A_0 = G^{-1} \mathcal{D}\left(\frac{1}{\lambda_\ell(0)}\right) G \quad \text{and} \quad A_0^* = G^* \mathcal{D}\left(\frac{1}{\lambda_\ell(0)}\right) (G^*)^{-1}$$

when

$$A_0^{-1} \mathcal{Q} = -A_0^{-1} (\mathbf{Id} - \lambda A_0)^{-1} b b^* (A_0^* - \lambda \mathbf{Id}),$$

we have

$$(\mathbf{Id} - \lambda A_0)^{-1} = G^* \mathcal{D}\left(\frac{\lambda_\ell(0)}{1-\lambda\lambda_\ell(0)}\right)(G^*)^{-1}$$

and

$$A_0^{-1}\mathcal{Q} = -G^{-1}\mathcal{D}(\lambda_\ell(0))\mathcal{D}\left(\frac{\lambda_\ell(0)}{\lambda_\ell(0)\lambda}\right) Gbb^*(A^* - \lambda\mathbf{Id})^{-1}.$$

First of all, we can compute that when $\lambda = \lambda_i(0)$

$$\prod_{j=1}^{p}(\lambda - \lambda_j(0)) \cdot \mathrm{trace}(A_0^{-1} - \lambda\mathbf{Id})^{-1} = \prod_{i\neq j}(\lambda_i(0) - \lambda_j(0))$$

and

$$\prod_{j=1}^{p}(\lambda - \lambda_j(0))(A_0^{-1}\mathcal{Q}) = \left(G^{-1}\mathcal{D}(\lambda_\ell(0)\prod_{j\neq\ell}(\lambda_i(0)-\lambda_j(0)))\right)\cdot G \cdot bb^*(A_0^* - \lambda_j(0)\mathbf{Id})^{-1}$$

then we will get that

$$\begin{aligned}
\mathcal{I}(\vartheta) &= \sum_{i=1}^{p}\left(G^{-1}\mathcal{D}\left(\frac{\lambda_\ell(0)\prod_{s\neq\ell}(\lambda_i(0)-\lambda_s(0))}{\lambda_i(0)\prod_{i\neq j}(\lambda_i(0)-\lambda_j(0))}\right)G\right)bb^*(A_0^* - \lambda_i(0)\mathbf{Id})^{-1} \\
&= \sum_{i=1}^{p} G^{-1}e^j Gbb^*\left(\frac{1}{\lambda_i(0)}A_0^* - \mathbf{Id}\right)^{-1} \\
&= -\sum_{i=1}^{p} G^{-1}e^j Gbb^*(\mathbf{Id} - x_j A_0^*)^{-1}
\end{aligned}$$

where e^i is $p \times p$ matrix with the $(i,i)-th$ component is 1, the others are 0, x_i is the eigenvalue of A_0. With Taylor's development, we have

$$\begin{aligned}
\mathcal{I}(\vartheta) &= \sum_{i=1}^{p}\sum_{n\geq 0} G^{-1}e^j Gbb^* x_j^n (A_0^*)^n \\
&= \sum_{n\geq 0}\left(\sum_{i=1}^{p} G^{-1}e^j Gbb^* x_j^n\right)(A_0^*)^n \\
&= \sum_{n\geq 0} G^{-1}\mathcal{D}(x_j^n)Gbb^*(A_0^*)^n
\end{aligned}$$

as $G^{-1}\mathcal{D}(x_j^n)G = A_0^n$, we have

$$\mathcal{I}(\vartheta) = \sum_{n\geq 0} A_0^n bb^*(A_0^*)^n.$$

It is easy to verify that $\mathcal{I}(\vartheta)$ is the unique solution of Lyapunov equation (1.9). □

5 Proofs

5.1 Proof of Theorem 1.1

The statement of Theorem follows from Lemma 1.4 since (1.45) implies immediately that

$$\mathbf{P}_\vartheta - \lim_{N\to\infty} \frac{1}{N}\langle M\rangle_N = \mathcal{I}(\vartheta), \tag{1.46}$$

and, hence also due to the central limit theorem for martingales,

$$\frac{1}{\sqrt{N}} M_N \overset{law}{\Rightarrow} \mathcal{N}\left(\mathbf{0},\, \mathcal{I}(\vartheta)\right).$$

5.2 Proof of Theorem 1.2

Due to the strong law of large numbers for martingales, in order to proof the strong consistency we have only to check that

$$\lim_{N\to\infty} \langle M\rangle_N = +\infty \quad a.s.,$$

or, equivalently that for a one fixed constant $\mu > 0$

$$\lim_{N\to\infty} \mathbf{E}_\vartheta \exp\left(-\frac{\mu}{2}\langle M\rangle_N\right) = 0. \tag{1.47}$$

But in the case when $p = 1$ the ingredients in the right hand side of formulas (1.33)-(1.34) with $\alpha = 1$ can be given more explicitly:

$$\prod_{n=1}^{N} \det \mathbf{A}_n = \vartheta^N \sigma_{N+1}^2,$$

and

$$\sigma_{N+1}^2 \Psi_{N+1}^1 = \frac{1-\lambda_-}{\lambda_+ - \lambda_-} S_N\left(\frac{\lambda_+}{\vartheta}\right) + \frac{\lambda_+ - 1}{\lambda_+ - \lambda_-} S_N\left(\frac{\lambda_-}{\vartheta}\right), \tag{1.48}$$

where the matrix $S_N(a)$ is defined by equation (1.39),

$$\frac{\lambda_\pm}{\vartheta} = \frac{\vartheta^2 + \mu + 1 \pm \sqrt{(\mu + (1-\vartheta)^2)(\mu + (1+\vartheta)^2)}}{2\vartheta}$$

are the two eigenvalues of the matrix $\mathcal{A}_\mu = \begin{pmatrix} \frac{1}{\vartheta} & \frac{1}{\vartheta} \\ \mu\frac{1}{\vartheta} & \mu\frac{1}{\vartheta} + \vartheta \end{pmatrix}$ Note that $\frac{\lambda_+}{\vartheta}\frac{\lambda_-}{\vartheta} = 1$, $\left|\frac{\lambda_+}{\vartheta}\right| > 1$ and $\lambda_+ > 1$ for every $\mu > 0$ and $\vartheta \in \mathbb{R}$. Equations (1.48) and (1.39) imply that for $\kappa = \frac{\lambda_+ - 1}{1 - \lambda_-}$

$$\det \Psi_{N+1}^1 = \left(\frac{1}{\sigma_{N+1}^2}\right)^2 \left(\frac{1-\lambda_-}{\lambda_+ - \lambda_-}\right)^2 \det\left(S_N\left(\frac{\lambda_+}{\vartheta}\right)\right) \det\left(\mathbf{Id}_{2\times 2} + \kappa\left(S_N\left(\frac{\lambda_+}{\vartheta}\right)\right)^{-1} S_N\left(\frac{\lambda_-}{\vartheta}\right)\right)$$

and that

$$\det\left(S_N\left(\frac{\lambda_+}{\vartheta}\right)\right) = \vartheta^{-N}\lambda_+^N \sigma_{N+1}^2.$$

Following from the Lemma 1.2, we have

$$\det\left(\mathbf{Id}_{2\times 2}+\kappa(S_N(\frac{\lambda_+}{\vartheta}))^{-1}S_N(\frac{\lambda_-}{\vartheta})\right)$$

is uniformly bounded and separated from 0 when μ is sufficiently large (and so $a=\frac{\lambda_-}{\vartheta}$ is sufficiently small). Since $\lambda_+>1$, we obtain that

$$\lim_{N\to\infty}L_N^{\vartheta}(\mu)=c\lim_{N\to\infty}\lambda_+^{-\frac{N}{2}}=0.$$

The uniform consistency and the uniform convergence of the moments on compacts $\mathbb{K}\subset(-1,\,1)$ follow from the estimates (see [48], Eq.17.51):

$$\mathbf{E}_{\vartheta}\left(\frac{1}{N}\langle M\rangle_N\right)^{-q}\le(1-\vartheta^2)^{-q},$$

$$\mathbf{E}\left(\frac{1}{\sqrt{N}}M_N\right)^{q}\le\left(\sqrt{1-\vartheta^2}\right)^{q}.$$

Remark 6. *It is worth mentioning that even in a stationary autoregressive models of order* 1 *with strongly dependent noises the Least Square Estimator* $\widetilde{\vartheta}_N=\frac{\sum_{n=1}^{N}X_{n-1}X_n}{\sum_{n=1}^{N}X_{n-1}^2}$ *is not consistent.*

Appendix – Laplace transforms of quadratic forms for general Gaussian vector sequences

In this part, we consider the Laplace transforms of quadratic forms corresponding to the Gaussian vector sequence $(X_t,\,t=0,1,\cdots)$ and the given deterministic symmetric matrix sequence $Q(t),\,t=0,1,\cdots$:

$$\mathcal{L}(t)=\mathbf{E}\exp\left(-\frac{1}{2}\sum_{s=0}^{t}X_s'Q(s)X_s\right).$$

We state our main result:

Theorem A.1 For any $t\ge0$ the following equality holds:

$$\mathcal{L}(t)=\prod_{s=0}^{t}[\det(\mathbf{Id}+\gamma(s,s)Q(s)]^{-1/2}\exp\left\{-\frac{1}{2}\sum_{s=0}^{t}z'(s)Q(s)(\mathbf{Id}+\gamma(s,s)Q(s))^{-1}z(s)\right\},\tag{1.49}$$

where $(\gamma(t,s),\;0\le s\le t)$ is the unique solution of the equation

$$\gamma(t,s)=K(t,s)-\sum_{r=0}^{s-1}\gamma(t,r)[\mathbf{Id}+\gamma(r,r)Q(r)]^{-1}Q(r)\gamma(s,r)',\;1\le s\le t;\quad\gamma(t,0)=K(t,0),\tag{1.50}$$

and $(z_s,\;0\le s\le t)$ is the unique solution of the equation

$$z_s=m_s-\sum_{r=0}^{s-1}\gamma(s,r)[\mathbf{Id}+\gamma(r,r)Q(r)]^{-1}Q(r)z_r\,,\;1\le s\le t;\quad z_0=m_0\,.\tag{1.51}$$

where

$$m_t=\mathbf{E}(X_t),\;K(t,s)=\mathbf{E}(X_t-m_t)(X_s-m_s)'.$$

To prove this Theorem, we need the following Lemma:

Lemma A.2 Let V be a random variable and U be a p-dimension random vector, Q is a symmetric matrix, then

$$\begin{aligned}\frac{\mathbf{E}e^{-V-\frac{1}{2}U'QU}}{\mathbf{E}e^{-V}} &= [\det(\mathbf{Id}+\gamma_{UU}Q)]^{-1/2} \\ &\times \exp\left\{-\frac{1}{2}(\mathbf{E}(U)-\gamma_{UV})'Q(\mathbf{Id}+\gamma_{UU})^{-1}(\mathbf{E}(U)-\gamma_{UV})\right\}.\end{aligned}$$

where

$$\gamma_{UU}=\mathbf{E}(U-\mathbf{E}(U))(U-\mathbf{E}(U))' \quad and \quad \gamma_{UV}=\mathbf{E}(U-\mathbf{E}(U))(V-\mathbf{E}(V))$$

Proof. Let us define $\zeta=\begin{pmatrix}U\\V\end{pmatrix}$, $a=\mathbf{E}\zeta=\begin{pmatrix}\mathbf{E}(U)\\\mathbf{E}(V)\end{pmatrix}$, $\Gamma=\mathbf{E}(\zeta-a)(\zeta-a)'=\begin{pmatrix}\gamma_{UU}&\gamma_{UV}\\\gamma'_{UV}&\gamma_{VV}\end{pmatrix}$, $d=\begin{pmatrix}\mathbf{0}_{p\times 1}\\-1\end{pmatrix}$ and $D=\begin{pmatrix}Q&\mathbf{0}_{p\times 1}\\\mathbf{0}_{1\times p}&0\end{pmatrix}$. With the formula

$$\begin{aligned}\mathbf{E}\exp[d'\zeta-\zeta'D\zeta] &= \left(\frac{\det[2D+\Gamma^{-1}]^{-1}}{\det\Gamma}\right)^{1/2} \\ &\times \exp\left\{\frac{1}{2}\left[-a'\Gamma^{-1}a+(d'+a'\Gamma^{-1})(2D+\Gamma^{-1})^{-1}(d'+a'\Gamma^{-1})'\right]\right\}.\end{aligned}$$

we have

$$\begin{aligned}\frac{\mathbf{E}e^{-V-\frac{1}{2}U'QU}}{\mathbf{E}e^{-V}} &= \left(\frac{\det[D+\Gamma^{-1}]^{-1}}{\det\Gamma}\right)^{1/2} \\ &\times \exp\left\{\frac{1}{2}\left[(d'+a'\Gamma^{-1})[(D+\Gamma^{-1})^{-1}-\Gamma](d+\Gamma^{-1}a)\right]\right\}\end{aligned}$$

First of all,

$$\frac{\det[D+\Gamma^{-1}]^{-1}}{\det\Gamma}=\det[\Gamma(D+\Gamma^{-1})]^{-1}=[\det(\mathbf{Id}+\gamma_{UU}Q)]^{-1},$$

on the other hand, we can calculate that

$$(D+\Gamma^{-1})^{-1}-\Gamma=\begin{pmatrix}(\mathbf{Id}+\gamma_{UU}Q)^{-1}-\mathbf{Id}&\mathbf{0}_{p\times 1}\\-\gamma'_{UV}Q(\mathbf{Id}+\gamma_{UU}Q)^{-1}&0\end{pmatrix}\times\Gamma$$

so we have

1)

$$d'[(D+\Gamma^{-1})^{-1}-\Gamma]d=-\gamma'_{UV}Q(\mathbf{Id}+\gamma_{UU}Q)^{-1}\gamma_{UV},$$

2)

$$d'[(D+\Gamma^{-1})^{-1}-\Gamma]\Gamma^{-1}a=\gamma'_{UV}Q(\mathbf{Id}+\gamma_{UU}Q)^{-1}\mathbf{E}(U),$$

3)

$$\begin{aligned}
a^{'}\Gamma^{-1}[(D+\Gamma^{-1})^{-1}-\Gamma]\Gamma^{-1}a &= a^{'}\Gamma^{-1}\begin{pmatrix} (\mathbf{Id}+\gamma_{UU}Q)^{-1}-\mathbf{Id} & \mathbf{0}_{p\times 1} \\ -\gamma_{UV}^{'}Q(\mathbf{Id}+\gamma_{UU}Q)^{-1} & 0 \end{pmatrix} a \\
&= a^{'}\begin{pmatrix} -Q(\mathbf{Id}+\gamma_{UU}Q)^{-1} & \mathbf{0}_{p\times 1} \\ \mathbf{0}_{1\times p} & 0 \end{pmatrix} a \\
&= -\mathbf{E}(U)^{'}Q(\mathbf{Id}+\gamma_{UU}Q)^{-1}\mathbf{E}(U).
\end{aligned}$$

Because

$$\begin{aligned}
(d^{'}+a^{'}\Gamma^{-1})[(D+\Gamma^{-1})^{-1}-\Gamma](d+\Gamma^{-1}a) &= 2d^{'}[(D+\Gamma^{-1})^{-1}-\Gamma]\Gamma^{-1}a \\
&+ a^{'}\Gamma^{-1}[(D+\Gamma^{-1})^{-1}-\Gamma]\Gamma^{-1}a + d^{'}[(D+\Gamma^{-1})^{-1}-\Gamma]d
\end{aligned}$$

we have

$$\begin{aligned}
\frac{\mathbf{E}e^{-V-\frac{1}{2}U^{'}QU}}{\mathbf{E}e^{-V}} &= [\det(\mathbf{Id}+\gamma_{UU}Q)]^{-1/2} \\
&\times \exp\left\{-\frac{1}{2}(\mathbf{E}(U)-\gamma_{UV})^{'}Q(\mathbf{Id}+\gamma_{UU})^{-1}(\mathbf{E}(U)-\gamma_{UV})\right\}.
\end{aligned}$$

□

Now we introduce the problems appropriate for computing the Ltqf. Let $(\varepsilon_t,\ t=0,1,\dots)$ be a sequence of i.i.d. standard Gaussian random vectors which is independent of the given process $(X_t,\ t=0,1,\dots)$. Let us define the auxiliary sequences $(Y_t,\ t=0,1,\dots)$ and $(\xi_t,\ t=0,1,\dots)$ by

$$\begin{aligned}
Y_t &= Q(t)X_t + Q^{1/2}(t)\varepsilon_t\,, \\
\xi_t &= \sum_{s=0}^{t} X_s^{'}Y_s\,.
\end{aligned} \tag{1.52}$$

The notation $\pi_s(X_t)$ is used for the conditional expectation of X_t given σ-field $\mathcal{Y}_t=\sigma(\{Y_s\,,\ 0\le s\le t\})$:

$$\pi_s(X_t)=\mathbf{E}(X_t|\mathcal{Y}_s).$$

Moreover we make the convention that $\pi_{-1}(X_t)=\mathbf{E}(X_t)$. We shall be concerned with one-step prediction for X from Y and with filtering ξ from Y. Here, clearly the pair (X,Y) is jointly Gaussian, and hence the optimal one-step predictor is the Gaussian distribution defined by the conditional mean $\pi_{t-1}(X_t)$ and the conditional variance matrix $\gamma_{XX}(t)=\mathbf{E}[(X_t-\pi_{t-1}(X_t))(X_t-\pi_{t-1}(X_t))^{'}/\mathcal{Y}_{t-1}]$ which actually is deterministic *i.e.*,

$$\gamma_{XX}(t)=\mathbf{E}[X_t-\pi_{t-1}(X_t)][X_t-\pi_{t-1}(X_t)]^{'}\,,\ t\ge 1\,;\quad \gamma_{XX}(0)=K(0,0)\,. \tag{1.53}$$

Of course, the joint distribution of (X,ξ,Y) is not Gaussian, but we observe that the conditional distribution of (X_t,ξ_{t-1}) given $\mathcal{Y}_{t-1}$ is Gaussian. Hence, in particular, the optimal filter for ξ is the Gaussian distribution defined by the conditional mean $\pi_t(\xi_t)$ and the corresponding conditional covariance (which is random). Actually the other main characteristic which is involved in the sequel is the following conditional covariance :

$$\gamma_{X\xi}(t)=\mathbb{E}[(X_t-\pi_{t-1}(X_t))(\xi_{t-1}-\pi_{t-1}(\xi_{t-1}))/\mathcal{Y}_{t-1}]\,,\ t\ge 1\,;\quad \gamma_{X\xi}(0)=\mathbf{0}\,. \tag{1.54}$$

Now we can state the announced key property :

Lemma A.3 For any $t = 0, 1, \ldots$ the following equality holds:

$$\begin{aligned}\mathcal{L}(t) &= \prod_{s=0}^{t} \left[\det(\mathbf{Id} + \gamma_{XX}(s)Q(s))\right]^{-1/2} \\ &\times \exp\left\{ -\frac{1}{2}\sum_{s=0}^{t}[\pi_{s-1}(X_s) - \gamma_{X\xi}(s)]Q(s)(\mathbf{Id} + \gamma_{XX})^{-1}[\pi_{s-1}(X_s) - \gamma_{X\xi}(s)]' \right\}\end{aligned}$$

Proof. Setting

$$I_{t-1} = \frac{1}{2}\sum_{s=0}^{t-1} X_s Q(s) X_s',$$

we can write

$$\frac{\mathcal{L}(t)}{\mathcal{L}(t-1)} = \frac{\mathbf{E}(\exp\left\{-I_{t-1} - \frac{1}{2}X_t Q(t) X_t'\right\})}{\mathbf{E}(\exp\{-I_{t-1}\})}. \tag{1.55}$$

Let us define a new probability measure $\widetilde{\mathbf{P}}$ by

$$d\widetilde{\mathbf{P}} = \exp\{-\zeta_{t-1}\}d\mathbf{P}\,; \quad \zeta_{t-1} = \sum_{s=0}^{t-1} X_s' Q^{1/2}(s)\varepsilon_s + \frac{1}{2}\sum_{s=0}^{t-1} X_s Q(s) X_s'. \tag{1.56}$$

Under $\widetilde{\mathbf{P}}$ the distribution of X is the same as under $\mathbf{P}$ and X is independent of $(Y_s, 0 \le s \le t-1)$. Hence we can rewrite the equality (1.55) as

$$\frac{\mathcal{L}(t)}{\mathcal{L}(t-1)} = \frac{\widetilde{\mathbf{E}}(\exp\left\{-I_{t-1} - \frac{1}{2}X_t Q(t) X_t'\right\}/\mathcal{Y}_{t-1})}{\widetilde{\mathbf{E}}(\exp\{-I_{t-1}\}/\mathcal{Y}_{t-1})},$$

where $\widetilde{\mathbf{E}}(./\mathcal{Y}_{t-1})$ denotes a conditional expectation computed with respect to $\widetilde{\mathbf{P}}$. Then, using the classical Bayes formula, again we can rewrite (1.55) as

$$\frac{\mathcal{L}(t)}{\mathcal{L}(t-1)} = \frac{\mathbf{E}(\exp\left\{-I_{t-1} - \frac{1}{2}X_t Q(t) X_t'\right\}\exp\{-\zeta_{t-1}\}/\mathcal{Y}_{t-1})}{\mathbf{E}(\exp\{-I_{t-1}\}\exp\{-\zeta_{t-1}\}/\mathcal{Y}_{t-1})}.$$

Since from the definitions (1.52) and (1.56) we have $\xi_{t-1} = I_{t-1} + \zeta_{t-1}$, this means that

$$\frac{\mathcal{L}(t)}{\mathcal{L}(t-1)} = \frac{\mathbf{E}(\exp\left\{-\xi_{t-1} - \frac{1}{2}X_(t)Q(t)X_t'\right\}/\mathcal{Y}_{t-1})}{\mathbf{E}(\exp\{-\xi_{t-1}\}/\mathcal{Y}_{t-1})}.$$

with Lemma 1, we get

$$\begin{aligned}\frac{\mathcal{L}(t)}{\mathcal{L}(t-1)} &= \left[\det(\mathbf{Id} + \gamma_{XX}(t)Q(t))\right]^{-1/2} \\ &\times \exp\left\{ -\frac{1}{2}[\pi_{t-1}(X_t) - \gamma_{X\xi}(t)]Q(s)(\mathbf{Id} + \gamma_{XX})^{-1}[\pi_{t-1}(X_s) - \gamma_{X\xi}(t)]' \right\}\end{aligned}$$

which achieves the proof. □

The last thing to prove Theorem 1 is to get the equation (1.50). Since for the general setting the analysis is quite similar, for simplicity of notation we deal only with the case $Q \equiv \mathbf{Id}$, *i.e.*, $Y_t = X_t + \varepsilon_t$. Since the joint distribution of (X_r', Y_s) for any $r\,,s$ is Gaussian we can apply the Note following Theorem 13.1 in [49]. For any l we can write

$$\begin{cases} \pi_l(X_t) = \pi_{l-1}(X_t) + \gamma(t,l)\langle\nu\rangle_l^{-1}\nu_l, \\ \pi_{-1}(X_t) = m_t, \end{cases} \tag{1.57}$$

where $\nu_l = Y_l - \mathbf{E}(Y_l/\mathcal{Y}_{l-1}) = Y_l - \pi_{l-1}(X_l)$ is the innovation and $\langle\nu\rangle_l$ is its variance matrix

$$\langle\nu\rangle_l = \mathbf{Id} + \gamma(l,l)\,,$$

with

$$\gamma(t,l) = \mathbf{E}(X_t - \pi_{l-1}(X_t))(X_l - \pi_{l-1}(X_l))'\,. \tag{1.58}$$

By the definition (1.58), we see for $l = t$ that the variance matrix $\gamma_{XX}(t) = \gamma(t,t)$ and Now, equality (1.57) implies

$$\pi_l(X_t) = m_t + \sum_{r=0}^{l} \gamma(t,r)(\mathbf{Id} + \gamma_{XX}(r))[Y_r - \pi_{r-1}(X_r)]\,,$$

and putting $l = t-1$ we get nothing but equation

$$\pi_{t-1}(X_t) = m_t + \sum_{s=0}^{t-1} \gamma(t,s)(\mathbf{Id} + \gamma_{XX}(s))[Y_s - \pi_{s-1}(X_s)]\,, t \geq 0\,, \tag{1.59}$$

Let us define

$$\delta_X(t,l) = X_t - \pi_l(X_t),$$

According to (1.57) we can write

$$\delta_X(t,l) = \delta_X(t,l-1) - \gamma(t,l)\langle\nu\rangle_l^{-1}\nu_l,$$

and so

$$\mathbf{E}\delta_X(t^1,l)\delta_X'(t^2,l) = \mathbf{E}\delta_X(t^1,l-1)\delta_X'(t^2,l-1) - \gamma(t^1,l)\langle\nu\rangle_l^{-1}\gamma(t^2,l)'\,,$$

or

$$\mathbf{E}\delta_X(t^1,l)\delta_X'(t^2,l) = \mathbf{E}\delta_X^{(}t^1,-1)\delta_X'(t^2,-1) - \sum_{r=0}^{l} \gamma(t^1,r)\langle\nu\rangle_r^{-1}\gamma(t^2,r)'. \tag{1.60}$$

Taking $t^1 = t\,, t^2 = s\,, l = s-1$ in (1.60), it is readily seen that equation (1.50) holds for $\gamma(t,s)$. Now we analyze the difference $\pi_{t-1}(X_t) - \gamma_{X\xi}(t)$. Using the representation $\xi_t = \sum_{r=0}^{t} X_r' Y_r$ we can rewrite $\gamma_{X\xi}(t)$ in the following form

$$\begin{aligned} \gamma_{X\xi}(t) &= \pi_{t-1}(X_t - \pi_{t-1}(X_t))(\xi_{t-1} - \pi_{t-1}(\xi_{t-1})) \\ &= \sum_{r=0}^{t-1} \pi_{t-1}((X_t - \pi_{t-1}(X_t))(X_r - \pi_{t-1}(X_r))')Y_r \\ &= \sum_{r=0}^{t-1} \mathbf{E}((X_t - \pi_{t-1}(X_t))(X_r - \pi_{t-1}(X_r))')Y_r\,. \end{aligned}$$

So we have

$$\gamma_{X\xi}(t) = \sum_{r=0}^{t-1} \widetilde{\gamma}(t,r)Y_r\,, \tag{1.61}$$

where

$$\widetilde{\gamma}(t,r) = \mathbf{E}((X_t - \pi_{t-1}(X_t))(X_r - \pi_{t-1}(X_r))^{'}) = \gamma(r,t)^{'} . \qquad (1.62)$$

Using the definitions (1.58) and (1.62) we can write

$$\widetilde{\gamma}(t,r) - \gamma(t,r) = -\mathbf{E}X_t(\pi_{t-1}(X_r) - \pi_{r-1}(X_r))^{'}.$$

Again, applying the Note following Theorem 13.1 in [49], we can write also

$$\pi_l(X_r) = \pi_{l-1}(X_r) + \gamma(r,l)\langle\nu\rangle_l^{-1}\nu_l,$$

This means that

$$\pi_{t-1}(X_r) - \pi_{r-1}(X_r) = \sum_{l=r}^{t-1}\gamma(r,l)\langle\nu\rangle_l^{-1}\nu_l\,,$$

so

$$\mathbf{E}X_t(\pi_{t-1}(X_r) - \pi_{r-1}(X_r))^{'} = \sum_{l=r}^{t-1}\gamma(t,l)\langle\nu\rangle_l^{-1}\widetilde{\gamma}(l,r).$$

Hence we have proved the following relation

$$\widetilde{\gamma}(t,r) - \gamma(t,r) = -\sum_{l=r}^{t-1}\gamma(t,l)\langle\nu\rangle_l^{-1}\widetilde{\gamma}(l,r)\,. \qquad (1.63)$$

Now we can show that the difference $z_t = \pi_{t-1}(X_t) - \gamma_{X\xi}(t)$ satisfies the equation (1.51). Using (1.59), (1.61) and (1.63), we obtain

$$\begin{aligned}
z_t &= m_t + \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}(Y_r - \pi_{r-1}(X_r)) - \sum_{r=0}^{t-1}\widetilde{\gamma}(t,r)Y_r \\
&= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}\pi_{r-1}(X_r) + \sum_{r=0}^{t-1}(\gamma(t,r)\langle\nu\rangle_r^{-1} - \widetilde{\gamma}(t,r))Y_r \\
&= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}\pi_{r-1}(X_r) + \sum_{r=0}^{t-1}(\gamma(t,r)\langle\nu\rangle_r^{-1} - (\gamma(t,r) - \sum_{l=r}^{t-1}\gamma(t,l)\langle\nu\rangle_l^{-1}\widetilde{\gamma}(l,r)))Y_r \\
&= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}\pi_{r-1}(X_r) - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}\gamma(r,r))Y_r + \sum_{l=0}^{t-1}\gamma(t,l\langle\nu\rangle_l^{-1}(\sum_{r=0}^{l}\widetilde{\gamma}(l,r)Y_r\,,
\end{aligned}$$

Now, using 1.61 again and the property $\gamma(r,r) = \widetilde{\gamma}(r,r)$, we can write

$$\begin{aligned}
z_t &= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}\pi_{r-1}(X_r) + \sum_{l=0}^{t-1}\gamma(t,l)\langle\nu\rangle_l^{-1}\gamma_{X\xi}(l) \\
&= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}(\pi_{r-1}(X!r) - \gamma_{X\xi}(r)) \\
&= m_t - \sum_{r=0}^{t-1}\gamma(t,r)\langle\nu\rangle_r^{-1}z_r\,,
\end{aligned}$$

which is nothing else but equation (1.51).

Chapter 2

Controlled drift estimation in fractional diffusion linear systems

1 Introduction

1.1 Historical survey

The experiment design has been given a great deal of interest over the last decades from the early statistics literature (see *e.g.* [34, 76, 79]) as well as in the engineering literature (see *e.g.* [24–26]).

Many of these works focused on identification of directly observed dynamic system parameters. The classical approach for experiment design consists on a two-step procedure: maximize the Fisher information under energy constraint of the input and find an adaptive estimation procedure. In this area, there are several approaches like sequential design and Bayesian design (see *e.g.* [26, 46, 53] and the references therein).

For partially observed systems, even in the linear case we can only mention [4, 47, 53, 54, 57], where linear signal - observation model perturbed by the white noise has been considered.

On the other hand, large sample asymptotic properties of the Maximum Likelihood Estimator (MLE) of the drift of a fractional Ornstein-Uhlenbeck process [18, 37] have also been studied in the directly observed case (see [37] for consistency and [8, 10, 19] for asymptotical normality).

The work which is presented in this part is a direct continuation of what has been initiated in [10–12]. We focus on the determination of the asymptotical optimal input for the estimation of the drift parameter in a partially observed but controlled fractional Ornstein-Uhlenbeck process.

More precisely, we present here a technique that allows us to use both methods developed in [12] for computing the asymptotical optimal input and in [10] for deducing the drift MLE asymptotical properties. The remainder term appearing can be treated by Laplace transform computations.

This chapter falls into four parts. In this introduction, we state the setting and the main results. In the second part, we present two key elements which are the model transformation and the Fisher information decomposition. In the third part, we focus on the proof for the partially observable case. In the fourth part, we give some remarks on the proof for the directly observed problem. Finally technical proofs of lemmas are postponed in the last part.

1.2 The setting and the main result

We consider real-valued processes $X = (X_t,\, t \geq 0)$ and $Y = (Y_t,\, t \geq 0)$, representing the signal and the observation respectively. In the fully observable case, they are governed by:

$$dY_t = dX_t = -\vartheta X_t dt + u(t)dt + dV_t^H,\; t > 0 \tag{2.1}$$

and in the partially observable case, they are governed by the following linear system of stochastic differential equation:

$$\begin{cases} dX_t &= -\vartheta X_t dt + u(t)dt + dV_t^H, \\ dY_t &= \mu X_t dt + dW_t^H, \end{cases} \quad t > 0, \tag{2.2}$$

with initial condition $X_0 = Y_0 = 0$. Here, $V^H = (V_t^H,\, t \geq 0)$ and $W^H = (W_t^H, t \geq 0)$ are independent normalized fBm's with the same known[1] Hurst parameter $H \in (0,1)$ and the coefficients ϑ and $\mu \neq 0$ are real constants. The unobserved signal process $X = (X_t,\, t \geq 0)$, is controlled by the real-valued function $u = (u(t),\, t \geq 0)$.

The system has a uniquely defined solution process (X, Y) which is, due to the well known properties of the fBm, Gaussian but neither Markovian nor a semimartingale for $H \neq \frac{1}{2}$ (see, *e.g.*, [49], page 238).

Suppose that parameter $\vartheta > 0$ is unknown and is to be estimated given the observed trajectory $Y^T = (Y_t,\, 0 \leq t \leq T)$.

For a fixed value of the parameter ϑ, let $\mathbf{P}_\vartheta^T$ denote the probability measure, induced by (X^T, Y^T) on the function space $\mathcal{C}_{[0,T]} \times \mathcal{C}_{[0,T]}$ and let $\mathcal{F}_t^Y$ be the natural filtration of Y, $\mathcal{F}_t^Y = \sigma\left(Y_s, 0 \leq s \leq t\right)$.

Let $\mathcal{L}(\vartheta, Y^T)$ be the likelihood, *i.e.* the Radon-Nikodym derivative of $\mathbf{P}_\vartheta^T$, restricted to $\mathcal{F}_T^Y$ with respect to some reference measure on $\mathcal{C}_{[0,T]}$. In this setting, Fisher information stands for :

$$\mathcal{I}_T\left(\vartheta, u\right) = -\mathbf{E}_\vartheta \frac{\partial^2}{\partial \vartheta^2} \ln \mathcal{L}(\vartheta, Y^T)\,.$$

Let us denoted $\mathcal{U}_T$ some functional space of controls, that is defined by equations (2.12) and (2.13) page 37. Let us therefore note

$$\mathcal{J}_T(\vartheta) = \sup_{u \in \mathcal{U}_T} \mathcal{I}_T(\vartheta, u).$$

Our main goal is to find estimator $\overline{\vartheta}_T$ of the parameter ϑ which are asymptotically efficient in the sense that, for any compact $\mathbb{K} \subset \mathbb{R}_*^+ = \{\vartheta \in \mathbb{R}, \vartheta > 0\}$,

$$\sup_{\vartheta \in \mathbb{K}} \mathcal{J}_T(\vartheta) \mathbf{E}_\vartheta \left(\overline{\vartheta}_T - \vartheta\right)^2 = 1 + o(1)\,, \tag{2.3}$$

as $T \to \infty$. We claim that:

[1] In the continuous-time observation setting, there is no statistical error made for the Hurst parameter H estimation with classical methods, see for instance quadratic generalized variations method in [32].

Theorem 2.1. *The asymptotical optimal input in the class of controls* $\mathcal{U}_T$ *is* $u_{opt}(t) = \frac{\kappa}{\sqrt{2\lambda}} t^{H-\frac{1}{2}}$ *for* $H > \frac{1}{2}$ *(and* $u_{opt}(t) = \frac{\kappa}{\sqrt{2\lambda}} t^{\frac{1}{2}-H}$ *for* $H < \frac{1}{2}$*) where*

$$\kappa = 2H\Gamma\left(\frac{3}{2} - H\right)\Gamma\left(\frac{1}{2} + H\right) \quad \text{and} \quad \lambda = \frac{H\Gamma(3-2H)\Gamma(H+\frac{1}{2})}{2(1-H)\Gamma(\frac{3}{2}-H)} \tag{2.4}$$

and Γ *stands for the Gamma function. Moreover,*

$$\lim_{T\to+\infty} \frac{\mathcal{J}_T(\vartheta)}{T} = \mathcal{I}(\vartheta)$$

where

$$\mathcal{I}(\vartheta) = \begin{cases} \dfrac{1}{2\vartheta} + \dfrac{1}{\vartheta^2} & \text{(fully observable case)} \\ \dfrac{1}{2\vartheta} - \dfrac{2\vartheta}{\alpha(\alpha+\vartheta)} + \dfrac{\vartheta^2}{2\alpha^3} + \dfrac{\mu^2}{\alpha^2\vartheta^2} & \text{(partially observable case)} \end{cases} \tag{2.5}$$

and $\alpha = \sqrt{\mu^2 + \vartheta^2}$.

Remark 2.1. *In order to compare, we can see that the Fisher information for the problems* (2.1) *and* (2.2) *with no input* $u(t) = 0$ *is*

$$\lim_{T\to+\infty} \frac{\mathcal{I}_T(\vartheta, 0)}{T} = \begin{cases} \dfrac{1}{2\vartheta} & \text{(fully observable case)} \\ \dfrac{1}{2\vartheta} - \dfrac{2\vartheta}{\alpha(\alpha+\vartheta)} + \dfrac{\vartheta^2}{2\alpha^3} & \text{(partially observable case).} \end{cases}$$

This values have been obtained in [8, 10, 19] for the fully observable case and in [10] for the partially observed case, all for the fractional setting.On the other hand, when $H = \frac{1}{2}$ *(classical Wiener case), the optimal input is* $u(t) = 1$.

As the optimal input does not depend on ϑ (see Theorem 2.1), a possible candidate is the Maximum Likelihood Estimator (MLE) $\hat{\vartheta}_T$, defined as the maximum of the likelihood:

$$\hat{\vartheta}_T = \arg\max_{\vartheta>0} \mathcal{L}(\vartheta, Y^T).$$

Moreover, MLE reaches efficiency and we deduce its large sample asymptotic properties:

Theorem 2.2. *The MLE is uniformly consistent on compacts* $K \subset \mathbb{R}_*^+$*, i.e. for any* $\nu > 0$*,*

$$\lim_{T\to\infty} \sup_{\vartheta\in\mathbb{K}} \mathbf{P}_\vartheta^T \left\{ \left| \hat{\vartheta}_T - \vartheta \right| > \nu \right\} = 0\,,$$

uniformly on compacts asymptotically normal: as T *tends to* $+\infty$*,*

$$\lim_{T\to\infty} \sup_{\vartheta\in\mathbb{K}} \left| \mathbf{E}_\vartheta f\left(\sqrt{T}\left(\hat{\vartheta}_T - \vartheta\right)\right) - \mathbf{E} f(\xi) \right| = 0 \quad \forall f \in \mathcal{C}_b$$

and ξ *is a zero mean Gaussian random variable of variance* $\mathcal{I}(\vartheta)^{-1}$ *(see* (2.5) *for the explicit value)* **which does not depend on** H *and we have the uniform on* $\vartheta \in \mathbb{K}$ *convergence of the moments: for any* $p > 0$*,*

$$\lim_{T\to\infty} \sup_{\vartheta\in\mathbb{K}} \left| \mathbf{E}_\vartheta \left| \sqrt{T}\left(\hat{\vartheta}_T - \vartheta\right) \right|^p - \mathbf{E}\left|\xi\right|^p \right| = 0.$$

Finally, the MLE is efficient in the sense of (2.3).

Remark 2.2. *The MLE satisfies all the properties in Theorem 2.2 with the same $\mathcal{I}(\vartheta)$ when $H = \frac{1}{2}$. To the best of our knowledge, the result is also new in this case but with the same method we will present .*

Because the proof for the two cases is same, so in the following parts, we will only deal with the partially observable case and give the explanation of the fully observable case in section 4

2 Preliminaries

2.1 Transformation of the model

The explicit representation of the likelihood function can be written thanks to the transformation of observation model proposed in [38]. In what follows, all random variables and processes are defined on a given stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ satisfying the usual conditions and processes are $(\mathcal{F}_t)-$ adapted. Moreover the *natural filtration* of a process is understood as the $\mathbf{P}$-completion of the filtration generated by this process. Let us define for $H > \frac{1}{2}$ (for the case $H < \frac{1}{2}$, see Section 5.1):

$$k_H(t,s) = \kappa^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H},\ w_H(t) = \frac{1}{2\lambda(2-2H)} t^{2-2H} \tag{2.6}$$

$$N_t = \int_0^t k_H(t,s) dW_t^H \quad \text{and} \quad M_t = \int_0^t k_H(t,s) dV_t^H \tag{2.7}$$

where κ and λ are defined in (2.4).

Then the process $N = (N_t, t \geq 0)$ is a Gaussian martingale, called in [56] the *fundamental martingale*, whose variance function is nothing but w_H. Moreover, the natural filtration of the martingale N coincides with the natural filtration of the fBm W^H. Similarly $M = (M_t, t \geq 0)$ stands for the fundamental martingale of V^H.

Following [38], let us introduce a process $Z = (Z_t, 0 \leq t \leq T)$ the fundamental semimartingale associated to Y, defined as

$$Z_t = \int_0^t k_H(t,s) dY_s.$$

Note that Y can be represented as $Y_t = \int_0^t K_H(t,s) dZ_s$, where $K_H(t,s) = H(2H-1)\int_s^t r^{H-\frac{1}{2}}(r-s)^{H-\frac{3}{2}} dr$ for $0 \leq s \leq t$ and therefore the natural filtrations of Y and Z coincide. Moreover, we have the following representation:

$$dZ_t = \mu\lambda \ell(t)^* \zeta_t d\langle N\rangle_t + dN_t,\ Z_0 = 0, \tag{2.8}$$

where $\zeta = (\zeta_t,\ t \geq 0)$ is the solution of the stochastic differential equation

$$d\zeta_t = -\lambda\vartheta \mathbf{A}(t)\zeta_t d\langle M\rangle_t + b(t)v(t) d\langle M\rangle_t + b(t) dM_t, \quad \zeta_0 = 0, \tag{2.9}$$

with

$$\ell(t) = \begin{pmatrix} t^{2H-1} \\ 1 \end{pmatrix},\ \mathbf{A}(t) = \begin{pmatrix} t^{2H-1} & 1 \\ t^{4H-2} & t^{2H-1} \end{pmatrix},\ b(t) = \begin{pmatrix} 1 \\ t^{2H-1} \end{pmatrix} \tag{2.10}$$

and $*$ standing for the transposition. Here, for a control $u(t)$, we have defined the function $v(t)$ by the following equation

$$v(t) = \frac{d}{dw_H(t)} \int_0^t k_H(t,s)u(s)ds; \tag{2.11}$$

provided that the fractional derivative exists. Let us define the space of control for $v(t)$:

$$\mathcal{V}_T = \left\{ v \,\middle|\, \frac{1}{T}\int_0^T |v(t)|^2 dw_H(t) \leq 1 \right\}. \tag{2.12}$$

Remark that with (2.11) the following relation between control u and its transformation v holds:

$$u(t) = \frac{d}{dt}\int_0^t K_H(t,s)v(s)dw_H(s). \tag{2.13}$$

We can set the admissible controls as $\mathcal{U}_T = \{u \,|\, v \in \mathcal{V}_T\}$. Note that these sets are non empty.

2.2 Likelihood function and the Fisher information

The classical Girsanov theorem and the general filtering theorem (see [49]) gives

$$\mathcal{L}(\vartheta, Z^T) = \exp\left(\mu\lambda \int_0^T \ell(t)^*\pi_t(\zeta)dZ_t - \frac{\mu^2\lambda^2}{2}\int_0^T \pi_t(\zeta)\ell(t)\ell(t)^*\pi_t(\zeta)^* d\langle N\rangle_t\right)$$

where the conditional expectation $\pi_t(\zeta) = \mathbf{E}_\vartheta(\zeta_t|\mathcal{F}_t^Y)$ satisfies the equation

$$d\pi_t(\zeta) = a_\vartheta(t)\pi_t(\zeta)d\langle M\rangle_t + \mu\lambda\gamma(t)\ell(t)dZ_t + b(t)v(t)d\langle M\rangle_t, \quad \pi_0(\zeta) = 0, \tag{2.14}$$

Here $a_\vartheta(t) = -\vartheta\lambda\mathbf{A}(t) - \mu^2\lambda^2\gamma(t)\ell(t)\ell(t)^*$ and $\gamma(t) = \mathbf{E}_\vartheta(\zeta_t - \pi_t(\zeta))^*(\zeta_t - \pi_t(\zeta))$ is the covariance of the filtering error, which is the unique solution of the Ricatti equation

$$\frac{d\gamma(t)}{d\langle M\rangle_t} = -\vartheta\lambda(\mathbf{A}(t)\gamma(t) + \gamma(t)\mathbf{A}(t)^*) + b(t)b(t)^* - \mu^2\lambda^2\gamma(t)\ell(t)\ell(t)^*\gamma(t) \tag{2.15}$$

with initial condition $\gamma(0) = 0$. Note that Equation (2.14) can be rewritten in the equivalent form

$$d\pi_t(\zeta) = -\vartheta\lambda\mathbf{A}(t)\pi_t(\zeta)d\langle M\rangle_t + b(t)v(t)d\langle M\rangle_t + \mu\lambda\gamma(t)\ell(t)d\nu_t \tag{2.16}$$

with initial condition $\pi_0(\zeta) = 0$ and where the innovation process $(\nu_t,\, t \geq 0)$ is defined by:

$$d\nu_t = dZ_t - \mu\lambda\ell(t)^*\pi_t(\zeta)d\langle N\rangle_t, \quad \nu_0 = 0. \tag{2.17}$$

The Fisher information stands for

$$\begin{aligned}\mathcal{I}_T(\vartheta, v) &= -\mathbf{E}_\vartheta \frac{\partial^2}{\partial\vartheta^2}\ln\mathcal{L}(\vartheta, Z^T) \\ &= \mathbf{E}_\vartheta\mu^2\lambda^2\int_0^T \frac{\partial\pi_t(\zeta)}{\partial\vartheta}\ell(t)\ell(t)^*\frac{\partial\pi_t(\zeta)}{\partial\vartheta}^* d\langle N\rangle_t.\end{aligned}$$

2.3 Fisher information decomposition

Contrary to what have been done in [12], it is hard to compute directly the Fisher information using $\frac{\partial \pi_t(\zeta)}{\partial \vartheta}$ in its implicit form. Let us introduce

$$\Pi_t^\vartheta = \begin{pmatrix} \pi_t(\zeta) \\ \frac{\partial \pi_t(\zeta)}{\partial \vartheta} \end{pmatrix} \quad \text{and} \quad \mathcal{C}(t)^* = \begin{pmatrix} \mathbf{0}_{1\times 2} & \ell(t)^* \end{pmatrix} \tag{2.18}$$

where $\mathbf{0}_{1\times 2}$ is the zero matrix of size 1×2. The Fisher information can be rewritten as

$$\mathcal{I}_T(\vartheta, v) = \mu^2 \lambda^2 \mathbf{E}_\vartheta \int_0^T \left(\mathcal{C}(t)^* \Pi_t^\vartheta\right)^2 d\langle N \rangle_t.$$

We will separate the Fisher information with two parts, one with the control, the other without. So we will focus on the following decomposition

$$\begin{aligned} \mathcal{I}_T(\vartheta, v) &= \mu^2 \lambda^2 \int_0^T \mathbf{E}_\vartheta \left(\mathcal{C}(t)^* (\Pi_t^\vartheta - \mathbf{E}_\vartheta \Pi_t^\vartheta + \mathbf{E}_\vartheta \Pi_t^\vartheta)\right)^2 d\langle N \rangle_t \\ &= \mathcal{I}_{1,T}(\vartheta, v) + \mathcal{I}_{2,T}(\vartheta, v) \end{aligned} \tag{2.19}$$

where

$$\mathcal{I}_{1,T}(\vartheta, v) = \mu^2 \lambda^2 \int_0^T \mathbf{E}_\vartheta \left(\mathcal{C}(t)^* (\Pi_t^\vartheta - \mathbf{E}_\vartheta \Pi_t^\vartheta)\right)^2 d\langle N \rangle_t$$

and

$$\mathcal{I}_{2,T}(\vartheta, v) = \mu^2 \lambda^2 \int_0^T \left(\mathcal{C}(t)^* \mathbf{E}_\vartheta \Pi_t^\vartheta\right)^2 d\langle N \rangle_t. \tag{2.20}$$

The deterministic function $\left(\mathcal{P}^\vartheta(t) = \mathbf{E}_\vartheta \Pi_t^\vartheta, t \geq 0\right)$ satisfies the following equation:

$$\frac{d\mathcal{P}^\vartheta(t)}{d\langle N \rangle_t} = \mathcal{A}^\vartheta(t) \mathcal{P}^\vartheta(t) + \mathcal{D}(t) v(t), \quad \mathcal{P}(0) = \mathbf{0}_{4\times 1}, \tag{2.21}$$

where

$$\mathcal{D}(t) = \begin{pmatrix} b(t) \\ \mathbf{0}_{2\times 1} \end{pmatrix} \quad \text{and} \quad \mathcal{A}^\vartheta(t) = \begin{pmatrix} -\lambda\vartheta \mathbf{A}(t) & \mathbf{0}_{2\times 2} \\ -\lambda \mathbf{A}(t) & a_\vartheta(t) \end{pmatrix}.$$

At the same time, the process $\overline{\mathcal{P}}_t = \left(\Pi_t^\vartheta - \mathcal{P}^\vartheta(t), t \geq 0\right)$ satisfies the following equation:

$$d\overline{\mathcal{P}}_t = \mathcal{A}^\vartheta(t) \overline{\mathcal{P}}_t \, d\langle N \rangle_t + \mu\lambda\gamma(t) \begin{pmatrix} \ell(t) \\ \ell(t) \end{pmatrix} d\nu_t \tag{2.22}$$

with initial condition $\overline{\mathcal{P}}_0 = \mathbf{0}_{4\times 1}$. Since $\mathcal{I}_{1,T}(\vartheta, v)$ does not depend on $v(t)$ (see equation (2.22)), we write it $\mathcal{I}_{1,T}(\vartheta)$. In fact, $\mathcal{I}_{1,T}(\vartheta)$ is the Fisher information of the initial system (2.2) when $u = 0$.

3 Proofs

3.1 Proof of Theorem 2.1

With the technique of separation (2.19) and the precedent remarks, we have

$$\mathcal{J}_T(\vartheta) = \mathcal{I}_{1,T}(\vartheta) + \mathcal{J}_{2,T}(\vartheta)$$

where

$$\mathcal{J}_{2,T}(\vartheta) = \sup_{v \in \mathcal{V}_T} \mathcal{I}_{2,T}(\vartheta, v).$$

From (2.21), we get

$$\mathcal{P}(t) = \varphi_\vartheta(t) \int_0^t \varphi_\vartheta^{-1}(s)\mathcal{D}(t)v(s)d\langle N\rangle_s \tag{2.23}$$

where $\varphi_\vartheta(t)$ is the matrix defined by

$$\frac{d\varphi_\vartheta(t)}{d\langle N\rangle_t} = \mathcal{A}^\vartheta(t)\varphi_\vartheta(t), \;\; \varphi_\vartheta(0) = \mathbf{Id}_{4\times 4} \tag{2.24}$$

with $\mathbf{Id}_{4\times 4}$ the 4×4 identity matrix. Substituting in (2.20), we get

$$\mathcal{I}_{2,T}(\vartheta, v) = \int_0^T \int_0^T \mathcal{K}_T(s,\sigma)\frac{s^{\frac{1}{2}-H}}{\sqrt{2\lambda}}v(s)\frac{\sigma^{\frac{1}{2}-H}}{\sqrt{2\lambda}}v(\sigma)dsd\sigma,$$

where

$$\mathcal{K}_T(s,\sigma) = \int_{\max(s,\sigma)}^T \mathcal{G}(t,s)\mathcal{G}(t,\sigma)dt,$$

and

$$\mathcal{G}(t,\sigma) = \frac{\mu^2}{2}\left(t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_\vartheta(t)\varphi_\vartheta^{-1}(\sigma)\mathcal{D}(\sigma)\sigma^{\frac{1}{2}-H}\right).$$

Then

$$\begin{aligned}
\mathcal{J}_{2,T}(\vartheta) &= T \sup_{\widetilde{v}\in L^2[0,T], \|\widetilde{v}\|\le 1} \int_0^T \int_0^T \mathcal{K}_T(s,\sigma)\widetilde{v}(s)\widetilde{v}(\sigma)dsd\sigma, \\
&= T \sup_{\widetilde{v}\in L^2[0,T], \|\widetilde{v}\|\le 1} (\mathcal{K}_T\widetilde{v}, \widetilde{v})
\end{aligned} \tag{2.25}$$

where $\tilde{v}(s) = \dfrac{s^{\frac{1}{2}-H}}{\sqrt{2\lambda}}\dfrac{v(s)}{\sqrt{T}}$ and $||.||$ stands for the usual norm in $L^2[0,T]$.
On one hand, we get from [10] that

$$\lim_{T\to\infty}\frac{\mathcal{I}_{1,T}(\vartheta)}{T} = \frac{1}{2\vartheta} - \frac{2\vartheta}{\alpha(\alpha+\vartheta)} + \frac{\vartheta^2}{2\alpha^3}$$

where $\alpha = \sqrt{\mu^2+\vartheta^2}$.
On the other hand, in order to prove the Theorem 2.1, we have to check that

$$\lim_{T\to\infty}\frac{\mathcal{J}_{2,T}(\vartheta)}{T} = \frac{\mu^2}{\alpha^2\vartheta^2}$$

or, equivalently, looking at equation (2.25) that

Lemma 2.1.

$$\lim_{T\to\infty} \sup_{\widetilde{v}\in L^2[0,T], \|\widetilde{v}\|\le 1} (\mathcal{K}_T\widetilde{v}, \widetilde{v}) = \frac{\mu^2}{\alpha^2\vartheta^2}$$

with an optimal input $v_{opt}(t) = \sqrt{2\lambda}t^{H-\frac{1}{2}}$ *belonging to the space of control* $\mathcal{V}_T$*.*

Proof. The proof is adapted from [12] and is postponed to section 5.2. □

3.2 Proof of Theorem 2.2

As the optimal input does not depend on ϑ, the MLE $\widehat{\vartheta}_T$ of ϑ in the system

$$\begin{cases} d\zeta_t &= -\lambda\vartheta\mathbf{A}(t)d\langle M\rangle_t + b(t)v_{opt}(t)d\langle M\rangle_t + b(t)dM_t\,, \quad \zeta_0 = 0, \\ dZ_t &= \mu\lambda\ell(t)\zeta_t + dN_t\,, \quad Z_0 = 0, \end{cases} \tag{2.26}$$

is a good candidate to reach efficiency in (2.3).

Ibragimov-Khasminskii program

The proof of Proposition 2.2 is base on [31, Theorem I.10.1]. Let us define the likelihood ratio

$$\mathcal{Z}_\vartheta^T(r) = \frac{\mathcal{L}(\vartheta + \frac{r}{\sqrt{T}}, Z^T)}{\mathcal{L}(\vartheta, Z^T)}, \quad r \in \mathcal{S}_\vartheta^T = \left\{ r \,:\, \vartheta + \frac{r}{\sqrt{T}} \in \mathbb{R}_*^+ \right\}.$$

Actually, to prove Proposition 2.2, it is sufficient to check the three following conditions on the likelihood ratio. For any compacts $\mathbb{K} \subset \mathbb{R}_*^+$,

(A.1) Let $\mathcal{Z}_\vartheta(r) = \exp\left(r\xi - \frac{u^2}{2}\mathcal{I}(\vartheta) \right)$ with $\xi \sim \mathcal{N}(0, \mathcal{I}(\vartheta))$, whose maximum is attained at the unique point $\hat{r} = \xi\,\mathcal{I}(\vartheta)^{-1}$, where $\mathcal{I}(\vartheta)$ is defined by (2.5). Uniformly in $\vartheta \in \mathbb{K}$, the marginal (finite-dimensional) distributions of the random function $\mathcal{Z}_\vartheta^T(r)$ converge to the marginal distributions of the random function $\mathcal{Z}_\vartheta(r)$.

(A.2) There exist $\chi > 0$ such that for all $r \in \mathcal{S}_\vartheta^T$,

$$\sup_{\vartheta \in \mathbb{K}} \mathbf{E}_\vartheta \mathcal{Z}_\vartheta^T(r)^{\frac{1}{2}} \leq \exp\left(-\chi r^2\right).$$

(A.3) There exist $a > 0$ and $b > 0$ (depending on $\mathbb{K}$) such that for any $R > 0$, for $|r_1| < R$, $|r_2| < R$

$$\sup_{\vartheta \in \mathbb{K}} \mathbf{E}_\vartheta \left| \mathcal{Z}_\vartheta^T(r_2)^{\frac{1}{2}} - \mathcal{Z}_\vartheta^T(r_1)^{\frac{1}{2}} \right|^2 \leq b(1 + R^a)|r_2 - r_1|^2.$$

Laplace transform proof

For any $\vartheta_1 > 0$, $\left(\pi_t^{\vartheta_1}(\zeta), t \geq 0\right)$ is the solution of (2.16) and $(\gamma_{\vartheta_1}(t), t \geq 0)$ the solution of (2.15), both when $\vartheta = \vartheta_1$. Let us denote $\delta_{\vartheta_1,\vartheta_2}(t) = \pi_t^{\vartheta_2}(\zeta) - \pi_t^{\vartheta_1}(\zeta)$ and $L_T(a, \vartheta_1, \vartheta_2)$ the Laplace transform of the integral of the quadratic form of the difference $\delta_{\vartheta_1,\vartheta_2}(t)$:

$$L_T(a, \vartheta_1, \vartheta_2) = \mathbf{E}_\vartheta \exp\left(-a\frac{\mu^2\lambda^2}{2} \int_0^T \delta_{\vartheta_1,\vartheta_2}(t)^* \ell(t)\ell(t)^* \delta_{\vartheta_1,\vartheta_2}(t) d\langle N\rangle_t \right).$$

It has been proved in [10] that, if there exists $a_0 < 0$ such that for all $a > a_0$, $\forall$ r_1, $r_2 \in \mathcal{S}_\vartheta^T$,

$$\lim_{T\to\infty} L_T(a, \vartheta + \frac{r_1}{\sqrt{T}}, \vartheta + \frac{r_2}{\sqrt{T}}) = \exp(-a\frac{(r_2 - r_1)^2}{2}\mathcal{I}(\vartheta)), \tag{2.27}$$

then properties (A.1-A.3) of the Ibragimov-Khasminskii program hold.

In the following, we compute the behavior of this Laplace transform. Let us define

$$\Pi_t = \begin{pmatrix} \pi_t^{\vartheta_1}(\zeta) \\ \delta_{\vartheta_1,\vartheta_2}(t) \end{pmatrix}$$

which is governed by:

$$d\Pi_t = \mathcal{A}(t)\Pi_t d\langle N\rangle_t + \mathcal{B}(t)d\nu_t^{\vartheta_1} + \mathcal{D}(t)v_{opt}(t)d\langle N\rangle_t, \quad \Pi_0 = \mathbf{0}_{4\times 1}, \tag{2.28}$$

where

$$a_{\vartheta_2}(t) = -\vartheta_2\lambda\mathbf{A}(t) - \mu^2\lambda^2\gamma_{\vartheta_2}(t)\ell(t)\ell(t)^*, \quad \mathbf{D}_\gamma^{\vartheta_1,\vartheta_2} = \gamma_{\vartheta_2}(t) - \gamma_{\vartheta_1}(t),$$

$$\mathcal{A}(t) = \begin{pmatrix} -\vartheta_1\lambda\mathbf{A}(t) & 0 \\ -(\vartheta_2-\vartheta_1)\lambda\mathbf{A}(t) & a_{\vartheta_2}(t) \end{pmatrix} \quad \text{and} \quad \mathcal{B}(t) = \mu\lambda \begin{pmatrix} \gamma_{\vartheta_1}(t) \\ \mathbf{D}_\gamma^{\vartheta_1,\vartheta_2} \end{pmatrix} \ell(t).$$

Following from [36]

$$\begin{aligned} L_T(a,\vartheta_1,\vartheta_2) &= \mathbf{E}_{\vartheta_1}\exp\left(-a\frac{\mu^2\lambda^2}{2}\int_0^T \delta_{\vartheta_1,\vartheta_2}(t)^*\ell(t)\ell(t)^*\delta_{\vartheta_1,\vartheta_2}(t)d\langle N\rangle_t\right) \\ &= \mathbf{E}_{\vartheta_1}\exp\left(-a\frac{\mu^2\lambda^2}{2}\int_0^T \Pi_t^*\mathcal{C}(t)\mathcal{C}(t)^*\Pi_t d\langle N\rangle_t\right) \\ &= \exp\left(-a\frac{\mu^2\lambda^2}{2}\int_0^T \left(\operatorname{trace}\left(\mathcal{H}(t)\mathcal{M}(t)\right) + Q(t)^*\mathcal{M}(t)Q(t)\right)d\langle N\rangle_t\right) \\ &= L_{1,T}(a,\vartheta_1,\vartheta_2)L_{2,T}(a,\vartheta_1,\vartheta_2) \end{aligned} \tag{2.29}$$

where

$$L_{1,T}(a,\vartheta_1,\vartheta_2) = \exp\left(-a\frac{\mu^2\lambda^2}{2}\int_0^T \operatorname{trace}\left(\mathcal{H}(t)\mathcal{M}(t)\right)d\langle N\rangle_t\right)$$

and

$$L_{2,T}(a,\vartheta_1,\vartheta_2) = \exp\left(-a\frac{\mu^2\lambda^2}{2}\int_0^T Q(t)^*\mathcal{M}(t)Q(t)d\langle N\rangle_t\right). \tag{2.30}$$

Here, $\mathcal{H}(t)$ satisfies the Ricatti equation

$$\frac{d\mathcal{H}(t)}{d\langle N\rangle_t} = \mathcal{A}(t)\mathcal{H}(t) + \mathcal{H}(t)\mathcal{A}(t)^* + \mathcal{B}(t)\mathcal{B}(t)^* - a\lambda^2\mu^2\mathcal{H}(t)\mathcal{M}(t)\mathcal{H}(t), \tag{2.31}$$

with

$$\mathcal{M}(t) = \mathcal{C}(t)\mathcal{C}(t)^*,$$

$\mathcal{C}(t)$ defined in (2.18) and $(Q(t), t\geq 0)$ satisfying the equation:

$$Q(t) = \mathbf{E}_{\vartheta_1}\Pi_t - a\lambda^2\mu^2\int_0^t \varphi(t)\varphi^{-1}(s)\mathcal{H}(s)\mathcal{M}(s)Q(s)d\langle N\rangle_s \tag{2.32}$$

where

$$\frac{d\varphi(t)}{d\langle N\rangle_t} = \mathcal{A}(t)\varphi(t), \quad \varphi(0) = \mathbf{Id}_{4\times 4}. \tag{2.33}$$

Since $L_{1,T}(a,\vartheta_1,\vartheta_2)$ has been studied in [10], checking the condition (2.27) (and therefore proving the Proposition 2.2) is only but to prove the following lemma:

Lemma 2.2. *With the previous notations, there exists* $a_0 < 0$ *such that for all* $a > a_0$, $\forall\ r_1,\ r_2 \in \mathcal{S}_\vartheta^T$,

$$\lim_{T\to\infty} L_{2,T}\left(a, \vartheta + \frac{r_1}{\sqrt{T}}, \vartheta + \frac{r_2}{\sqrt{T}}\right) = \exp\left(-a\frac{(r_2-r_1)^2}{2}\frac{\mu^2}{\alpha^2\vartheta^2}\right)$$

where $\alpha = \sqrt{\mu^2+\vartheta^2}$.

It is worth emphasizing, that using this technique of separation, the above lemma does not appear neither in [10] nor in [12].

4 Fully observable case

4.1 Optimal input

Let $H > \frac{1}{2}$ (for $H < \frac{1}{2}$, see Section 5.1) and let us denote by $Z = (Z_t,\ t \geq 0)$ the fundamental semimartingale associated to X, defined as

$$Z_t = \int_0^t k_H(t,s)dX_s$$

where $k_H(t,s)$ is defined in (2.6). Thanks to the this transformation, we can write the explicit expression of the likelihood function

$$\mathcal{L}_T(\vartheta, Z^T) = \exp\left(\int_0^T (-\vartheta\lambda\ell(t)^*\zeta_t + v(t))dZ_t - \frac{1}{2}\int_0^T (-\vartheta\lambda\ell(t)^*\zeta_t + v(t))^2 d\langle M\rangle_t\right)$$

where $\ell(t)$, λ and M_t are defined in (2.6),(2.7) and (2.10) respectively, $(\zeta_t, t\geq 0)$ is the solution of (2.9) and $(v(t), t \geq 0)$ is defined in (2.11). We can decompose the Fisher information with the same technique into the following form:

$$\begin{aligned}
\mathcal{I}_T(\vartheta, v) &= \mathbf{E}_\vartheta \int_0^T (\lambda\ell(t)^*\zeta_t)^2 d\langle N\rangle_t,\\
&= \mathbf{E}_\vartheta \int_0^T \left(\ell(t)^*(\zeta_t - \mathbf{E}_\vartheta\zeta_t + \mathbf{E}_\vartheta\zeta_t)\right)^2 d\langle N\rangle_t,\\
&= \mathbf{E}_\vartheta\lambda^2 \int_0^T \left(\ell(t)^*(\zeta_t - \mathbf{E}_\vartheta\zeta_t)\right)^2 d\langle M\rangle_t + \lambda^2\int_0^T \ell(t)^*(\mathbf{E}_\vartheta\zeta_t)^2 d\langle N\rangle_t,\\
&= \mathcal{I}_1(\vartheta, v) + \mathcal{I}_2(\vartheta, v).
\end{aligned}$$

The mean function $(\mathbf{E}_\vartheta\zeta_t, t \geq 0)$ satisfies

$$d\mathbf{E}_\vartheta\zeta_t = -\vartheta\lambda\mathbf{A}(t)\mathbf{E}_\vartheta\zeta_t d\langle N\rangle_t + b(t)v(t)d\langle N\rangle_t, \quad \mathbf{E}_\vartheta\zeta_0 = \mathbf{0}_{2\times 1}$$

and the process $(\zeta_t - \mathbf{E}_\vartheta\zeta_t, t \geq 0)$ satisfies

$$d\left(\zeta_t - \mathbf{E}_\vartheta\zeta_t\right) = -\vartheta\lambda\mathbf{A}(t)(\zeta_t - \mathbf{E}_\vartheta\zeta_t)d\langle N\rangle_t + b(t)dN_t, \quad \zeta_0 - \mathbf{E}_\vartheta\zeta_0 = \mathbf{0}_{2\times 1}. \quad (2.34)$$

We can see in (2.34) that $\mathcal{I}_1(\vartheta, v)$ does not depend on $v(t)$ and can noted $\mathcal{I}_1(\vartheta)$. Actually, it has been studied in [10] and

$$\lim_{T\to\infty}\frac{\mathcal{I}_1(\vartheta)}{T} = \frac{1}{2\vartheta}.$$

Then, for the optimal control computation, we will only consider $\mathcal{I}_2(\vartheta, v)$. Same kind of computations as in the partially observable case (explicit form of the kernel and maximization) leads to

$$\lim_{T\to\infty} \frac{\mathcal{I}_2(\vartheta, v_{opt})}{T} = \frac{1}{\vartheta^2}$$

where $v_{opt}(t) = \sqrt{2\lambda} t^{H-\frac{1}{2}}$.

4.2 Properties of Estimator

We have know that the optimal input does not depend on ϑ, we can get the explicit expression of MLE, that is

$$\widehat{\vartheta}_T = \vartheta - \frac{\int_0^T \lambda \ell(t)^* \zeta_t^o dN_t}{\int_0^T (\lambda \ell(t)^* \zeta_t^o)^2 d\langle N\rangle_t}, \tag{2.35}$$

where ζ_t^o is the solution of (2.9) with the input $v(t) = v_{opt}(t)$, $t \geq 0$.

To prove the asymptotical properties in Theorem 2.2, we will compute the Laplace transform of the denominator in (2.35):

$$L_T(a, \vartheta) = \mathbf{E}_\vartheta \exp\left(-\frac{a}{2}\int_0^T (\lambda \ell(t)^* \zeta_t^o)^2 d\langle N\rangle_t\right). \tag{2.36}$$

Following from [36], we can have

$$\begin{aligned} L_T(a, \vartheta) &= \exp\left(-\frac{1}{2}\int_0^T Q(t)^* \mathcal{M}(t) Q(t) dt + \operatorname{trace}(\gamma(t)\mathcal{M}(t))dt\right) \\ &= L_{1,T}(a, \vartheta) L_{2,T}(a, \vartheta) \end{aligned}$$

where

$$L_{1,T}(a, \vartheta) = \exp\left(-\frac{1}{2}\int_0^T \operatorname{trace}(\gamma(t)\mathcal{M}(t))dt\right)$$

and

$$L_{2,T}(a, \vartheta) = \exp\left(-\frac{1}{2}\int_0^T Q(t)^* \mathcal{M}(t) Q(t) dt\right).$$

In the previous equations, we denoted

$$\mathcal{M}(t) = \frac{a\lambda}{2}\ell(t)\ell(t)^* t^{1-2H} \quad \text{and} \quad Q(t) = \mathbf{E}_\vartheta \zeta_t^o - \int_0^t \gamma(t,s)\mathcal{M}(s)Q(s)ds.$$

Moreover, we defined $\gamma(t,s) = \Pi_t \Pi_s^{-1}\gamma(s)$ where Π_t satisfies the differential equation

$$d\Pi_t = -\frac{\vartheta}{2}\mathbf{A}_H(t)\Pi_t dt$$

and $\gamma(t)$ satisfies the following Ricatti equation:

$$\frac{d\gamma(t)}{dt} = -\frac{\vartheta}{2}\left(\mathbf{A}_H(t)\gamma(t) + \gamma(t)\mathbf{A}_H(t)^*\right) - \gamma(t)\mathcal{M}(t)\gamma(t) + \frac{1}{2\lambda}b(t)b(t)^* t^{1-2H}.$$

Asymptotical behavior of the first term $L_{1,T}(a, \vartheta)$ have been studied in [10]. Direct computations for the second term $L_{2,T}(a, \vartheta)$ leads to

$$\lim_{T\to\infty} L_T(\frac{a}{T}, \vartheta) = \exp\left(-\frac{a}{2}\left(\frac{1}{2\vartheta} + \frac{1}{\vartheta^2}\right)\right)$$

5 Technical proofs of Lemmas

5.1 From $H > \frac{1}{2}$ to $H < \frac{1}{2}$

Thanks to [33, Corollary 5.2], for $H < 1/2$, we have the relation between fBm processes of indexes H and $1-H$:

$$W_t^H = \aleph_H \int_0^t (t-s)^{2H-1}\, dW_s^{1-H}\,, \text{ with } \aleph_H = \left(\frac{2H}{\Gamma(2H)\Gamma(3-2H)}\right)^{\frac{1}{2}}. \qquad (2.37)$$

Using this relation, we can transform the observation model (2.2) to the following observation model:

$$\begin{cases} d\tilde{X}_t &= -\vartheta\tilde{X}_t dt + \tilde{u}(t)dt + dV_t^{1-H}\,, \quad \tilde{X}_0 = 0\,, \\ d\tilde{Y}_t &= \mu\tilde{X}_t dt + dW_t^{1-H}\,, \quad \tilde{Y}_0 = 0\,, \end{cases}$$

with

$$\tilde{X}_t = \aleph_{1-H}\int_0^t (t-s)^{1-2H}\, dX_s, \qquad \tilde{Y}_t = \aleph_{1-H}\int_0^t (t-s)^{1-2H}\, dY_s\,,$$

and

$$\tilde{u}(t) = \aleph_{1-H}\frac{d}{dt}\int_0^t (t-r)^{1-2H}u(r)dr = (1-2H)\aleph_{1-H}\int_0^t (t-r)^{-2H}u(r)dr.$$

It had been proved in [11] that the set of controls $\mathcal{U}_T$ (see (2.12) for the definition) remains unchanged after transformation (2.37). Then $1-H > \frac{1}{2}$ and the results of Proposition 2.1 and Proposition 2.2 are valid for any $H \in (0,1)$.

5.2 Proof of Lemma 2.1

First of all, we need a preliminary result. Let us denote

$$p(t) = \begin{pmatrix} t^{\frac{1}{2}-H} \\ t^{H-\frac{1}{2}} \\ 0 \\ 0 \end{pmatrix}$$

and recall that $\varphi_\vartheta(t)$ is the solution of equation (2.24) and $\alpha = \sqrt{\mu^2+\vartheta^2}$.

Lemma 2.3. *With the previous notations,*

$$\lim_{t\to\infty}\int_0^t t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_\vartheta(t)\varphi_\vartheta^{-1}(s)p(s)ds = -\frac{2}{\alpha\vartheta}. \qquad (2.38)$$

Proof. Due to the asymptotical behavior of $\varphi_\vartheta(t)$ as $t\to\infty$, we can plug it into the computation of the limit of the integral of (2.24). Using the asymptotical behavior of $\gamma_\vartheta(t)$ (see (2.15)), we get from [10] that $a_\vartheta(t) \underset{t\to\infty}{\sim} -\alpha\lambda\mathbf{A}(t)$. Therefore, the standard arguments (see, *e.g.*, [40]) imply that

$$\lim_{t\to\infty}\int_0^t t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_\vartheta(t)\varphi_\vartheta^{-1}(s)p(s)ds = \lim_{t\to\infty}\int_0^t t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_{\vartheta,\infty}(t)\varphi_{\vartheta,\infty}^{-1}(s)p(s)ds$$

where
$$\frac{d\varphi_{\vartheta,\infty}(t)}{dt} = \left(M \otimes \mathbf{A}_H(t)\right)\varphi_{\vartheta,\infty}(t), \quad \varphi_{\vartheta,\infty}(0) = \mathbf{Id}_{4\times 4},$$
$\otimes$ denotes the Kronecker product,
$$M = \begin{pmatrix} -\frac{\vartheta}{2} & 0 \\ -\frac{1}{2} & -\frac{\alpha}{2} \end{pmatrix} \quad \text{and} \quad \mathbf{A}_H(t) = \begin{pmatrix} 1 & t^{1-2H} \\ t^{2H-1} & 1 \end{pmatrix}.$$
Let us diagonal the matrix M
$$M = G\begin{pmatrix} -\frac{\vartheta}{2} & 0 \\ 0 & -\frac{\alpha}{2} \end{pmatrix}G^{-1} \quad ,G = \begin{pmatrix} \vartheta-\alpha & 0 \\ 1 & 1 \end{pmatrix},$$
when we define $\widetilde{\varphi}_{\vartheta,\infty}(t) = (G^{-1}\otimes \mathbf{Id_{2\times 2}})\varphi_{\vartheta,\infty}(t)$, it satisfies the following equation:
$$\frac{d\widetilde{\varphi}_{\vartheta,\infty}(t)}{dt} = \begin{pmatrix} -\frac{\vartheta}{2}\mathbf{A}_H(t) & \mathbf{0} \\ \mathbf{0} & -\frac{\alpha}{2}\mathbf{A}_H(t) \end{pmatrix}\widetilde{\varphi}_{\vartheta,\infty}(t)$$
with initial condition
$$\widetilde{\varphi}_{\vartheta,\infty}(0) = \frac{1}{\vartheta-\alpha}\begin{pmatrix} 1 & 0 \\ -1 & \vartheta-\alpha \end{pmatrix}\otimes \mathbf{Id}_{2\times 2}.$$
Then,
$$\int_0^t t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_{\vartheta,\infty}(t)\varphi^{-1}_{\vartheta,\infty}(s)p(s)ds = \int_0^t \frac{1}{\vartheta-\alpha}[g^{\vartheta}(t,s) - g^{\alpha}(t,s)]ds,$$
where $g^{\vartheta}(t,s) = t^{\frac{1}{2}-H}\ell(t)^*\rho_{\vartheta}(t)\rho_{\vartheta}^{-1}(s)b(s)s^{\frac{1}{2}-H}$ with
$$\frac{d\rho_{\vartheta}(t)}{dt} = -\frac{\vartheta}{2}\mathbf{A}_H(t)\rho_{\vartheta}(t),\ \rho_{\vartheta}(0) = \mathbf{Id}_{2\times 2}. \tag{2.39}$$
As $\int_0^t g^{\vartheta}(t,s)ds \to \frac{2}{\theta}$ when $t\to\infty$ (see [12]), which achieves the proof. □

Now let us return to the proof of the Proposition 2.1. With $v_{opt}(t) = \sqrt{2\lambda}t^{H-\frac{1}{2}}$, we can compute
$$\frac{\mathcal{I}_{2,T}(\vartheta, v_{opt})}{T} = \frac{\mu^2}{4T}\int_0^T \left(\int_0^t t^{\frac{1}{2}-H}\mathcal{C}(t)^*\varphi_{\vartheta}(t)\varphi_{\vartheta}^{-1}(s)p(s)ds\right)^2 dt \underset{T\to\infty}{\sim} \frac{\mu^2}{\alpha^2\vartheta^2}$$
and
$$\lim_{T\to\infty}\sup_{\widetilde{v}\in L^2[0,T],\|\widetilde{v}\|\le 1}(\mathcal{K}_T\widetilde{v},\widetilde{v}) \ge \frac{\mu^2}{\alpha^2\vartheta^2}.$$
To get the upper bound, let us introduce the Gaussian process $(\xi_t,\ 0\le t\le T)$
$$\xi_t = \left(\int_t^T \sigma^{\frac{1}{2}-H}\mathcal{C}(\sigma)^*\varphi_{\vartheta}(\sigma)\odot dW_{\sigma}\right)\varphi_{\vartheta}^{-1}(t)$$
where $(W_{\sigma}, \sigma\ge 0)$ is a Wiener process and $\odot$ denotes the Itô backward integral (see [65]). It is worth emphasizing that
$$\mathcal{K}_T(s,\sigma) = \frac{\mu^2}{4}\mathbf{E}\left(\xi_s\mathcal{D}(s)s^{\frac{1}{2}-H}\xi_{\sigma}\mathcal{D}(\sigma)\sigma^{\frac{1}{2}-H}\right) = \mathbf{E}(\mathcal{X}_{\sigma}\mathcal{X}_s).$$

where $\mathcal{X}$ is the centered Gaussian process defined by $\mathcal{X}_t = \frac{\mu}{2}\xi_t \mathcal{D}(t) t^{\frac{1}{2}-H}$. The process $(\xi_t, 0 \leq t \leq T)$ satisfies the following dynamic

$$-d\xi_t = \xi_t \mathcal{A}^\vartheta(t) d\langle M\rangle_t + \mathcal{C}(t)^* t^{\frac{1}{2}-H} \odot dW_t, \quad \xi_T = 0.$$

Obviously, $\mathcal{K}_T(s,\sigma)$ is a compact symmetric operator for fixed T, so we should estimate the spectral gap (the first eigenvalue $\nu_1(T)$) of the operator. The estimation of the spectral gap is based on the Laplace transform computation. Let us compute, for sufficiently small negative $a < 0$ the Laplace transform of $\int_0^T \mathcal{X}_t^2 dt$:

$$\begin{aligned} L_T(a) &= \mathbf{E}_\vartheta \exp\left(-a\int_0^T \mathcal{X}_t^2 dt\right) \\ &= \mathbf{E}_\vartheta \exp\left(-a\int_0^T \left(\frac{\mu}{2}\xi_t \mathcal{D}(t) t^{\frac{1}{2}-H}\right)^2 dt\right). \end{aligned}$$

On one hand, for $a > -\frac{1}{\nu_1(T)}$, since $\mathcal{X}$ is a centered Gaussian process with covariance operator $\mathcal{K}_T$, using Mercer's theorem and Parseval's inequality, $L_T(a)$ can be represented as :

$$L_T(a) = \prod_{i\geq 1}(1 + 2a\nu_i(T))^{-\frac{1}{2}}, \tag{2.40}$$

where $\nu_i(T)$, $i \geq 1$ is the sequence of positive eigenvalues of the covariance operator. On the other hand,

$$\begin{aligned} L_T(a) &= \mathbf{E}_\vartheta \left(-a\frac{\mu^2\lambda}{2}\int_0^T \xi_t \mathcal{D}(t)\mathcal{D}(t)^* \xi_t^* d\langle N\rangle_t\right) \\ &= \exp\left(\frac{1}{2}\int_0^T \mathrm{trace}(2\lambda \mathcal{H}^\vartheta(t)\mathcal{M}(t) d\langle N\rangle_t\right) \end{aligned}$$

where $\mathcal{H}^\vartheta(t)$ is the solution of Ricatti differential equation:

$$\frac{d\mathcal{H}^\vartheta(t)}{d\langle N\rangle_t} = \mathcal{H}^\vartheta(t)\mathcal{A}^\vartheta(t)^* + \mathcal{A}^\vartheta(t)\mathcal{H}^\vartheta(t) + 2\lambda \mathcal{H}^\vartheta(t)\mathcal{M}(t)\mathcal{H}^\vartheta(t) - a\mu^2\lambda \mathcal{D}(t)\mathcal{D}(t)^*$$

with initial condition $\mathcal{H}^\vartheta(0) = \mathbf{0}_{4\times 4}$, provided that the solution of this equation exists for any $0 \leq t \leq T$.

It is well know that if $\det \Psi_1(t) > 0$, for any $t \in [0,T]$, then $\mathcal{H}^\vartheta(t) = \Psi_1^{-1}(t)\Psi_2(t)$, where the pair of 4×4 matrices (Ψ_1, Ψ_2) satisfies the system of linear differential equations:

$$\frac{d\Psi_1(t)}{d\langle N\rangle_t} = -\Psi_1(t)\mathcal{A}^\vartheta(t) - 2\lambda\Psi_2(t)\mathcal{M}(t), \qquad \Psi_1(0) = \mathbf{Id}_{4\times 4},$$

$$\frac{d\Psi_2(t)}{d\langle N\rangle_t} = -a\mu^2\lambda\Psi_1(t)\mathcal{D}(t)\mathcal{D}(t)^* + \Psi_2(t)\mathcal{A}^\vartheta(t)^*, \quad \Psi_2(0) = \mathbf{0}_{4\times 4}$$

and

$$L_T(a) = \exp\left(-\frac{1}{2}\int_0^T \mathrm{trace}\left(\mathcal{A}^\vartheta(t)\right) d\langle N\rangle_t\right) (\det \Psi_1(T))^{-\frac{1}{2}}.$$

Here again, standard arguments (see [40]) imply that under the condition $\det \Psi_{1,\infty}(t) > 0$, for any $t \in [0,T]$,

$$L_T(a) \underset{T\to\infty}{\sim} \exp\left(-\frac{1}{2}\int_0^T \operatorname{trace}\left(\mathcal{A}_\infty^\vartheta(t)\right) d\langle N\rangle_t\right) \left(\det \Psi_{1,\infty}(T)\right)^{-\frac{1}{2}} \tag{2.41}$$

where

$$\mathcal{A}_\infty^\vartheta(t) = \begin{pmatrix} -\vartheta & 0 \\ -1 & -\alpha \end{pmatrix} \otimes \lambda \mathbf{A}(t)$$

and

$$\begin{aligned} \frac{d\Psi_{1,\infty}(t)}{d\langle N\rangle_t} &= -\Psi_{1,\infty}(t)\mathcal{A}_\infty^\vartheta(t) - 2\lambda\Psi_{2,\infty}(t)\mathcal{M}(t), \\ \frac{d\Psi_{2,\infty}(t)}{d\langle N\rangle_t} &= -a\mu^2\lambda\Psi_{1,\infty}(t)\mathcal{D}(t)\mathcal{D}(t)^* + \Psi_{2,\infty}(t)\mathcal{A}_\infty^\vartheta(t)^*, \end{aligned} \tag{2.42}$$

with initial conditions $\Psi_{1,\infty}(0) = \mathbf{Id}_{4\times 4}$ and $\Psi_{2,\infty}(0) = \mathbf{0}_{4\times 4}$. Rewriting the system (2.42) in the following form

$$\frac{d(\Psi_{1,\infty}(t),\ \Psi_{2,\infty}(t)\mathbf{J})}{d\langle N\rangle_t} = (\Psi_{1,\infty}(t),\ \Psi_{2,\infty}(t)\mathbf{J}) \cdot (\Upsilon \otimes \lambda \mathbf{A}(t)) \tag{2.43}$$

where $\mathbf{J} = \begin{pmatrix} J & J \\ J & J \end{pmatrix}$, $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\Upsilon = \begin{pmatrix} \vartheta & 0 & -a\mu^2 & 0 \\ 1 & \alpha & 0 & 0 \\ 0 & 0 & -\vartheta & -1 \\ 0 & -2 & 0 & -\alpha \end{pmatrix}$. When $-\frac{\vartheta^2\alpha^2}{2\mu^2} < a < 0$, , we have four real eigenvalue of the matrix Υ, we denote them $(x_i)_{i=1,2,3,4}$. It can be checked that there exists a constant $C > 0$ such that

$$\det \Psi_{1,\infty}(T) = \exp\left((x_1 + x_3)T\right)\left(C + \underset{T\to\infty}{O}\left(\frac{1}{T}\right)\right)$$

where $x_1 = \sqrt{\frac{\vartheta^2+\alpha^2+\sqrt{\mu^4-2a\mu^2}}{2}}$ and $x_3 = \sqrt{\frac{\vartheta^2+\alpha^2-\sqrt{\mu^4-2a\mu^2}}{2}}$. Therefore, due to the equality (2.41),

$$L_T(a) = \prod_{i\geq 1}(1 + 2a\nu_i(T)) \underset{T\to\infty}{\sim} \exp\left((\vartheta+\alpha)T\right)\left(\det \Psi_{1,\infty}(T)\right)^{-\frac{1}{2}} > 0. \tag{2.44}$$

Consequently, we have $\prod_{i\geq 1}(1 + 2a\nu_i(T)) > 0$ for any $a > -\frac{\vartheta^2\alpha^2}{2\mu^2}$ and $\lim_{T\to\infty} \nu_1(T) \leq \frac{\mu^2}{\vartheta^2\alpha^2}$ which achieves the proof.

5.3 Proof of Lemma 2.2

In this section we will prove that

$$\begin{aligned} \lim_{T\to\infty} \ln L_{2,T}\left(a, \vartheta + \frac{r_1}{\sqrt{T}}, \vartheta + \frac{r_2}{\sqrt{T}}\right) &= -a\frac{\mu^2\lambda^2}{2}\lim_{T\to\infty}\int_0^T Q(t)^*\mathcal{M}(t)Q(t)d\langle N\rangle_t \\ &= -a\frac{(r_2-r_1)^2}{2}\frac{\mu^2}{\alpha^2\vartheta^2}. \end{aligned}$$

Let us define the function $(\mathcal{P}(t) = \mathbf{E}_{\vartheta_1}\Pi_t, t \geq 0)$ which satisfies

$$\frac{d\mathcal{P}(t)}{d\langle N\rangle_t} = \mathcal{A}(t)\mathcal{P}(t) + \mathcal{D}(t)v(t), \quad \mathcal{P}(0) = \mathbf{0}_{4\times 1} \tag{2.45}$$

where

$$\mathcal{A}(t) = \begin{pmatrix} -\vartheta_1\lambda\mathbf{A}(t) & 0 \\ -(\vartheta_2 - \vartheta_1)\lambda\mathbf{A}(t) & a_{\vartheta_2}(t) \end{pmatrix}.$$

Using the asymptotical behavior of $\gamma_{\vartheta_2}(t)$ (see (2.15)), we get from [10] that $a_{\vartheta_2}(t) \sim -\alpha_2\lambda\mathbf{A}(t)$ (as $t \to \infty$) with $\alpha_2 = \sqrt{\mu^2 + \vartheta_2^2}$. Then, let us consider, in the following, the asymptotic behavior $\mathcal{P}_\infty(t)$, $\mathcal{H}_\infty(t)$ and $\varphi_\infty(t)$ of $\mathcal{P}(t)$, $\mathcal{H}(t)$ and $\varphi(t)$ respectively and solutions of equations (2.45), (2.31) and (2.33) substituying

$$\mathcal{A}_\infty(t) = \begin{pmatrix} -\vartheta_1 & 0 \\ -(\vartheta_2 - \vartheta_1) & -\alpha_2 \end{pmatrix} \otimes \lambda\mathbf{A}(t)$$

and

$$\mathcal{B}_\infty(t)\mathcal{B}_\infty(t)^* = \begin{pmatrix} g_1^2 & g_1g_2 \\ g_1g_2 & g_2^2 \end{pmatrix} \otimes \lambda\mathbf{A}(t)\mathbf{J}$$

in the place of $\mathcal{A}(t)$ and $\mathcal{B}(t)\mathcal{B}(t)^*$. Here $g_1 = \frac{\mu}{\sqrt{\lambda}(\alpha_1+\vartheta_1)}$ $g_2 = \frac{\mu}{\sqrt{\lambda}(\alpha_2+\vartheta_2)} - \frac{\mu}{\sqrt{\lambda}(\alpha_1+\vartheta_1)}$, $\alpha_1 = \sqrt{\mu^2 + \vartheta_1^2}$ and

$$\mathbf{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Let us also introduce the function $(Q_\infty(t), t \geq 0)$ defined by

$$Q_\infty(t) = \mathcal{P}_\infty(t) - a\lambda^2\mu^2 \int_0^t \varphi_\infty(t)\varphi_\infty^{-1}(s)\mathcal{H}_\infty(s)\mathcal{M}(s)Q_\infty(s)d\langle N\rangle_s. \tag{2.46}$$

then the asymptotic behavior of the equation (2.30) can be presented by:

$$\begin{aligned} \lim_{T\to\infty}\int_0^T Q(t)^*\mathcal{M}(t)Q(t)d\langle N\rangle_t &= \lim_{T\to\infty}\int_0^T \left(\mathcal{C}(t)^*Q(t)\right)^2 d\langle N\rangle_t \\ &= \lim_{T\to\infty}\int_0^T \left(\mathcal{C}(t)^*Q_\infty(t)\right)^2 d\langle N\rangle_t \\ &= \lim_{T\to\infty}\frac{1}{2\lambda}\left(\vartheta_2 - \vartheta_1\right)^2 \int_0^T \left(\frac{t^{\frac{1}{2}-H}\mathcal{C}(t)^*Q_\infty(t)}{\vartheta_2 - \vartheta_1}\right)^2 dt. \end{aligned}$$

Let us rewrite, using (2.46), the quantity $\overline{Q}_\infty(t) = \frac{t^{\frac{1}{2}-H}\mathcal{C}(t)Q_\infty(t)}{\vartheta_2-\vartheta_1}$ by

$$\overline{Q}_\infty(t) = \frac{t^{\frac{1}{2}-H}\mathcal{C}(t)^*\mathcal{P}_\infty(t)}{\vartheta_2 - \vartheta_1} - \frac{a\lambda\mu^2}{2}\int_0^t \left(F(t,s) + G(t,s)\right)\overline{Q}_\infty(s)ds \tag{2.47}$$

where

$$F(t,s) = \begin{pmatrix} t^{H-\frac{1}{2}} & t^{\frac{1}{2}-H} \end{pmatrix} \rho_{\alpha_2}(t)\rho_{\alpha_2}^{-1}(s) \begin{pmatrix} (\mathcal{H}_\infty^{3,3}(s)s^{2H-1} + \mathcal{H}_\infty^{3,4}(s))s^{\frac{1}{2}-H} \\ (\mathcal{H}_\infty^{4,3}(s)s^{2H-1} + \mathcal{H}_\infty^{4,4}(s))s^{\frac{1}{2}-H} \end{pmatrix}$$

and

$$\begin{aligned} G(t,s) &= \frac{\vartheta_2-\vartheta_1}{\vartheta_1-\alpha_2}\begin{pmatrix} t^{H-\frac{1}{2}} & t^{\frac{1}{2}-H}\end{pmatrix}(\rho_{\vartheta_1}(t)\rho_{\vartheta_1}^{-1}(s) \\ &\quad -\rho_{\alpha_2}(t)\rho_{\alpha_2}^{-1}(s))\begin{pmatrix} (\mathcal{H}_\infty^{1,3}(s)s^{2H-1}+\mathcal{H}_\infty^{1,4})s^{\frac{1}{2}-H} \\ (\mathcal{H}_\infty^{2,3}(s)s^{2H-1}+\mathcal{H}_\infty^{2,4})s^{\frac{1}{2}-H}\end{pmatrix}\end{aligned}$$

where $\mathcal{H}_\infty^{i,j}(s)$ is the i-th line and j-th row component of the matrix $\mathcal{H}_\infty(s)$ and ρ_ϑ is defined in (2.39). Let us fix $\vartheta_1=\vartheta+\frac{r_1}{\sqrt{T}}$ and $\vartheta_2=\vartheta+\frac{r_2}{\sqrt{T}}$. In order to prove the Lemma 2.2, we check successively :

1. $$\lim_{t\to+\infty}\frac{t^{\frac{1}{2}-H}\mathcal{C}(t)\mathcal{P}_\infty(t)}{\vartheta_2-\vartheta_1}=-\sqrt{\frac{2}{\lambda}}\frac{1}{\alpha_2\vartheta_1}$$
   with the same computations as in Lemma 2.3;

2. and for t and T large enough, there exists a constant $C_1>0$ such that
   $$\left|\int_0^t\left(F(t,s)+G(t,s)\right)\overline{Q}_\infty(s)ds\right|\le\frac{C_1}{T}.$$

On one hand, following from the proof of Lemma 2.1 in [12], we get:

$$|F(t,s)-(C_1\cdot e^{-\alpha_2(t-s)}+\frac{C_2}{t})(\mathcal{H}_\infty^{3,3}(s)s^{2H-1}+\mathcal{H}_\infty^{3,4}(s)+\mathcal{H}_\infty^{4,3}(s)s^{2H-1}+\mathcal{H}_\infty^{4,4}(s))|\le\frac{C_3}{(t\vee 1)^2}$$

where C_1, C_2 and C_3 are three constants. Moreover, it follows from [10] that when s is large enough

$$\begin{aligned}\mathcal{H}_\infty^{3,3}(s)s^{2H-1}+\mathcal{H}_\infty^{3,4}(s)+\mathcal{H}_\infty^{4,3}(s)s^{2H-1}+\mathcal{H}_\infty^{4,4}(s) &= \operatorname{trace}\left(\mathcal{H}_\infty(s)\mathcal{M}(s)\frac{d\langle N\rangle_s}{ds}\right) \\ &= \underset{T\to\infty}{O}(\frac{1}{T}).\end{aligned}$$

On the other hand,

$$|G(t,s)-\frac{\vartheta_2-\vartheta_1}{\vartheta_1-\alpha_2}\left(C_4(e^{-\vartheta_1(t-s)}-e^{-\alpha_2(t-s)})\right)R(s)|\le\frac{C_4}{(t\vee 1)^2}.$$

where $R(s)=\mathcal{H}_\infty^{1,3}(s)s^{2H-1}+\mathcal{H}_\infty^{1,4}(s)+\mathcal{H}_\infty^{2,3}(s)s^{2H-1}+\mathcal{H}_\infty^{2,4}(s)$ and C_4 is a constant. Let us define two processes I_t and J_t such that :

$$I_t=\begin{pmatrix} t^{H-\frac{1}{2}} & t^{\frac{1}{2}-H}\end{pmatrix}\pi_t^{\vartheta_1}(\zeta)\quad\text{and}\quad J_t=\begin{pmatrix} t^{H-\frac{1}{2}} & t^{\frac{1}{2}-H}\end{pmatrix}\delta_{\vartheta_1,\vartheta_2}.$$

It follows from [38] that

$$|R(s)-\operatorname{cov}(I_s,J_s)|\le\frac{C}{(s\vee 1)^2}$$

for s is large enough and C is a constant because

$$Var(I_t)=t^{2H-1}\mathcal{H}^{1,1}(t)+\mathcal{H}^{1,2}(t)+\mathcal{H}^{2,1}(t)+\mathcal{H}^{2,2}t^{1-2H}(t)$$

and

$$Var(J_t) = t^{2H-1}\mathcal{H}^{3,3}(t) + \mathcal{H}^{3,4}(t) + \mathcal{H}^{4,3}(t) + \mathcal{H}^{4,4}t^{1-2H}(t).$$

Let

$$\begin{aligned}
\phi_t &= \frac{d}{d\omega_H(t)}\int_0^t k_H(t,s)X_s ds \\
&= \lambda(t^{2H-1}\cdot Z_t^X + \int_0^t r^{2H-1}dZ_t^X) \\
&= \lambda\cdot\begin{pmatrix} t^{2H-1} & 1 \end{pmatrix}\zeta_t
\end{aligned}$$

where $Z_t^X = \int_0^t k_H(t,s)dX_s$. We can estimate the variance of I_t with

$$\begin{aligned}
VarI_t &= \frac{1}{\lambda}\mathbf{E}_{\vartheta_1}\left(t^{\frac{1}{2}-H}\pi_t^{\vartheta_1}(\phi) - t^{\frac{1}{2}-H}\mathbf{E}_{\vartheta_1}\phi_t^{\vartheta_1}\right)^2 \\
&= \frac{1}{\lambda}t^{1-2H}\left(\mathbf{E}_{\vartheta_1}(\pi_t^{\vartheta_1}(\phi))^2 - (\mathbf{E}_{\vartheta_1}\phi_t^{\vartheta_1})^2\right) \\
&\leq \frac{1}{\lambda}t^{1-2H}\left(\mathbf{E}_{\vartheta_1}(\phi_t^{\vartheta_1})^2 - (\mathbf{E}_{\vartheta_1}\phi_t^{\vartheta_1})^2\right) \\
&= \frac{1}{\lambda}t^{1-2H}Var(\phi_t^{\vartheta_1})
\end{aligned}$$

Following from [38], $Var(\phi_t^{\vartheta_1}) \underset{t\to\infty}{\sim} \frac{\lambda(\alpha_1-\vartheta_1)}{\mu^2}$, so with the Cauchy-Schwarz inequality $|\operatorname{cov}(I_t, J_t)| = \underset{T\to\infty}{O}(\frac{1}{T})$.
Finally, we get for t large enough that

$$\int_0^t |F(t,s) + G(t,s)|\, ds \leq \frac{C_2}{T}. \tag{2.48}$$

Consequently

$$\begin{aligned}
\left|\int_0^t (F(t,s) + G(t,s))\,\overline{Q}_\infty(s)ds\right| &\leq \int_0^t |(F(t,s) + G(t,s))|\,\left|\overline{Q}_\infty(s)\right| ds \\
&\leq C_5\int_0^t |(F(t,s) + G(t,s))|\, ds \\
&\leq \frac{C_1}{T}
\end{aligned}$$

provided that $\left|\overline{Q}_\infty(t)\right| \leq C_5$ for t and T large enough that will be explained below. To show this, let us defined the operator S defined by

$$S(f)(t) = \frac{a\lambda\mu^2}{2}\int_0^t |F(t,s) + G(t,s)|\, f(s)ds.$$

Equation (2.47) leads to

$$|\overline{Q}_\infty(t)| \leq \left|\frac{t^{\frac{1}{2}-H}\mathcal{C}(t)^*\mathcal{P}_\infty(t)}{\vartheta_2 - \vartheta_1}\right| + S(|\overline{Q}_\infty|)(t)$$

or

$$(I-S)\,(|\overline{Q}_\infty|)(t) \le \left|\frac{t^{\frac{1}{2}-H}\mathcal{C}(t)^*\mathcal{P}_\infty(t)}{\vartheta_2-\vartheta_1}\right| \le C_6 \quad \text{for } t \text{ large enough.}$$

Since, with (2.48), we have $\|S\| \le \frac{C_7}{T} < 1$ for T large enough, we can compute

$$\begin{aligned} |\overline{Q}_\infty(t)| &\le (I-S)^{-1}(C_4)(t) = \prod_{n=1}^{\infty} S^n(C_4)(t) \\ &\le C_5, \end{aligned}$$

that concludes the proof.

Chapter 3

Stochastic analysis of mixed fractional Brownian motion

1 Introduction

In this part we present a new perspective on the *mixed* fractional Brownian motion, i.e., the process

$$X_t = B_t + B_t^H, \quad t \in [0, T], \quad T > 0, \tag{3.1}$$

where $B = (B_t)$ is the standard Brownian motion and $B^H = (B_t^H)$ is independent fractional Brownian motion (fBm) with the Hurst exponent $H \in (0, 1]$, that is, a centered Gaussian process with covariance function

$$K(s,t) = \mathbf{E} B_t^H B_s^H = \frac{1}{2}\Big(t^{2H} + s^{2H} - |t-s|^{2H}\Big), \quad s, t \in [0, T].$$

The fBm B^H coincides with the standard Brownian motion at $H = \frac{1}{2}$, but otherwise differs from it in many ways, including path regularity, dependence range of the increments, etc. The diversity of properties makes it an interesting mathematical object as well as a powerful modeling tool in a variety of applications (see, e.g., [9], [55]).

The interest in the process (3.1) was triggered by P.Cheridito's paper [16], where the author discovered a curious change in the properties of X occurring at $H = \frac{3}{4}$. It turns out that X is a semimartingale in its own filtration if and only if either $H = \frac{1}{2}$ or $H \in \big(\frac{3}{4}, 1\big]$ and, moreover, in the latter case, the probability measure μ^X, induced by X on the measurable space of continuous functions $C([0, T])$, is equivalent to the standard Wiener measure μ^W.

Since B^H is not a semimartingale on its own, unless $H = \frac{1}{2}$ or $H = 1$, this assertion means that B^H can be "regularized" up to a semimartingale by adding to it an independent Brownian perturbation. In [16] this fact is discussed in the context of the option pricing problem from mathematical finance and arbitrage opportunities on non-semimartingale markets (see also [17]). A comprehensive survey of further related developments in finance can be found in [7]. Besides being of interest to finance community, the result in [16] also led to a number of elegant generalizations and alternative proofs, some of which are briefly recalled in Section 2.3 below.

Our objective is to develop the basic toolbox for analysis of the mixed fBm, based on the filtering theory of Gaussian processes. The main ingredient is the so called *fundamental martingale*, whose natural filtration coincides with the filtration generated by the process X, and with respect to which, X can be represented as a stochastic integral and vise versa. For more general additive models driven by the mixed fBm this notion naturally generalizes to *fundamental semimartingale* and leads to the Girsanov type change of measure (Theorems 3.1 and 3.7 and Corollaries 3.2 and 3.8).

Using our approach we give a new direct proof of the aforementioned Cheridito's regularization theorem, which besides establishing the already known semimartingality and equivalence properties, also yields a representation of the mixed fBm as a diffusion type process in its own filtration and a formula for the corresponding Radon-Nikodym derivative. Counterparts of these results are also derived for $H < \frac{1}{4}$, in which case μ^X was shown equivalent to the measure induced by B^H by H. van Zanten in [74] (Theorem 3.3).

Another area of potential applications is statistical analysis of models, driven by mixed fractional noises. To demonstrate the ideas, in this paper we consider only the basic linear regression setting and show how the Maximum Likelihood Estimator (MLE) can be defined and studied in the large sample asymptotic regime (Theorem 3.4).

The rest of this chapter has the following structure. Our main results are detailed in the next section. Some frequently used notations and auxiliary results are gathered in Section 3 and the proofs appear in Sections 4 to 7.

2 The main results

Let $\mathcal{F}^X = (\mathcal{F}^X_t)$ and $\mathcal{F} = (\mathcal{F}_t)$, $t \in [0,T]$ be the natural filtrations of X and (B, B^H) respectively and consider the filtering process

$$M_t = \mathbf{E}\big(B_t|\mathcal{F}^X_t\big), \quad t \in [0,T]. \tag{3.2}$$

Since B is an $\mathcal{F}$-martingale and $\mathcal{F}^X_t \subseteq \mathcal{F}_t$, the process M is an $\mathcal{F}^X$-martingale. Remarkably, M encodes many of the essential features of the process X, making its structure particularly transparent. As shown below, M and X generate the same filtrations and can be expressed as stochastic integrals with respect to each other. More precisely, M admits the representation:

$$M_t = \int_0^t g(s,t)dX_s, \quad \langle M\rangle_t = \int_0^t g(s,t)ds, \quad t \ge 0, \tag{3.3}$$

where the kernel $g(s,t)$ solves integro-differential equation

$$g(s,t) + H\frac{d}{ds}\int_0^t g(r,t)|s-r|^{2H-1}sign(s-r)dr = 1, \quad 0 < s < t \le T. \tag{3.4}$$

The family of functions $\big\{g(s,t), 0 \le s \le t \le T\big\}$ plays the key role in our approach to analysis of the mixed fBm.

The equation (3.4) is uniquely solved by $g(s,t) \equiv \frac{1}{2}$ and $g(s,t) \equiv 1/(1+t)$ for $H = \frac{1}{2}$ and $H = 1$ respectively and for other values of H can be rewritten as a simpler integral equation with a weakly singular kernel, whose precise formula is determined by the range of H. A particularly neat form is obtained for $H > \frac{1}{2}$, as elaborated in the following subsection.

2.1 Mixed fBm for $H > \frac{1}{2}$

For H in this range, the derivative and integration in (3.4) can be interchanged and it takes the form of integral equation:

$$g(s,t) + H(2H-1)\int_0^t g(r,t)|r-s|^{2H-2}dr = 1, \quad 0 \le s \le t \le T. \tag{3.5}$$

It will be convenient to extend definition of $g(s,t)$ to the domain $0 \le t \le s \le T$ by setting

$$g(s,t) := 1 - H(2H-1)\int_0^t g(r,t)|r-s|^{2H-2}dr, \tag{3.6}$$

so that $g(s,t)$ satisfies (3.5) for all $s,t \in [0,T]$.

For $H \in \left(\frac{1}{2}, 1\right]$ the kernel $\kappa(s,r) := H(2H-1)|r-s|^{2H-2}$, $s,r \in [0,T]$ has a weak (integrable) singularity on the diagonal. In this case the equation (3.5) is well known to have unique continuous solution (see, e.g., [72]), which satisfies various regularity properties, implicitly required by our results and elaborated in the course of the proofs. For example, the derivative $\dot{g}(s,t) = \frac{\partial}{\partial t}g(s,t)$ explodes at the endpoints of the interval $[0,t]$ but, nevertheless, belongs to $L^2([0,t])$ if $H > 3/4$ (Lemma 3.4).

Though this equation reduces to a particular instance of the Riemann boundary value problem (see, e.g., [58]), its solution does not admit an explicit form. Nevertheless, it can be efficiently approximated numerically (see, e.g., [73]).

Along with the function $g(s,t)$, let us define

$$R(s,t) := \frac{\dot{g}(s,t)}{g(t,t)}, \quad \dot{g}(s,t) := \frac{\partial}{\partial t}g(s,t), \quad s \ne t, \tag{3.7}$$

and

$$G(s,t) := 1 - \frac{1}{g(s,s)}\int_0^t R(\tau,s)d\tau, \quad 0 \le s \le t \le T. \tag{3.8}$$

As we show below $g(t,t) > 0$ for all $t \ge 0$ and the functions in (3.7) and (3.8) are well defined. The following theorem summarizes a number of useful representation formulas:

Theorem 3.1. *The $\mathcal{F}^X$-martingale M, defined in* (3.2)*, satisfies* (3.3) *and*

$$\langle M \rangle_t = \int_0^t g^2(s,s)ds, \tag{3.9}$$

where $g(s,t)$ is the unique solution of equation (3.5)*. Moreover,*

$$X_t = \int_0^t G(s,t)dM_s, \quad t \in [0,T], \tag{3.10}$$

with G, defined by (3.8)*, and, in particular, $\mathcal{F}_t^X = \mathcal{F}_t^M$, $\mathbf{P}$-a.s. for all $t \in [0,T]$.*

The equality in (3.9) suggests that the martingale M admits *innovation* type representation, which can be used to analyze the structure of the mixed fBm with stochastic drifts and to derive an analogue of Girsanov's theorem:

Corollary 3.2. *Consider a process $Y = (Y_t)$ defined by*

$$Y_t = \int_0^t f(s)ds + X_t, \quad t \in [0, T], \tag{3.11}$$

where $f = \big(f(t)\big)$ is a process with continuous paths and $\mathbf{E}\int_0^T \big|f(t)\big|dt < \infty$, adapted to a filtration $\mathcal{G} = (\mathcal{G}_t)$, with respect to which M is a martingale. Then Y admits the representation

$$Y_t = \int_0^t G(s,t)dZ_s \tag{3.12}$$

with G, defined in (3.8), where the process $Z = (Z_t)$

$$Z_t = \int_0^t g(s,t)dY_s, \quad t \in [0, T]$$

is a $\mathcal{G}$-semimartingale with the Doob–Meyer decomposition

$$Z_t = M_t + \int_0^t \Phi(s)d\langle M\rangle_s, \tag{3.13}$$

where

$$\Phi(t) = \frac{d}{d\langle M\rangle_t}\int_0^t g(s,t)f(s)ds. \tag{3.14}$$

In particular, $\mathcal{F}_t^Y = \mathcal{F}_t^Z$, $\mathbf{P}$-a.s. for all $t \in [0,T]$ and, if

$$\mathbf{E}\exp\left\{-\int_0^T \Phi(t)dM_t - \frac{1}{2}\int_0^T \Phi^2(t)d\langle M\rangle_t\right\} = 1,$$

then the measures μ^X and μ^Y are equivalent and the corresponding Radon-Nikodym derivative is given by

$$\frac{d\mu^Y}{d\mu^X}(Y) = \exp\left\{\int_0^T \hat{\Phi}(t)dZ_t - \frac{1}{2}\int_0^T \hat{\Phi}^2(t)d\langle M\rangle_t\right\}, \tag{3.15}$$

where $\hat{\Phi}(t) = \mathbf{E}\big(\Phi(t)|\mathcal{F}_t^Y\big)$.

Remark 7. *The choice of the filtration $\mathcal{G}$ is obvious in typical applications. For example, in filtering problems $f(t)$ plays the role of the unobserved state process and X is interpreted as the observation noise. If the state process and the noise are independent, the corollary applies with $\mathcal{G}_t := \mathcal{F}_t^f \vee \mathcal{F}_t^X$.*

If $f(t)$ is a function of Y_t, then (3.11) becomes a stochastic differential equation with respect to the mixed fBm X. In this case, $f(t)$ is adapted to $\mathcal{F}^X$ itself and hence the natural choice is $\mathcal{G}_t := \mathcal{F}_t^X$. For example, $f(t) := \theta Y_t$ with $\theta \in \mathbb{R}$ corresponds to the mixed fractional Ornstein– Uhlenbeck process:

$$Y_t = \theta\int_0^t Y_s ds + X_t, \quad t \in [0, T]. \tag{3.16}$$

Remark 8. *Equality of the filtrations $\mathcal{F}^X$ and $\mathcal{F}^M$ means that the information contained in X is preserved progressively in M. Therefore, following the terminology of [56], [45] and [37], M merits to be called a* fundamental martingale *associated with the mixed fBm X. Similarly, Z is a* fundamental semimartingale *associated with the process Y.*

2.2 Mixed fBm for $H < \frac{1}{2}$

In this case the derivative and integration in (3.4) are no longer interchangeable, but nevertheless it can still be reduced to a weakly singular integral equation (Theorem 3.5). Being somewhat more involved, the details are deferred to Subsection 5.2 below, where Theorem 3.1 and Corollary 3.2 are generalized to all $H \in (0,1]$.

Instead let us briefly describe an alternative "indirect" approach, which can also be used to derive results analogous to those in the previous subsection. The trick is to transform X into

$$\widetilde{X}_t = \int_0^t \widetilde{\rho}(s,t)dX_s, \quad t \in [0,T],$$

where the kernel $\widetilde{\rho}(s,t)$, whose explicit formula appears in (3.59) below, is such that the process

$$\widetilde{B}_t = \int_0^t \widetilde{\rho}(s,t)dB_s^H,$$

is a standard Brownian motion. The main point of this transformation is that the Gaussian process

$$\widetilde{U}_t = \int_0^t \widetilde{\rho}(s,t)dB_s$$

has covariance function with integrable partial derivative:

$$\widetilde{\kappa}(s,t) := \frac{\partial^2}{\partial s \partial t}\mathbf{E}\widetilde{U}_s\widetilde{U}_t = |t-s|^{-2H}\chi\left(\frac{s\wedge t}{s\vee t}\right), \quad s \neq t, \tag{3.17}$$

where $\chi(\cdot)$ is a continuous function, specified in (3.60). Therefore the process $\widetilde{X} = \widetilde{B} + \widetilde{U}$ with $H < \frac{1}{2}$ has the structure, similar to the original process X with $H > \frac{1}{2}$. In particular, the martingale $\widetilde{M}_t = \mathbf{E}\big(\widetilde{B}_t|\mathcal{F}_t^{\widetilde{X}}\big)$ admits the representation

$$\widetilde{M}_t = \int_0^t \widetilde{g}(s,t)\widetilde{X}_s, \quad t \in [0,T] \tag{3.18}$$

where $\widetilde{g}(s,t)$ satisfies the weakly singular equation (cf. (3.5)):

$$\widetilde{g}(s,t) + \int_0^t \widetilde{g}(r,t)\widetilde{\kappa}(r,s)dr = 1, \quad 0 \le s \le t \le T. \tag{3.19}$$

It can be shown that all three processes X, $\widetilde{X}$ and $\widetilde{M}$ generate the same filtrations and thus the martingale $\widetilde{M}$ is also fundamental. Moreover, it turns out that the original martingale M, defined in (3.2), and the martingale $\widetilde{M}$ can be represented as stochastic integrals with respect to each other. In fact, both are generated by the same innovation Brownian motion (see Lemma 3.10). Analogs of Theorem 3.1 and Corollary 3.2 can now be readily obtained, using the same techniques as in the case $H > \frac{1}{2}$ (see Subsection 5.1 for the details).

2.3 Semimartingale structure of X

As mentioned above, P. Cheridito showed in [16] that X is a semimartingale in its own filtration if and only if $H \in \left\{\frac{1}{2}\right\} \cup \left(\frac{3}{4}, 1\right]$ and, moreover, $\mu^X \sim \mu^W$ for $H > \frac{3}{4}$.

This statement is evident for $H = \frac{1}{2}$, for which X is just a sum of two independent Brownian motions. It also holds by a simple argument for $H \in \left(0, \frac{1}{2}\right)$. Indeed, as is well known, the p-power variation of B^H is finite and positive for $p = \frac{1}{H}$ (see, e.g., Section 1.8 in [55] for precise definitions and related results). Hence for $H < \frac{1}{2}$, the quadratic variation of B^H and thus also of X is infinite, preventing it from being a semimartingale and, a fortiori, from being equivalent to B. A more delicate argument is required for $H \in \left(\frac{1}{2}, 1\right]$, since in this range the quadratic variation of B^H vanishes, and consequently X has the same quadratic variation as B.

To show that X is not a semimartingale for $H \in (\frac{1}{2}, \frac{3}{4}]$, the author first argues in [16] that X cannot be a semimartingale if it is not a *quasimartingale*, i.e. does not satisfy the property

$$\sup_{\tau} \sum_{j=0}^{n-1} \mathbf{E} \left| \mathbf{E}\left(X_{t_{j+1}} - X_{t_j} \middle| \mathcal{F}^X_{t_j}\right) \right| < \infty$$

where τ is the set of all finite partitions $0 = t_0 < t_1 < ... < t_n = T$. Then he shows that the above sums are unbounded for $H \in (\frac{1}{2}, \frac{3}{4}]$ on the sequence of uniform partitions.

The equivalence of X and B for $H > \frac{3}{4}$ and the consequent semimartingale property of X are shown in [16] using the Hida-Hitsuda [29] criterion for equivalence of measures in terms of the relative entropies between the restrictions of these measures to finite partitions.

F. Baudoin and D. Nualart [6] noticed that the Hida-Hitsuda criterion actually applies in the more general setting and showed that the process $X := B + V$, where V is a centered Gaussian process with covariance function K, is equivalent to a Brownian motion, if $\partial^2 K/\partial s \partial t \in L^2([0,T]^2)$. In particular, for $V_t := B^H_t$, this partial derivative is square integrable for $H > 3/4$, confirming the result in [16].

The next extension of Cheridito's result is due to H. van Zanten [74, 75], who addresses the question of equivalence of a linear combination $\xi = \sum_{k=1}^n \alpha_k B^{H_k}$ of n independent fBm's with the Hurst exponents $H_1 < ... < H_n$ and nonzero constants $\alpha_1, ..., \alpha_n$, to a single fBm. Using spectral techniques for processes with stationary increments, van Zanten shows that X and $\alpha_1 B^{H_1}$ are equivalent if $H_2 - H_1 > \frac{1}{4}$, and, conversely, if X is equivalent to a multiple of an fBm, then it must be equivalent to $\alpha_1 B^{H_1}$ and $H_2 - H_1 > \frac{1}{4}$. The Radon-Nikodym derivative between the measures is given in [74] in terms of certain reproducing kernels, but the author points out that it might be hard to obtain more explicit expression (see remark (iii) on page 63). Also the results in [74] do not imply semimartingality of X.

The following theorem gives a representation of X as a diffusion type process in its own filtration and a formula for the Radon-Nikodym derivative in terms of the solution of equations (3.5) and (3.19):

Theorem 3.3.

1. *The process X defined in* (3.1) *is a semimartingale in its own filtration if and only if $H \in \{\frac{1}{2}\} \cup (\frac{3}{4}, 1]$. For $H \in (\frac{3}{4}, 1]$, X is a diffusion type process:*

$$X_t = W_t - \int_0^t \varphi_s(X) ds, \quad t \in [0, T],$$

where

$$W_t = \int_0^t \frac{1}{g(s,s)} dM_s \tag{3.20}$$

is an $\mathcal{F}^X$*-Brownian motion and* $\varphi_t(X) = \int_0^t R(s,t)dX_s$*, with* $R(s,t)$ *defined in* (3.7)*. Moreover, the measures* μ^X *and* μ^W *are equivalent and*

$$\frac{d\mu^X}{d\mu^W}(X) = \exp\left\{ -\int_0^T \varphi_t(X)dX_t - \frac{1}{2}\int_0^T \varphi_t^2(X)dt \right\}.$$

2. The measures μ^X *and* μ^{B^H} *are equivalent if and only if* $H < \frac{1}{4}$ *and*

$$\frac{d\mu^X}{d\mu^{B^H}}(X) = \exp\left\{ -\int_0^T \widetilde{\varphi}_t(\widetilde{X})d\widetilde{X}_t - \frac{1}{2}\int_0^T \widetilde{\varphi}_t^2(\widetilde{X})dt \right\}, \tag{3.21}$$

where $\widetilde{\varphi}_t(\widetilde{X}) = \int_0^t \widetilde{R}(s,t)d\widetilde{X}_s$*,* $\widetilde{R}(s,t) := \dfrac{\dot{\widetilde{g}}(s,t)}{\widetilde{g}(t,t)}$ *and* $\widetilde{g}(s,t)$ *is the solution of* (3.19)*.*

2.4 Drift estimation in mixed fractional noise

As another application, we consider the problem of construction and large sample asymptotic analysis of the Maximum Likelihood Estimator (MLE) of the unknown drift parameter of the mixed fBm. Let

$$Y_t = \theta t + \beta B_t + \alpha B_t^H, \quad t \in [0,T] \tag{3.22}$$

where β, α and H are known constants and B and B^H are standard and fractional Brownian motions respectively. It is required to estimate the unknown parameter $\theta \in \mathbb{R}$, given the sample $Y^T = \{Y_t,\ t \in [0,T]\}$. While various reasonable estimators can be suggested for this purpose, the MLE is traditionally of a special interest due to its well known large sample optimality properties. Sometimes the performance of MLE is considered as a benchmark for estimators with simpler structure, such as, e.g., least squares estimator, and an explicit formula for the asymptotic variance of MLE often comes handy.

The problem of constructing the MLE is elementary in the case $\alpha = 0$, i.e. in absence of the fractional component. In the case of purely fractional noise, i.e. when $\beta = 0$, it was solved in [45]. Parameter estimation in models with mixed fBm such as (3.22), was considered in the recent monographs [55] and [61], where the construction of the MLE for θ appears as an open problem (see Remark (iii) page 181 in [61] and the discussion on page 354 in [55]). The following theorem aims at filling this gap (w.l.o.g. $\alpha = \beta = 1$ will be assumed hereafter).

Theorem 3.4. *The MLE of* θ *is given by*

$$\hat{\theta}_T(Y) = \frac{\int_0^T g(s,T)dY_s}{\int_0^T g(s,T)ds}, \tag{3.23}$$

where the function $g(s,T)$, $s \in [0,T]$ is the unique solution of equation (3.4) *with $t := T$. For $H \in (0,1)$ this estimator is strongly consistent and the corresponding estimation error is normal*

$$\hat{\theta}_T - \theta \sim N\left(0, \frac{1}{\int_0^T g(s,T)ds}\right), \tag{3.24}$$

with the following asymptotic variance:

1. *for $H > \frac{1}{2}$,*

$$\lim_{T\to\infty} T^{2-2H}\mathbf{E}(\hat{\theta}_T - \theta)^2 = \frac{2H\Gamma(H+\frac{1}{2})\Gamma(3-2H)}{\Gamma(\frac{3}{2}-H)}, \tag{3.25}$$

where $\Gamma(\cdot)$ is the standard Gamma function.

2. *for $H < \frac{1}{2}$,*

$$\lim_{T\to\infty} T\mathbf{E}\big(\hat{\theta}_T - \theta\big)^2 = 1. \tag{3.26}$$

Remark 9. *The constant in the right hand side of* (3.25) *coincides with the asymptotic variance, obtained in [45] for the problem of estimating the drift θ from the observations with purely fractional noise:*

$$\bar{Y}_t = \theta t + B_t^H, \quad t \in [0,T].$$

Hence either the Brownian or the fractional Brownian component is asymptotically negligible for $H > \frac{1}{2}$ and $H < \frac{1}{2}$ respectively.

Remark 10. *The fundamental martingales M and $\widetilde{M}$, introduced above, are also expected to play a key role in the statistical analysis of models more general than* (3.22)*, such as, e.g., the mixed fractional Ornstein– Uhlenbeck process* (3.16)*. The progress in this direction will be reported elsewhere.*

3 Notations and Auxiliary Results

3.1 Notations

Throughout we assume that all the random variables are supported on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We will frequently use the constants

$$c_H = \frac{1}{2H\Gamma\big(\frac{3}{2}-H\big)\Gamma\big(H+\frac{1}{2}\big)}, \quad \lambda_H = \frac{2H\Gamma\big(H+\frac{1}{2}\big)\Gamma(3-2H)}{\Gamma\big(\frac{3}{2}-H\big)}$$
$$\beta_H = c_H^2\left(\frac{1}{2}-H\right)^2\frac{\lambda_H}{2-2H}.$$

For a measurable function f on $[0,T]$ and $t \in [0,T]$, we define

$$K_f(s,t) = -2H \frac{d}{ds} \int_s^t f(r) r^{H-1/2} (r-s)^{H-1/2}\, dr\,, \quad 0 \le s \le t\,,$$

and

$$Q_f(s) = \frac{d}{ds} \int_0^s f(r) r^{1/2-H} (s-r)^{1/2-H}\, dr, \quad 0 \le s \le t.$$

These operators are readily related to the Riemann−Liouville fractional integrals and derivatives (see [67]). The respective inversion formulas are

$$f(s) = -c_H s^{\frac{1}{2}-H} \frac{d}{ds} \int_s^t K_f(r,t)(r-s)^{\frac{1}{2}-H} dr, \tag{3.27}$$

and

$$f(s) = 2Hc_H s^{H-\frac{1}{2}} \frac{d}{ds} \int_0^s Q_f(r)(s-r)^{H-\frac{1}{2}} dr, \tag{3.28}$$

where the equalities hold almost everywhere.

Following the notations of [35] and [59], define the space

$$\Lambda_t^{H-\frac{1}{2}} := \left\{ f : [0,t] \mapsto \mathbb{R} \text{ such that } \int_0^t \left(s^{\frac{1}{2}-H} K_f(s,t) \right)^2 ds < \infty \right\},$$

with the scalar product

$$\langle f, g \rangle_{\Lambda_t^{H-\frac{1}{2}}} := \frac{2-2H}{\lambda_H} \int_0^t s^{1-2H} K_f(s,t) K_g(s,t) ds. \tag{3.29}$$

For $H > \frac{1}{2}$

$$\Lambda_t^{H-\frac{1}{2}} = \left\{ f : [0,t] \mapsto \mathbb{R} \text{ such that } \int_0^t \int_0^t f(u) f(v) |u-v|^{2H-2} du dv < \infty \right\},$$

and the inclusion $L^2([0,t]) \subset \Lambda_t^{H-\frac{1}{2}}$ holds. This inclusion fails for $H < \frac{1}{2}$ (see Remark 4.2 in [59]).

For any $H \in (0,1)$ and $\phi, \psi \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$, the following identity holds

$$\int_0^t \phi(s)\psi(s) ds = c_H \int_0^t K_\phi(s,t) Q_\psi(s) ds. \tag{3.30}$$

Recall that for $0 < \alpha, \beta < 1$

$$\int_0^T |s-r|^{-\alpha} |r-t|^{-\beta} dr \le \begin{cases} C_1 |s-t|^{1-\alpha-\beta} & \alpha+\beta > 1 \\ C_2 \log \frac{T}{|s-t|} + C_3 & \alpha+\beta = 1 \\ C_4 & \alpha+\beta < 1 \end{cases} \tag{3.31}$$

Here and below C_i, c_i, $i = 1, 2, \ldots$ stand for constants depending only on H and T, whose precise values are of no importance, often changing from line to line.

For $H > \frac{1}{2}$, define

$$\kappa(s,r) := H(2H-1)|r-s|^{2H-2}.$$

This kernel is *weakly singular*, in the sense $\sup_{s\in[0,T]}\int_0^T k(s,r)dr < \infty$. We will denote by $\kappa^{(m)}$, $m=1,2,...$ the m-th iteration of the kernel κ, that is, $\kappa^{(1)}(s,t)=\kappa(s,t)$ and

$$\kappa^{(m)}(s,t)=\int_0^t \kappa^{(m-1)}(s,r)\kappa(r,t)dr, \quad m=2,3,...$$

By (3.31), for $H>\frac{1}{2}$, $\kappa^{(m)}(\cdot,t)\in L^2([0,t])$ for $m>\frac{1}{4H-2}$ and, moreover, $\kappa^{(m)}(\cdot,t)\in C([0,t])$ for $m>\frac{1}{2H-1}$. Similar relations hold for the kernel $\widetilde{\kappa}$ with $H<\frac{1}{2}$, defined in (3.17).

3.2 Martingale representation lemma

For the reader's convenience, let us briefly recall some relevant properties of the integrals with respect to fBm. For the simple function of the form,

$$f(u)=\sum_{k=1}^n f_k \mathbf{1}_{u\in[u_k,u_{k+1})}, \quad f_k\in\mathbb{R}, \quad 0=u_1<u_2<...<u_k=t,$$

the stochastic integral with respect to B^H is defined as

$$\int_0^t f(s)dB_s^H := \sum_{k=1}^n f_k\big(B^H_{u_{k+1}}-B^H_{u_k}\big).$$

Since simple functions are dense in $\Lambda_t^{H-\frac{1}{2}}$ (see Theorem 4.1 in [59]), the definition of $\int_0^t f(s)dB_s^H$ extends to $f\in\Lambda_t^{H-\frac{1}{2}}$ through the limit

$$\int_0^t f(s)dB_s^H := \lim_n \int_0^t f_n(s)dB_s^H,$$

where f_n is any sequence of simple functions, such that $\lim_n \|f-f_n\|_{\Lambda_t^{H-\frac{1}{2}}}=0$. Moreover, for $f,g\in\Lambda_t^{H-\frac{1}{2}}$, cf. (3.29),

$$\mathbf{E}\int_0^t f(s)dB_s^H\int_0^t g(s)dB_s^H=\frac{2-2H}{\lambda_H}\int_0^t s^{1-2H}K_f(s,t)K_g(s,t)ds. \tag{3.32}$$

For $H>1/2$, the formula in the right hand side of (3.32) simplifies to

$$\mathbf{E}\int_0^t f(s)dB_s^H\int_0^t g(s)dB_s^H=\int_0^t\int_0^t f(r)g(s)\kappa(r,s)drds. \tag{3.33}$$

It turns out however (see Section 5 of [59]), that for $H>\frac{1}{2}$ the image of $\Lambda_t^{H-\frac{1}{2}}$ under the map $f\mapsto\int_0^t f(s)dB_s^H$ is a strict subset of $\overline{\mathrm{sp}}_{[0,t]}(B^H)$, the closure in $L^2(\Omega,\mathcal{F},\mathbf{P})$ of all possible linear combinations of the increments of B^H. In other words, some linear functionals of B^H cannot be realized as stochastic integrals of the above type and thus the representation of M as a stochastic integral is not entirely evident at the outset.

Lemma 3.1. *Let η be a Gaussian random variable, such that (η, X_t), $t \in [0,T]$ forms a Gaussian process. Then there exists a unique function $h(\cdot,t) \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$, such that*

$$\mathbf{E}(\eta|\mathcal{F}_t^X) = \mathbf{E}\eta + \int_0^t h(s,t)dX_t, \quad \mathbf{P}-a.s.$$

Proof. Following the arguments of the proof of Lemma 10.1 in [?], let $t_i = ti/2^n$, $i = 0,...,2^n$ and $\mathcal{F}_{t,n}^X = \sigma\{X_{t_i} - X_{t_{i-1}}, i = 1,...,2^n\}$. Then $\mathcal{F}_{t,n}^X \nearrow \mathcal{F}_t^X$ and by martingale convergence

$$\lim_n \mathbf{E}(\eta|\mathcal{F}_{t,n}^X) = \mathbf{E}(\eta|\mathcal{F}_t^X), \quad \mathbf{P}-a.s. \tag{3.34}$$

Since $\mathbf{E}(\eta|\mathcal{F}_{t,n}^X)$ are uniformly integrable, this convergence also holds in $L^2(\Omega,\mathbf{P})$. By the normal correlation theorem,

$$\mathbf{E}(\eta|\mathcal{F}_{t,n}^X) = \mathbf{E}\eta + \sum_{i=1}^{2^n} h_{i-1}^n\big(X_{t_i} - X_{t_{i-1}}\big),$$

with some constants h_{i-1}^n, $i = 1,...,2^n$. Define

$$h_n(s,t) := \sum_{i=1}^{2^n} h_{i-1}^n \mathbf{1}_{s\in[t_{i-1},t_i)},$$

then

$$\mathbf{E}\big(\eta|\mathcal{F}_{t,n}^X\big) = \mathbf{E}\eta + \int_0^t h_n(s,t)dB_s + \int_0^t h_n(s,t)dB_s^H,$$

and

$$\mathbf{E}\Big(\mathbf{E}\big(\eta|\mathcal{F}_{t,n}^X\big) - \mathbf{E}\big(\eta|\mathcal{F}_{t,m}^X\big)\Big)^2 = \int_0^t \big(h_n(s,t) - h_m(s,t)\big)^2 ds + \frac{2-2H}{\lambda_H}\int_0^t \Big(s^{\frac{1}{2}-H}(K_{h_n} - K_{h_m})(s,t)\Big)^2 ds.$$

Since the convergence (3.34) holds in $L_2(\Omega,\mathbf{P})$,

$$\lim_n \sup_{m\ge n}\left(\|h_n - h_m\|_2 + \|h_n - h_m\|_{\Lambda_t^{H-\frac{1}{2}}}\right) \le \lim_n \sup_{m\ge n} \mathbf{E}\Big(\mathbf{E}\big(\eta|\mathcal{F}_{t,n}^X\big) - \mathbf{E}\big(\eta|\mathcal{F}_{t,m}^X\big)\Big)^2 = 0,$$

and, by completeness of $L_2([0,t])$, there exists a function $h(\cdot,t) \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$, such that $\lim_n \|h - h_n\|_2 = \lim_n \|h - h_n\|_{\Lambda_t^{H-\frac{1}{2}}} = 0$. The claimed representation now follows, since

$$\mathbf{E}\left(\mathbf{E}(\eta|\mathcal{F}_t^X) - \mathbf{E}\eta - \int_0^t h(s,t)dB_s - \int_0^t h(s,t)dB_s^H\right)^2 \le$$
$$3\mathbf{E}\Big(\mathbf{E}(\eta|\mathcal{F}_t^X) - \mathbf{E}(\eta|\mathcal{F}_{t,n}^X)\Big)^2 + 3\int_0^t \big(h_n(s,t) - h(s,t)\big)^2 ds +$$
$$+ 3\int_0^t \Big(s^{\frac{1}{2}-H}(K_{h_n} - K_h)(s,t)\Big)^2 ds \xrightarrow{n\to\infty} 0.$$

The uniqueness of h is obvious. □

3.3 The equation (3.4) and its alternative forms

In this subsection we verify the representation (3.3) and explore the associated equation (3.4), rewriting it in a number of convenient alternative forms.

Theorem 3.5. *The representation* (3.3) *holds with* $g(s,t)$*,* $s \in [0,t]$ *being the unique continuous solution of the following equations:*

i. for $H \in (0,1]$*, the integro-differential equation* (3.4)

ii. for $H \in (0,1)$*, the fractional integro-differential equation*

$$c_H Q_g(s) + \frac{2-2H}{\lambda_H} K_g(s,t) s^{1-2H} = c_H Q_1(s), \quad s \in (0,t] \tag{3.35}$$

iii. for $H \in (\frac{1}{2},1]$*, the weakly singular integral equation* (3.5)

iv. for $H \in (0,\frac{1}{2})$*, the weakly singular integral equation*

$$g(s,t) + \beta_H t^{-2H} \int_0^t g(r,t) \bar{\kappa}\left(\frac{r}{t},\frac{s}{t}\right) dr = c_H s^{1/2-H}(t-s)^{1/2-H}, \quad s \in [0,t], \tag{3.36}$$

with the kernel

$$\bar{\kappa}(u,v) = (uv)^{1/2-H} \int_{u\vee v}^1 r^{2H-1}(r-u)^{-1/2-H}(r-v)^{-1/2-H}\, dr. \tag{3.37}$$

Proof. By Lemma 3.1, applied to $\eta := M_t$, a function $h(\cdot,t) \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$ exists, so that

$$M_t = \mathbf{E}(B_t|\mathcal{F}_t^X) = \int_0^t h(s,t)dX_s, \quad \mathbf{P}-a.s.$$

To verify the representation (3.3), we have to check that $h(s,t)$ uniquely solves each one of the equations in (i)-(iv). To this end, we will argue that $h(s,t)$ satisfies the equation from (ii) for almost every $s \in [0,t]$. Then we show that this equation reduces to (iii) for $H > \frac{1}{2}$ and to (iv) for $H < \frac{1}{2}$, which are well known to have unique continuous solutions and therefore $h(s,t)$ must satisfy (ii) for all $s \in [0,t]$. Finally we will argue that (ii) and (i) share the same solution.

For any test function $\varphi \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$, the orthogonality property of the conditional expectation implies

$$\begin{aligned} 0 = \mathbf{E}\left(B_t - \int_0^t h(s,t)dX_s\right)\int_0^t \varphi(s)dX_s = \\ \int_0^t \varphi(s)ds - \int_0^t \varphi(s)h(s,t)ds - \frac{2-2H}{\lambda_H}\int_0^t s^{1-2H}K_h(s,t)K_\varphi(s,t)ds = \\ \int_0^t K_\varphi(s,t)\left(c_H Q_1(s)ds - c_H Q_h(s,t) - \frac{2-2H}{\lambda_H} s^{1-2H} K_h(s,t)\right) ds, \end{aligned} \tag{3.38}$$

where we used the identity (3.30). Since φ can be an arbitrary differentiable function, $h(s,t)$ satisfies (3.35) for almost all $s \in [0,t]$.

Applying the transformation (3.28) with $H > \frac{1}{2}$ to equation (3.35), a direct calculation shows that $h(s,t)$ satisfies (3.5). This weakly singular equation is well known to have unique solution (see, e.g., [72]), continuous on $[0,t]$. Since the transformation (3.28) is invertible, $h(s,t)$ is also the unique continuous solution of (3.35) for $H > \frac{1}{2}$.

Similarly, for $H < \frac{1}{2}$, applying the invertible transformation

$$\psi \mapsto -\frac{d}{ds}\int_s^t (r-s)^{1/2-H} r^{2H-1}\psi(r)dr \tag{3.39}$$

to (3.35), it can be seen that $h(s,t)$ satisfies (3.36).

Changing the integration variable in (3.37) to $x := \frac{1-v}{u-v}\frac{r-u}{1-r}$ we get

$$\bar{\kappa}(u,v) = |u-v|^{-2H} M(u,v), \quad u,v \in (0,1),$$

where

$$M(u,v) = \left(\frac{a}{b}\right)^{\frac{1}{2}-H} \int_0^\infty x^{-\frac{1}{2}-H}(1+x)^{-\frac{1}{2}-H}\left(1+\left(1-\frac{a}{b}\right)x\right)^{2H-1} dx,$$

with

$$a = \frac{u}{1-u} \wedge \frac{v}{1-v}, \quad b = \frac{u}{1-u} \vee \frac{v}{1-v}.$$

For $H < \frac{1}{2}$ the function $M(a,b)$ is bounded and thus $\bar{\kappa}(u,v)$ is a weakly singular kernel. Moreover, since the right hand side of (3.36) is a continuous function for $H < \frac{1}{2}$, this equation has a unique solution, continuous on $[0,t]$. This completes the proof of (iv) and, in turn, of (ii).

Further, the identity (3.40) from Lemma 3.2 below and the orthogonality property (3.38) imply

$$0 = \int_0^t \varphi(s)ds - \int_0^t \varphi(s)h(s,t)ds - \frac{2-2H}{\lambda_H}\int_0^t s^{1-2H}K_h(s,t)K_\varphi(s,t)ds =$$
$$\int_0^t \varphi(s)\left(1-h(s,t)-H\frac{d}{ds}\int_0^t g(r,t)|s-r|^{2H-1}sign(s-r)dr\right)ds.$$

The assertion (i) follows, in view of arbitrariness of φ and unique solvability of (3.35).

Finally, for $t \in [0,T]$,

$$\langle M\rangle_t = \mathbf{E}M_t^2 = \mathbf{E}(M_tB_t) = \mathbf{E}\left(B_t\int_0^t g(s,t)dX_s\right) = \int_0^t g(s,t)ds.$$

□

The following lemma proves the identity, we used in the proof of Theorem 3.5:

Lemma 3.2. *For any $\phi, \psi \in L^2([0,t]) \bigcap \Lambda_t^{H-\frac{1}{2}}$*

$$\frac{2-2H}{\lambda_H}\int_0^t s^{1-2H}K_\phi(s,t)K_\psi(s,t)ds =$$
$$H\int_0^t \phi(r)\frac{d}{dr}\int_0^t \psi(u)|r-u|^{2H-1}sign(r-u)du\,dr \tag{3.40}$$

Proof. For $H > \frac{1}{2}$ the identity (3.40) follows directly from (3.32) and (3.33). To prove it for $H < \frac{1}{2}$, let us first show that

$$\frac{d}{ds}\int_0^s \tau^{1/2-H}(s-\tau)^{1/2-H}L_\psi(\tau,t)d\tau = \frac{2-2H}{\lambda_H c_H}s^{1-2H}K_\psi(s,t), \tag{3.41}$$

where $L_\psi(\tau,t) := H\frac{d}{d\tau}\int_0^t \psi(u)|\tau-u|^{2H-1}sign(\tau-u)du$. To this end,

$$\begin{aligned}
&\int_0^s \tau^{1/2-H}(s-\tau)^{1/2-H}L_\psi(\tau,t)d\tau = \\
&-H\int_0^t \psi(u)\int_0^s \frac{d}{d\tau}\left(\tau^{1/2-H}(s-\tau)^{1/2-H}\right)|\tau-u|^{2H-1}sign(\tau-u)d\tau du \overset{\dagger}{=} \\
&\frac{1}{c_H}\int_0^s \psi(r)dr + \frac{2H(2-2H)(H-1/2)}{c_H\lambda_H}\int_s^t \psi(r)r^{H-1/2}\int_0^s v^{1-2H}(r-v)^{H-3/2}dv\,dr = \\
&\frac{1}{c_H}\int_0^s \psi(r)dr - \frac{2H(2-2H)}{c_H\lambda_H}\int_s^t \psi(r)r^{H-1/2}s^{1-2H}(r-s)^{H-1/2}dr + \\
&\frac{2H(2-2H)(1-2H)}{c_H\lambda_H}\int_s^t \psi(r)r^{H-1/2}\int_0^s v^{-2H}(r-v)^{H-1/2}dv\,dr,
\end{aligned}$$

where the equality † holds by Proposition 2.1 from [56] and the other two equalities hold by integration by parts. The identity (3.41) is obtained by taking the derivative of both sides.

The formula (3.40) now follows from (3.41):

$$\begin{aligned}
&\frac{2-2H}{\lambda_H}\int_0^t K_\phi(s,t)K_\psi(s,t)s^{1-2H}ds = \\
&c_H\big(1/2-H\big)\int_0^t K_\phi(s,t)\int_0^s (s-\tau)^{-1/2-H}\tau^{1/2-H}L_\psi(\tau,t)d\tau ds = \\
&c_H\big(1/2-H\big)\int_0^t L_\psi(\tau,t)\tau^{1/2-H}\int_\tau^t (s-\tau)^{-1/2-H}K_\phi(s,t)ds = \\
&\int_0^t L_\psi(\tau,t)\phi(\tau)d\tau = H\int_0^t \phi(r)\frac{d}{dr}\int_0^t \psi(u)|r-u|^{2H-1}sign(r-u)du\,dr,
\end{aligned}$$

where we used (3.27). □

3.4 The integral equation (3.5) with $H > \frac{1}{2}$

In this section we derive several useful properties of the family of solution $\{g(s,t) : 0 \le s \le t\}$, $t \in [0,T]$ of the equation (3.5), assuming $H > \frac{1}{2}$.

Properties of $g(s,t)$ on the diagonal

Lemma 3.3. *The function $g(t,t)$, $t \in [0,T]$ satisfies the properties:*

i. $g(t,t)$ is continuous on $[0,T]$ with $g(0,0) := \lim_{t\to 0} g(t,t) = 1$

ii. $g(t,t) > 0$ for all $t \in [0,T]$.

iii. for all $t \in [0, T]$,

$$\int_0^t g(s,t)ds = \int_0^t g^2(s,s)ds. \tag{3.42}$$

Proof. (i) Let n_0 be the least integer greater than $\frac{1}{4H-2}$ and note that $\kappa^{(n_0)}(\cdot,t) \in L^2([0,t])$. Iterating the equation (3.5), we get

$$g(t,t) = 1 + \sum_{m=1}^{n_0-1}(-1)^m \int_0^t \kappa^{(m)}(r,t)dr + (-1)^{n_0}\int_0^t \kappa^{(n_0)}(r,t)g(r,t)dr$$

and

$$\big|g(t,t)-1\big| \le C_1 \sum_{m=1}^{n_0-1} t^{(2H-1)m} + \left(\int_0^t \left(\kappa^{(n_0)}(r,t)\right)^2 dr\right)^{1/2}\left(\int_0^t g^2(r,t)dr\right)^{1/2}. \tag{3.43}$$

Multiplying (3.5) by $g(s,t)$, integrating and using positive definiteness of the kernel κ, we get

$$\int_0^t g^2(s,t)ds \le \int_0^t g(s,t)ds \le t^{1/2}\left(\int_0^t g^2(s,t)ds\right)^{1/2},$$

that is, $\left(\int_0^t g^2(s,t)ds\right)^{1/2} \le t^{1/2}$. Plugging this back into (3.43) gives $\lim_{t\to 0} g(t,t) = 1$. Continuity of $g(t,t)$ on $(0,T]$ follows from continuity of $r \mapsto g(r,t)$ for all $r \in [0,t]$ and differentiability of $g(r,t)$ in t for any $r \in (0,t)$, guaranteed by Lemma 3.4 below.

(ii) The function $g(s,t)$ is differentiable at $s \in (0,t)$ (see, e.g., [72]). Letting $g'(s,t) := \frac{\partial}{\partial s} g(s,t)$ and taking the derivative of (3.5), we obtain

$$g'(s,t) = -\frac{\partial}{\partial s}\int_0^t g(r,t)\kappa(r,s)dr = -\frac{\partial}{\partial s}\left(\int_{-s}^{t-s} g(u+s,t)\kappa(u,0)du\right) =$$
$$-\int_0^t g'(r,t)\kappa(r,s)dr + g(t,t)\Big(\kappa(s,t)-\kappa(s,0)\Big),$$

where we used the obvious symmetry $g(t-s,t) = g(s,t)$ and $g(t,t) = g(0,t)$ in particular. Now suppose $g(t,t) = 0$ for some $t > 0$. Then

$$g'(s,t) + \int_0^t g'(r,t)\kappa(r,s)dr = 0, \quad s \in [0,t].$$

This equation has the unique solution $g'(s,t) \equiv 0$, i.e., $g(s,t)$ is a constant function. But since $g(t,t) = 0$, it follows that $g(s,t) = 0$ for all $s \in [0,t]$, which contradicts (3.5). Hence $g(t,t) \neq 0$ for all $t \ge 0$ and, in fact $g(t,t) > 0$, as $g(0,0) = 1$.

(iii) Next, multiplying (3.5) by $g(s,t)$ and integrating we obtain

$$\int_0^t g^2(s,t)ds + \int_0^t g(s,t)\int_0^t g(r,t)\kappa(r,s)drds = \int_0^t g(s,t)ds$$

and hence

$$\frac{d}{dt}\int_0^t g(s,t)ds = g^2(t,t) + 2g(t,t)\int_0^t g(r,t)\kappa(r,t)dr +$$
$$2\int_0^t \dot g(s,t)\left(g(s,t)+\int_0^t g(r,t)\kappa(r,s)dr\right)ds =$$
$$g^2(t,t)+2g(t,t)\Big(1-g(t,t)\Big)+2\int_0^t \dot g(s,t)ds =$$
$$-g^2(t,t)+2\left(g(t,t)+\int_0^t \dot g(s,t)ds\right) = -g^2(t,t)+2\frac{d}{dt}\int_0^t g(s,t)ds.$$

This implies $g^2(t,t)=\frac{d}{dt}\int_0^t g(s,t)ds$ and in turn (3.42).

□

Properties of $\dot g(s,t)=\dfrac{\partial}{\partial t}g(s,t)$

Lemma 3.4. *The solution $g(s,t)$ of (3.5) satisfies the following properties*

i. $g(s,t)$ is continuously differentiable at $t\in(0,T]$ for any $s>0$, $s\neq t$;

ii. the derivative $\dot g(s,t):=\dfrac{\partial}{\partial t}g(s,t)$ satisfies the equation

$$\dot g(s,t)+\int_0^t \dot g(r,t)\kappa(r,s)dr = -g(t,t)\kappa(s,t), \quad s\in(0,t),\ t>0. \tag{3.44}$$

iii. $\dot g(\cdot,t)\in L^2([0,t])$ for $H>3/4$.

Proof.

(i) The function $g_t(u):=g(ut,t)$, $u\in[0,1]$, $t>0$ satisfies the integral equation

$$g_t(u)+t^{2H-1}\int_0^1 g_t(v)\kappa(u,v)dv = 1, \quad u\in[0,1].$$

This equation has a unique continuous solution for any $t>0$ (see [72]) and in terminology of [64], any point $\lambda:=t^{2H-1}$ is regular. Since for $H>1/2$ the kernel belongs to $L_1([0,1])$, the corresponding operator maps $L^2([0,1])$ into itself (see, e.g., Theorem 9.5.1 in [22]). It follows from, e.g., Theorem on page 154 in [64], that the solution $g_t(u)$ is analytic at $t>0$. By [72] the solution $g_t(u)$ is continuously differentiable at $u\in(0,1)$ and hence the function $g(s,t)=g_t(s/t)$ is continuously differentiable at $t>0$ for any $s\in(0,t)$ and also for any $s>t$ by the extension (3.6).

(ii) The equation (3.44) is obtained by taking the derivative of both sides of (3.5).

(iii) Multiplying (3.44) by $\dot g(s,t)$, integrating and using positive definiteness of the kernel κ, we get

$$\int_0^t \dot g^2(s,t)ds \le -g(t,t)\int_0^t \dot g(s,t)\kappa(s,t)ds \le C_1\left(\int_0^t \dot g^2(s,t)ds\right)^{1/2}\left(\int_0^t s^{4H-4}ds\right)^{1/2}.$$

The right hand side is finite for $H>\frac{3}{4}$, which completes the proof.

□

Additional properties of $g(s,t)$

In this subsection we derive several properties, related to invertibility of the integral transform with the kernel $g(s,t)$, needed in the proof of (3.10).

Lemma 3.5. *The function* $R = \big\{R(t,s),\ 0 \le s,t \le T\big\}$, *defined in* (3.7), *satisfies the equation*

$$R(s,t) + \int_0^t R(r,t)\kappa(r,s)dr = -\kappa(s,t) \quad s,t \in [0,T],\ s \ne t, \tag{3.45}$$

and the identity

$$R(s,t) - R(t,s) = \int_s^t R(s,\tau)R(t,\tau)d\tau, \quad s < t. \tag{3.46}$$

Proof. The equation (3.45) follows from the definition of R and Lemma 3.4. To prove (3.46) we will use Krein's method of solving integral equations on a finite interval. Let $y(s,t)$ satisfy the equation

$$y(s,t) + \int_0^t y(r,t)\kappa(r,s)dr = \phi(s,t), \quad s \in (0,t),$$

where $\phi(\cdot,t)$ is an integrable function. Then

$$y(s,t) = F(s,t)g(s,s) + \int_s^t F(\tau,t)\dot{g}(s,\tau)d\tau, \tag{3.47}$$

where g is the solution of equation (3.5) and

$$F(\tau,t) = \frac{1}{g^2(\tau,\tau)}\frac{d}{d\tau}\int_0^\tau g(s,\tau)\phi(s,t)ds. \tag{3.48}$$

For the solution of (3.45), the formula (3.48) reads

$$F(\tau,t) = -\frac{1}{g^2(\tau,\tau)}\frac{\partial}{\partial\tau}\int_0^\tau g(r,\tau)\kappa(r,t)dr = -\frac{1}{g^2(\tau,\tau)}\frac{\partial}{\partial\tau}\Big(1 - g(t,\tau)\Big) = \frac{\dot{g}(t,\tau)}{g^2(\tau,\tau)},$$

and applying Krein's formula (3.47), we get

$$R(s,t) = \frac{\dot{g}(t,s)}{g(s,s)} + \int_s^t \frac{\dot{g}(t,\tau)}{g^2(\tau,\tau)}\dot{g}(s,\tau)d\tau = R(t,s) + \int_s^t R(s,\tau)R(t,\tau)d\tau.$$

□

Singular perturbations

Analysis of the large sample asymptotic of MLE in Theorem 3.4 leads to a singularly perturbed problem. Fix $\varepsilon > 0$ and let g_ε be the solution of the equation:

$$\varepsilon g_\varepsilon^{(\varphi)}(u) + \int_0^1 g_\varepsilon^{(\varphi)}(v)\kappa(u,v)dv = \varphi(u), \quad u \in [0,1], \tag{3.49}$$

where φ is a sufficiently smooth function. Let $g^{(\varphi)}$ be the solution of auxiliary integral equation of the first kind

$$\int_0^1 g^{(\varphi)}(v)\kappa(u,v)\,dv = \varphi(u). \tag{3.50}$$

The unique solution to the this equation is given by an explicit formula, which is not of immediate interest for our purposes. For example, in the special case of $\varphi \equiv 1$,

$$g^{(1)}(s) = c_H s^{\frac{1}{2}-H}(1-s)^{\frac{1}{2}-H}, \tag{3.51}$$

Clearly, $g^{(1)} \in L^2([0,1])$ for $H > \frac{1}{2}$.

As ε decreases, the first term on the left hand side of the equation (3.49) disappears and it degenerates to the equation (3.50). Hence the convergence $g_\varepsilon^{(\varphi)} \to g^{(\varphi)}$ as $\varepsilon \to 0$ should be expected. To this end, we have the following estimate:

Lemma 3.6. *Let $\psi(u)$ be a function, such that $g^{(\psi)}$ exists, then*

$$\left|\int_0^1 \big(g_\varepsilon^{(\varphi)}(s) - g^{(\varphi)}(s)\big)\psi(s)\,ds\right| \le 2\varepsilon\left(\int_0^1 \big(g^{(\psi)}(u)\big)^2 du\right)^{1/2}\left(\int_0^1 \big(g^{(\varphi)}(u)\big)^2\,du\right)^{1/2}.$$

Proof. The assertion of the lemma is trivial if either of the norms in the right hand is infinite. Otherwise, $g_\varepsilon^{(\varphi)} \in L^2([0,1])$ and $\delta_\varepsilon := g_\varepsilon^{(\varphi)} - g^{(\varphi)}$ satisfies

$$\varepsilon\delta_\varepsilon(u) + \int_0^1 \delta_\varepsilon(v)\kappa(v,u)dv = -\varepsilon g^{(\varphi)}(u).$$

Multiplying by δ_ε and integrating we obtain

$$\varepsilon\int_0^1 \delta_\varepsilon^2(u)du + \int_0^1\int_0^1 \delta_\varepsilon(u)\delta_\varepsilon(v)\kappa(u,v)dudv = \varepsilon\left|\int_0^1 g^{(\varphi)}(u)\delta_\varepsilon(u)du\right|,$$

and, in particular,

$$\int_0^1 \delta_\varepsilon^2(u)du \le \left|\int_0^1 g^{(\varphi)}(u)\delta_\varepsilon(u)du\right|.$$

On the other hand, by the Cauchy−Schwarz inequality,

$$\left|\int_0^1 g^{(\varphi)}(u)\delta_\varepsilon(u)du\right|^2 \le \int_0^1 \big(g^{(\varphi)}(u)\big)^2du\int_0^1 \delta_\varepsilon^2(u)du$$

and hence

$$\int_0^1 \delta_\varepsilon^2(u)du \le \int_0^1 \big(g^{(\varphi)}(u)\big)^2du. \tag{3.52}$$

The function δ_ε also satisfies

$$\varepsilon g_\varepsilon^{(\varphi)}(u) + \int_0^1 \delta_\varepsilon(v)\kappa(u,v)dv = 0,$$

and hence for any ψ such that $g^{(\psi)} \in L^2([0,1])$

$$\left|\int_0^1 \delta_\varepsilon(u)\psi(u)du\right| = \left|\int_0^1 \delta_\varepsilon(u)\int_0^1 g^{(\psi)}(v)\kappa(u,v)dvdu\right| =$$
$$\left|\int_0^1 g^{(\psi)}(v)\int_0^1 \delta_\varepsilon(u)\kappa(u,v)dudv\right| = \varepsilon\left|\int_0^1 g^{(\psi)}(u)g_\varepsilon^{(\varphi)}(u)du\right| =$$
$$\varepsilon\left|\int_0^1 g^{(\psi)}(u)\delta_\varepsilon(u)du + \int_0^1 g^{(\psi)}(u)g^{(\varphi)}(u)du\right| \le$$
$$2\varepsilon\left(\int_0^1 \left(g^{(\psi)}(u)\right)^2 du\right)^{1/2}\left(\int_0^1 \left(g^{(\varphi)}(u)\right)^2 du\right)^{1/2},$$

where we used (3.52). □

Remark 11. *The statement of Lemma 3.6 is valid for any symmetric nonnegative weakly singular kernel.*

Remark 12. *While the qualitative theory of integral equations with weakly singular kernels is quite mature (see, e.g., [60], [73]), singular perturbations of such equations, somewhat surprisingly, have never been addressed so far. Most of the available literature deals with singularly perturbed equations, whose kernels have mild discontinuities (see [69] and the references therein).*

If one fixes a function φ and thinks of ψ as a test function in the above lemma, its assertion can be interpreted as a particular type of weak convergence $g_\varepsilon^{(\varphi)} \to g^{(\varphi)}$ as $\varepsilon \to 0$. Such convergence is sufficient for the purposes of asymptotic analysis in the regression problem of Theorem 3.4. However, preliminary calculations show that in other problems, such as drift estimation of the mixed fractional Ornstein– Uhlenbeck process (3.16)*, stronger, pointwise limit is required. This type of convergence is apparently much harder to obtain and progress in this direction will be reported elsewhere.*

4 Mixed fBm for $H > \frac{1}{2}$

4.1 Proof of Theorem 3.1

The representation (3.9) holds by Theorem 3.5 and

$$\langle M\rangle_t = \int_0^t g(s,t)ds = \int_0^t g^2(s,s)ds,$$

where the last equality holds by (3.42).

To derive the representation (3.10), we will show that $\widehat{X}_t := \mathbf{E}(X_t|\mathcal{F}_t^M)$ coincides with X_t, $\mathbf{P}$-a.s. To this end, similarly to Lemma 3.1, there exists a square integrable function $H(s,t)$, $s \le t$ such that

$$\widehat{X}_t = \int_0^t H(s,t)dM_s, \quad t \in [0,T],$$

and by the normal correlation theorem

$$H(s,t) = \frac{1}{g^2(s,s)}\frac{\partial}{\partial s}\mathbf{E}X_t M_s = \frac{1}{g^2(s,s)}\frac{\partial}{\partial s}\left(\int_0^s g(r,s)\frac{\partial}{\partial r}\mathbf{E}X_t X_r dr\right) =$$
$$\frac{1}{g^2(s,s)}\frac{\partial}{\partial s}\left(\int_0^s g(r,s)\left(1+\int_0^t \kappa(\tau,r)d\tau\right)dr\right) \stackrel{\dagger}{=}$$
$$1+\frac{1}{g^2(s,s)}\frac{\partial}{\partial s}\int_0^s g(r,s)\int_0^t \kappa(\tau,r)d\tau dr =$$
$$1+\frac{1}{g^2(s,s)}\left(g(s,s)\int_0^t \kappa(\tau,s)d\tau + \int_0^t\int_0^s \dot{g}(r,s)\kappa(\tau,r)drd\tau\right) =$$
$$1-\frac{1}{g^2(s,s)}\int_0^t \dot{g}(\tau,s)d\tau = 1-\frac{1}{g(s,s)}\int_0^t R(\tau,s)d\tau = G(s,t),$$

where the equality † holds by (3.42). To prove the claim we will show that

$$\mathbf{E}\big(X_t-\widehat{X}_t\big)^2 = \mathbf{E}X_t^2 - \mathbf{E}\widehat{X}_t^2 = 0. \tag{3.53}$$

Since $X_0=\widehat{X}_0=0$, $\mathbf{P}$-a.s., (3.53) holds if

$$\frac{\partial^2}{\partial t\partial s}\int_0^{t\wedge s} G(r,t)G(r,s)d\langle M\rangle_r = \kappa(t,s), \quad s<t.$$

By (3.9), the latter holds if

$$\dot{G}(s,t)G(s,s)g^2(s,s) + \int_0^s \dot{G}(r,t)\dot{G}(r,s)g^2(r,r)dr = \kappa(t,s). \tag{3.54}$$

By the definition (3.8) and (3.42)

$$G(t,t) = 1-\frac{1}{g^2(t,t)}\int_0^t \dot{g}(s,t)ds = 1-\frac{1}{g^2(t,t)}\left(\frac{d}{dt}\int_0^t g(s,t)ds - g(t,t)\right) =$$
$$1-\frac{1}{g^2(t,t)}\big(g^2(t,t)-g(t,t)\big) = \frac{1}{g(t,t)},$$

and, since $\dot{G}(s,t)g(s,s) = -R(t,s)$, (3.54) reads

$$-R(t,s) + \int_0^s R(t,r)R(s,r)dr = \kappa(t,s). \tag{3.55}$$

Recall that the function R, satisfies the equation (3.45). Rearranging the terms, multiplying by $R(s,u)$ and integrating gives

$$\int_0^s R(t,u)R(s,u)du + \int_0^s \kappa(t,u)R(s,u)du = -\int_0^s\int_0^u R(r,u)R(s,u)\kappa(r,t)drdu =$$
$$-\int_0^s\left(\int_r^s R(r,u)R(s,u)du\right)\kappa(r,t)dr = -\int_0^s \Big(R(r,s)-R(s,r)\Big)\kappa(r,t)dr$$

where we used Lemma 3.5. The second term on the left hand side and the last term on the right hand side cancel out and we get

$$\int_0^s R(t,u)R(s,u)du = -\int_0^s R(r,s)\kappa(r,t)dr = R(t,s)+\kappa(s,t),$$

which verifies (3.55) and therefore (3.53), thus completing the proof. □

4.2 Proof of Corollary 3.2

The representation (3.13) is obvious in view of (3.3) and the definition (3.14). To prove the inversion formula (3.12) we should check that

$$\int_0^t f(s)ds = \int_0^t G(s,t)\Phi(s)\, d\langle M\rangle_s, \quad t\in[0,T]. \tag{3.56}$$

Since this is a pathwise statement, no generality will be lost if f is assumed deterministic. But for a deterministic $f\in L^2([0,t])$ we have

$$\mathbf{E}\left(\int_0^t f(s)dB_s\Big|\mathcal{F}_t^X\right) = \mathbf{E}\left(\int_0^t f(s)dB_s\Big|\mathcal{F}_t^M\right) = \int_0^t \frac{d}{d\langle M\rangle_s}\left(\mathbf{E}M_s\int_0^s f(r)dB_r\right)dM_s = \int_0^t \frac{d}{d\langle M\rangle_s}\left(\mathbf{E}\int_0^s g(r,t)dX_r\int_0^s f(r)dB_r\right)dM_s = \int_0^t \Phi(s)dM_s,$$

and, using the representation (3.10), we obtain (3.56):

$$\int_0^t f(s)ds = \mathbf{E}X_t\int_0^t f(s)dB_s = \mathbf{E}X_t\int_0^t \Phi(s)dM_s = \int_0^t G(s,t)\Phi(s)d\langle M\rangle_s.$$

The formula (3.15) follows from Theorem 7.13 in [?], once we check

$$\int_0^T \Phi^2(\tau)\, d\langle M\rangle_\tau = \int_0^T \Phi^2(\tau)g^2(\tau,\tau)\, d\tau < \infty, \quad \mathbf{E}-a.s \tag{3.57}$$

and

$$\mathbf{E}\int_0^T \big|\Phi(\tau)\big|\, d\langle M\rangle_\tau < \infty. \tag{3.58}$$

By the definition (3.14) and continuity of f

$$\Phi(\tau)g(\tau,\tau) = f(\tau) + \int_0^\tau R(s,\tau)f(s)\, ds,$$

where R is given by (3.7). Let m_0 be the least integer greater than $\frac{1}{2H-1}$ and define

$$\bar{R}(s,\tau) := R(s,\tau) - \sum_{m=1}^{m_0-1}\kappa^{(m)}(s,\tau).$$

Since R solves the equation (3.45), the function $\bar{R}$ is the unique solution of

$$\bar{R}(s,\tau) + \int_0^\tau \bar{R}(r,\tau)\kappa(r,s)dr = -\kappa^{(m_0)}(s,\tau).$$

By the choice of m_0, the right hand side is a continuous function and hence $\bar{R}$ is uniformly bounded. Consequently, $|R(s,\tau)|\le C_1|s-\tau|^{2H-2}$ with a constant C_1 and

$$\left|\int_0^\tau R(s,\tau)f(s)\, ds\right| \le \left(\int_0^\tau \big|R(s,\tau)\big|\, f^2(s)\, ds\right)^{1/2}\left(\int_0^\tau \big|R(s,\tau)\big|\, ds\right)^{1/2} \le C_2\left(\int_0^T \big|R(s,\tau)\big|\, f^2(s)\, ds\right)^{1/2}$$

where $C_2^2 = C_1 \sup_{\tau\in[0,T]} \int_0^T |s-\tau|^{2H-2}\,ds$. Hence

$$\int_0^T \Phi^2(\tau)g^2(\tau,\tau)\,d\tau \le 2\int_0^T f^2(\tau)d\tau + 2\int_0^T \left(\int_0^\tau R(s,\tau)f(s)\,ds\right)^2 d\tau \le$$
$$2\int_0^T f^2(\tau)d\tau + 2C_2^2\int_0^T f^2(s)\int_0^T \big|R(s,\tau)\big|d\tau ds \le 2(1+C_2^4)\int_0^T f^2(\tau)d\tau < \infty,$$

which proves (3.57). The condition (3.58) is verified similarly:

$$\mathbf{E}\int_0^T \big|\Phi(\tau)\big|\,d\langle M\rangle_\tau \le C_3\mathbf{E}\int_0^T \big|f(\tau)\big|d\tau + C_3\mathbf{E}\int_0^T \big|f(s)\big|\int_0^T \big|R(s,\tau)\big|dsd\tau \le$$
$$C_3\big(1+C_2^2\big)\mathbf{E}\int_0^T \big|f(\tau)\big|d\tau < \infty$$

where $C_3 := \sup_{\tau\in[0,T]} g(\tau,\tau)$. □

5 Mixed fBm for $H < \frac{1}{2}$

5.1 Indirect approach

In this subsection we work out the details of the approach to analysis of the mixed fBm for $H < \frac{1}{2}$, outlined in Section 2.2.

Properties of the process $\widetilde{X}$

Consider the process

$$\widetilde{X}_t = \int_0^t \widetilde{\rho}(s,t)dX_s,$$

where

$$\widetilde{\rho}(s,t) = \sqrt{\beta_H}s^{1/2-H}\int_s^t \tau^{H-1/2}(\tau-s)^{-1/2-H}\,d\tau, \quad 0\le s\le t. \tag{3.59}$$

The process $\widetilde{X}$ admits the following decomposition:

Lemma 3.7. *$\widetilde{X} = \widetilde{B} + \widetilde{U}$, where $\widetilde{B}$ is an $\mathcal{F}$-Brownian motion and $\widetilde{U}$ is a centered Gaussian process with the covariance function, satisfying*

$$\widetilde{\kappa}(s,t) := \frac{\partial^2}{\partial s\partial t}\,\mathbb{E}\widetilde{U}_s\widetilde{U}_t = |t-s|^{-2H}\chi\left(\frac{s\wedge t}{s\vee t}\right), \quad s\ne t,$$

where

$$\chi(u) = \beta_H u^{1/2-H}L\left(\frac{u}{1-u}\right), \quad u\in[0,1], \tag{3.60}$$

and

$$L(v) = \int_0^v r^{-1/2-H}(1+r)^{-1/2-H}\left(1-\frac{r}{v}\right)^{1-2H} dr.$$

Moreover, $\mathcal{F}_t^X = \mathcal{F}_t^{\widetilde{X}}$, $\mathbf{P}$-a.s. for all $t\in[0,T]$.

Proof. It is well known (see, e.g., [56]), that the integral transformation with kernel $\widetilde{\rho}(s,t)$, defined in (3.59), is invertible:

$$X_t = \int_0^t \rho(s,t)d\widetilde{X}_s, \quad t \in [0,T], \tag{3.61}$$

where

$$\rho(s,t) = \frac{1}{2H}\sqrt{\frac{2-2H}{\lambda_H}}s^{1/2-H}K_1(s,t)$$

In particular, $\mathcal{F}_t^X = \mathcal{F}_t^{\widetilde{X}}$, $t \in [0,T]$.

Further, it follows from [56] that the process $\widetilde{B}_t = \int_0^t \widetilde{\rho}(s,t)dB_s^H$ is an $\mathcal{F}^{B^H}$-Brownian motion. Hence $\widetilde{X} = \widetilde{B} + \widetilde{U}$ with

$$\widetilde{U}_t = \int_0^t \widetilde{\rho}(s,t)dB_s.$$

Plugging in the expression for $\widetilde{\rho}(s,t)$, we get

$$\widetilde{\kappa}(s,t) = \frac{\partial^2}{\partial s \partial t}\mathbf{E}\widetilde{U}_t\widetilde{U}_s = \frac{\partial^2}{\partial s \partial t}\int_0^{s\wedge t}\widetilde{\rho}(r,s)\widetilde{\rho}(r,t)dr = \int_0^{s\wedge t}\dot{\widetilde{\rho}}(r,s)\dot{\widetilde{\rho}}(r,t)dr,$$

where $\dot{\widetilde{\rho}}(s,t) = \frac{\partial}{\partial t}\widetilde{\rho}(s,t)$ and we used the property $\widetilde{\rho}(s,s) = 0$. The expression in (3.17) is obtained by direct calculation, using the expression (3.59).

□

Properties of the equation (3.19)

For $H < \frac{1}{2}$ the function χ is continuous on $[0,1]$ and hence the kernel $\widetilde{\kappa}$, defined in (3.17), has a weak singularity. Consequently, the equation (3.19) has unique continuous solution. All the results of Section 3.4, except for (ii) of Lemma 3.3, have been derived without using the difference structure of the kernel κ and hence remain valid for $\widetilde{g}(s,t)$ with the obvious adjustments. The proof of (ii) of Lemma 3.3 requires a different argument:

Lemma 3.8. *$\widetilde{g}(t,t) > 0$ for all $t \in [0,T]$.*

Proof. As before, we will show that the assumption $\widetilde{g}(t,t) = 0$ for some $t \in [0,T]$ leads to a contradiction. To this end, changing the integration variable, the equation (3.19) can be rewritten as

$$\widetilde{g}(s,t) + s^{1-2H}\int_0^{t/s}\widetilde{g}(su,t)|1-u|^{-2H}\chi(u)\,du = 1$$

The solution $\widetilde{g}(s,t)$ is differentiable at $s \in (0,t)$ (see [72]) with $\widetilde{g}'(s,t) := \frac{\partial}{\partial s}\widetilde{g}(s,t)$ satisfying

$$0 = \widetilde{g}'(s,t) + (1-2H)s^{-2H}\int_0^{t/s}\widetilde{g}(su,t)|1-u|^{-2H}\chi(u)\,du \\ + s^{1-2H}\int_0^{t/s}u\widetilde{g}'(su,t)|1-u|^{-2H}\chi(u)du$$

where we used the assumption $\widetilde{g}(t,t)=0$. Multiplying by s and changing back the variables, we obtain

$$s\widetilde{g}'(s,t)+\int_0^t r\widetilde{g}'(r,t)\widetilde{\kappa}(r,s)dr=(2H-1)\big(1-\widetilde{g}(s,t)\big). \tag{3.62}$$

Multiplying (3.62) by $\widetilde{g}(s,t)$ and integrating, we get

$$t\widetilde{g}(t,t)=2H\int_0^t \widetilde{g}(s,t)ds+(1-2H)\int_0^t \widetilde{g}^2(s,t)ds=2H\int_0^t \widetilde{g}^2(s,s)ds+(1-2H)\int_0^t \widetilde{g}^2(s,t)ds$$

where the last equality holds by (iii). By continuity of $\widetilde{g}(s,t)$, the latter implies $\widetilde{g}(s,t)=0$ for all $s\in[0,t]$, which contradicts (3.19). □

The rest of the properties of $\widetilde{g}(s,t)$ are verified exactly as in the previous section:

Lemma 3.9. *The solution $\widetilde{g}(s,t)$ of* (3.19) *satisfies the properties* (i) *and* (ii) *(with κ replaced by $\widetilde{\kappa}$) of the Lemma 3.4 and $\dot{\widetilde{g}}(\cdot,t)\in L^2([0,t])$ for $H<\frac{1}{4}$.*

The assertion of Lemma 3.5 remains valid with $R(s,t)$ replaced by $\widetilde{R}(s,t)=\dfrac{\dot{\widetilde{g}}(s,t)}{\widetilde{g}(t,t)}$.

The representation formulas

Corollary 3.6. *The assertions of Theorem 3.1 and Corollary 3.2 hold with M, $g(s,t)$, $G(s,t)$ and Y_t replaced, respectively, by $\widetilde{M}=\mathbf{E}\big(\widetilde{B}_t|\mathcal{F}_t^X\big)$, $\widetilde{g}(s,t)$, $\widetilde{G}(s,t)$ (defined accordingly as in* (3.8)*) and*

$$\widetilde{Y}_t=\int_0^t \widetilde{\rho}(s,t)dY_s.$$

Proof. As explained above, the solution $\widetilde{g}(s,t)$ of (3.19) satisfies the same properties as the solution $g(s,t)$ of equation (3.5). Consequently the arguments, used in the proof of Theorem 3.1 and Corollary 3.2 apply to the process $\widetilde{X}$, rather than X itself. □

5.2 The martingale M for $H<\frac{1}{2}$

The analysis of mixed fBm in the previous sections was based on different martingales, depending on the range of H. For $H>\frac{1}{2}$ it is natural to work directly with the martingale M, since the general equations (3.4) and (3.35) reduce in this case to the simpler integral equation (3.5). For $H<\frac{1}{2}$, a similar integral equation (3.19) is obtained, if the martingale $\widetilde{M}$ from (3.18) is used instead.

In this subsection, we revisit the "direct" approach based on the martingale M in the case $H<\frac{1}{2}$. The obtained formulas involve single transformation with the kernel $g(s,t)$, rather than composition of two transformations with kernels $\widetilde{\rho}(s,t)$ and $\widetilde{g}(s,t)$. This can be somewhat more convenient in statistical applications. For $H<\frac{1}{2}$ the equation (3.4) reduces to the integral equation (3.36), from which analogs of Theorem 3.1 and Corollary 3.2 can be deduced directly.

Next lemma reveals that the two martingales are, in fact, closely related.

Lemma 3.10. *The martingales M and $\widetilde{M}$ are generated by the same innovation Brownian motion*

$$W_t = \int_0^t \frac{d\widetilde{M}_s}{\widetilde{g}(s,s)},$$

and $\mathcal{F}_t^M = \mathcal{F}_t^{\widetilde{M}}$, $\mathbf{P}$*-a.s.* $t \in [0,T]$.

Proof. Recall that $\mathcal{F}_t^{\widetilde{M}} = \mathcal{F}_t^X$, $\mathbf{P}$-a.s. and hence

$$M_t = \mathbf{E}\big(M_t|\mathcal{F}_t^{\widetilde{M}}\big) = \int_0^t \frac{d\langle M, \widetilde{M}\rangle_s}{d\langle \widetilde{M}\rangle_s} d\widetilde{M}_s =: \int_0^t q(s) d\widetilde{M}_s.$$

The assertion of the lemma follows, since

$$\mathbf{E}\big(\widetilde{M}_t|\mathcal{F}_t^M\big) = \int_0^t \frac{d\langle M, \widetilde{M}\rangle_s}{d\langle M\rangle_s} dM_s = \int_0^t \frac{1}{q^2(s)} \frac{d\langle M, \widetilde{M}\rangle_s}{d\langle \widetilde{M}\rangle_s} q(s) d\widetilde{M}_s = \widetilde{M}_t.$$

□

The structure of the martingale M for $H < \frac{1}{2}$ and its relation to the process $\widetilde{X}$ are elaborated in the following lemma:

Lemma 3.11. *For $H < \frac{1}{2}$ and $t \in [0,T]$,*

$$M_t = \int_0^t p(s,t) d\widetilde{X}_s, \quad \langle M\rangle_t = \int_0^t p^2(s,s) ds, \tag{3.63}$$

where

$$p(s,t) := \sqrt{\frac{2-2H}{\lambda_H}} s^{1/2-H} K_g(s,t) \tag{3.64}$$

solves the equation (cf. (3.19)*)*

$$p(s,t) + \int_0^t p(r,t)\widetilde{\kappa}(r,s) dr = \sqrt{\frac{2-2H}{\lambda_H}} s^{1/2-H}, \quad 0 \le s \le t \le T. \tag{3.65}$$

Proof. The equation (3.65) is obtained from equation (3.35) by replacing g in the first term with (see (3.27))

$$g(s,t) = -c_H s^{\frac{1}{2}-H} \frac{d}{ds} \int_s^t K_g(r,t)(r-s)^{\frac{1}{2}-H} dr =$$
$$- c_H \sqrt{\frac{\lambda_H}{2-2H}} s^{\frac{1}{2}-H} \frac{d}{ds} \int_s^t p(r,t) r^{H-\frac{1}{2}} (r-s)^{\frac{1}{2}-H} dr.$$

Indeed

$$c_H Q_g(s) = -c_H^2 \sqrt{\frac{\lambda_H}{2-2H}} \frac{d}{ds} \int_0^s r^{1-2H}(s-r)^{\frac{1}{2}-H} \frac{d}{dr} \int_r^t p(u,t) u^{H-\frac{1}{2}} (u-r)^{\frac{1}{2}-H} du dr =$$
$$c_H^2 \sqrt{\frac{\lambda_H}{2-2H}} \big(1/2-H\big)^2 \int_0^s r^{1-2H}(s-r)^{-\frac{1}{2}-H} \int_r^t p(u,t) u^{H-\frac{1}{2}} (u-r)^{-\frac{1}{2}-H} du dr =$$
$$s^{\frac{1}{2}-H} \sqrt{\frac{2-2H}{\lambda_H}} \int_0^t p(u,t) \beta_H (su)^{H-\frac{1}{2}} \int_0^{s\wedge u} r^{1-2H}(s-r)^{-\frac{1}{2}-H}(u-r)^{-\frac{1}{2}-H} dr du =$$
$$s^{\frac{1}{2}-H} \sqrt{\frac{2-2H}{\lambda_H}} \int_0^t p(u,t) \widetilde{\kappa}(u,s) du,$$

where we used the definition (3.17) of $\widetilde{\kappa}$. The equation (3.65) follows, since

$$c_H Q_1(s) = \frac{2-2H}{\lambda_H} s^{1-2H}.$$

Let us now prove the representation formulas in (3.63). Being the solution of weakly singular equation (3.65), the function $p(s,t)$ is differentiable with respect to the first variable and the derivative $p'(s,t) = \frac{\partial}{\partial s} p(s,t)$ satisfies the equation

$$p'(s,t) + \int_0^t p'(r,t)\widetilde{\kappa}(r,s)dr = p(t,t)\Big(\widetilde{\kappa}(s,t) - \widetilde{\kappa}(s,0)\Big) + \sqrt{\frac{2-2H}{\lambda_H}}(1/2-H)s^{-1/2-H}.$$

The right hand side is an integrable function and so is the derivative $p'(s,t)$, $s \in (0,t)$. The first identity in (3.63) now follows by integration by parts:

$$\int_0^t p(s,t)d\widetilde{X}_s = p(t,t)\widetilde{X}_t - \int_0^t \widetilde{X}_s p'(s,t)ds = \int_0^t \int_r^t \widetilde{\rho}'(r,s)p(s,t)dsdX_r = \int_0^t g(s,t)dX_r,$$

where the last equality holds by direct calculation, using the definitions (3.59) and (3.64).

The second identity in (3.63) is obtained, using the identity (3.30):

$$\langle M \rangle_t = \int_0^t g(s,t)ds = \frac{2-2H}{\lambda_H}\int_0^t s^{1-2H} K_g(s,t)ds =$$
$$\sqrt{\frac{2-2H}{\lambda_H}}\int_0^t p(s,t)s^{1/2-H}ds = \int_0^t p^2(s,s)ds,$$

where the last equality is verified as in (iii) of Lemma 3.3. □

The following theorem generalizes Theorem 3.1 to all $H \in (0,1]$:

Theorem 3.7. *The $\mathcal{F}^X$-martingale M, defined in* (3.2)*, satisfies* (3.3) *and*

$$\langle M \rangle_t = \int_0^t \left(g^2(s,s) + \frac{2-2H}{\lambda_H}\left(s^{1/2-H}K_g(s,s)\right)^2 \right) ds, \tag{3.66}$$

where $g(s,t)$ is the unique solution of equation (3.4) *(or, equivalently,* (3.35)*). Moreover,*

$$X_t = \int_0^t G(s,t)dM_s, \quad t \in [0,T], \tag{3.67}$$

with

$$G(s,t) := 1 - \frac{d}{d\langle M \rangle_s}\int_0^t g(\tau,s)d\tau, \quad 0 \le s \le t \le T, \tag{3.68}$$

where $g(\tau,s)$ is defined as in (3.6) *for $\tau > s$. In particular, $\mathcal{F}_t^X = \mathcal{F}_t^M$, $\mathbf{P}$-a.s. $t \in [0,T]$.*

Proof. The representation (3.3) holds for all $H \in (0,1]$ by Theorem 3.5. For $H > \frac{1}{2}$,

$$K_g(s,t) = 2H(H-1/2)\int_s^t g(r,t)r^{H-1/2}(r-s)^{H-3/2}\,dr$$

and hence $K_g(t,t) = 0$ for all $t \geq 0$ and (3.66) reduces to the assertion (3.9) of Theorem 3.1. For $H < \frac{1}{2}$, the kernel $\bar{\kappa}(u,v)$ in (3.36) satisfies $\bar{\kappa}(u,1) = \bar{\kappa}(u,0) = 0$ and therefore $g(t,t) = 0$, $t \geq 0$ and (3.66) holds by Lemma 3.11.

Since, by Lemma 3.10, $\mathcal{F}_t^M = \mathcal{F}_t^{\widetilde{M}} = \mathcal{F}_t^X$, $\mathbf{P}$-a.s., it follows that $X_t = \mathbf{E}(X_t|\mathcal{F}_t^M)$, $\mathbf{P}$-a.s. and, similarly to Lemma 3.1, there exists a square integrable function $G(s,t)$, $s \leq t$ such that

$$X_t = \mathbf{E}(X_t|\mathcal{F}_t^M) = \int_0^t G(s,t)dM_s, \quad t \in [0,T].$$

By the normal correlation theorem

$$G(s,t) = \frac{d}{d\langle M\rangle_s}\mathbf{E}X_t M_s,$$

and the formula (3.68) holds, since

$$\mathbf{E}X_t M_s = \int_0^s g(r,s)\frac{\partial}{\partial r}\mathbf{E}X_t X_r dr = \int_0^s g(r,s)dr + H\int_0^s g(r,s)\big(r^{2H-1} + (t-r)^{2H-1}\big)dr =$$
$$\langle M\rangle_s + \int_0^t H\frac{d}{d\tau}\int_0^s g(r,s)|r-\tau|^{2H-1}sign(r-\tau)drd\tau = \langle M\rangle_s + \int_0^t (1-g(\tau,s))d\tau.$$

□

Corollary 3.8. *The assertion of Corollary 3.2 remains valid for all $H \in (0,1]$.*

Proof. Given the representation (3.67), the arguments from the proof of Corollary 3.2 apply for all $H \in (0,1]$ once we check (3.57) and (3.58) for $H < \frac{1}{2}$.

Since the kernel in (3.37) is weakly singular, as in (i) of Lemma 3.4, the solution $g(s,t)$ of (3.36) is differentiable with respect to the second (forward) variable. Taking the derivative of (3.35), we obtain

$$c_H Q_{\dot{g}}(s) + \frac{2-2H}{\lambda_H}K_{\dot{g}}(s,t)s^{1-2H} = 0, \quad 0 < s < t \leq T,$$

where the identity $g(t,t) = 0$ for $H < \frac{1}{2}$ was used. Applying the transformation (3.39), a direct calculation reveals that $\dot{g}(s,t)$ satisfies the equation (cf. (3.44)):

$$\dot{g}(s,t) + \beta_H t^{-2H}\int_0^t \dot{g}(r,t)\bar{\kappa}\left(\frac{r}{t},\frac{s}{t}\right)dr = p(t,t)\dot{\widetilde{\rho}}(s,t), \quad 0 \leq s \leq t \leq T,$$

where $\widetilde{\rho}(s,t)$ and $p(s,t)$ are defined in (3.59) and (3.64) respectively. Since $\dot{\widetilde{\rho}}(\cdot,t) \in L^1([0,t])$ and the kernel $\bar{\kappa}$ is weakly singular, (3.57) and (3.58) are now verified as in Corollary 3.2.

□

6 Semimartingale structure of X:proof

Here we will prove the theorem 3.3.

6.1 Proof of (1)

As mentioned in the introduction, B^H and hence also X have infinite quadratic variation for $H \in \big(0, \frac{1}{2}\big)$. Hence X is not a semimartingale in its own filtration and a fortiori μ^X and μ^W are singular. For $H = \frac{1}{2}$ the statement of the theorem is evident. Below we focus on the case $H \in \big(\frac{1}{2}, 1\big]$.

Remark 13. *The fact that X is not a semimartingale for $H \in (\frac{1}{2}, \frac{3}{4}]$ implies singularity of μ^X and μ^W, but not vise versa. For the sake of completeness, we prove both assertions directly, showing how they stem from the same property of the kernel κ.*

Equivalence for $H \in (\frac{3}{4}, 1]$

By Theorem 3.1

$$\langle M \rangle_t = \int_0^t g(s,t)ds = \int_0^t g^2(s,s)ds, \quad t \in [0,T].$$

Hence by the Lévy theorem and Theorem 3.1, $W = (W_t)$, $0 \le t \le T$, given by equation (3.20), is a Brownian motion with respect to $\mathcal{F}^X$. On the other hand,

$$M_t = \int_0^t g(s,t)dX_s = \int_0^t g(s,s)dX_s + \int_0^t \Big(g(r,t) - g(r,r)\Big)dX_r =$$
$$\int_0^t g(s,s)dX_s + \int_0^t \int_r^t \dot{g}(r,s)dsdX_r = \int_0^t g(s,s)dX_s + \int_0^t \int_0^s \dot{g}(r,s)dX_r ds,$$

where the last equality holds since $\dot{g}(\cdot, s) \in L^2([0,s])$ (see Lemma 3.4). Hence

$$W_t = \int_0^t \frac{1}{g(s,s)} dM_s = X_t + \int_0^t \int_0^s \frac{\dot{g}(r,s)}{g(s,s)} dX_r ds =: X_t + \int_0^t \varphi_s(X)ds.$$

The desired claim follows from Girsanov's theorem (Theorem 7.7 in [49]), once we check

$$\int_0^T \mathbf{E}\varphi_t^2(W)dt < \infty \quad \text{and} \quad \int_0^T \mathbf{E}\varphi_t^2(X)dt < \infty. \tag{3.69}$$

Since $\varphi_t(\cdot)$ is additive and $X_t = B_t + B_t^H$, where B and B^H are independent, it is enough to check only the latter condition. By Lemma 3.5 the function $R(s,t) = \dfrac{\dot{g}(s,t)}{g(t,t)}$ satisfies (3.45) and hence for $H > 3/4$,

$$\mathbf{E}\varphi_t^2(X) = \mathbf{E}\left(\int_0^t R(r,t)dX_r\right)^2 =$$
$$\int_0^t R^2(s,t)ds + \int_0^t \int_0^t R(s,t)R(r,t)\kappa(r,s)drds =$$
$$\int_0^t R(s,t)\left(R(s,t) + \int_0^t R(r,t)\kappa(r,s)dr\right)ds =$$
$$-\int_0^t R(s,t)\kappa(s,t)ds \le \left(\int_0^t R^2(s,t)ds\right)^{1/2}\left(\int_0^t \kappa^2(s,t)ds\right)^{1/2} =$$
$$C_1\left(\int_0^t R^2(s,t)ds\right)^{1/2} t^{2H-3/2}.$$

Since the kernel is positive definite, multiplying (3.45) by $R(s,t)$ and integrating gives

$$\int_0^t R^2(s,t)ds \le -\int_0^t R(s,t)\kappa(s,t)ds \le C_1\left(\int_0^t R^2(s,t)ds\right)^{1/2} t^{2H-3/2},$$

and consequently

$$\left(\int_0^t R^2(s,t)ds\right)^{1/2} \le C_1 t^{2H-3/2}.$$

Plugging this bound back gives $\mathbf{E}\varphi_t^2(X) \le C_1^2 t^{4H-3}$ and in turn

$$\int_0^T \mathbf{E}\varphi_t^2(X)dt \le C_1^2\int_0^T t^{4H-3}dt = C_2T^{4H-2},$$

which verifies (3.69) and completes the proof. □

Singularity for $H \in \left(\frac{1}{2}, \frac{3}{4}\right]$

As shown in the previous section, the process

$$M_t = \int_0^t g(s,t)dX_s, \quad t \in [0,T]$$

is a martingale. Suppose there exists a probability measure $\mathbf{Q}$, equivalent to $\mathbf{P}$, so that X is a Brownian motion in its natural filtration. Since the semimartingale property is preserved under equivalent change of measure, M must be a semimartingale under $\mathbf{Q}$, or, equivalently, the process

$$L_t := \int_0^t g(s,t)dW_s,$$

where W is the Brownian motion defined in (3.20), must be a semimartingale under $\mathbf{P}$. We will argue that this is impossible for $H \le \frac{3}{4}$, arriving at a contradiction and thus proving the claim.

To this end, define

$$\psi(s,t) = -\int_s^t g(r,r)\sum_{m=1}^{n_0-1}(-1)^m\kappa^{(m)}(r,s)dr, \quad 0 < s < t \le T,$$

where n_0 is the least integer greater than $\frac{1}{4H-2}$. Note that $\psi(\cdot,t) \in L^2([0,t])$ and define

$$U_t := \int_0^t \psi(s,t)dW_s$$
$$V_t := \int_0^t \big(g(s,t) - g(s,s) + \psi(s,t)\big)dW_s.$$

Then

$$L_t = V_t + \int_0^t g(s,s)dW_s - U_t.$$

The second term is an $\mathcal{F}^X$-martingale and hence, to argue that L is not a semimartingale, it is enough to show that

i. U has zero quadratic variation, but unbounded first variation

ii. V has bounded first variation.

Proof of (i). To check this assertion we will need an estimate for the variance of increments of U. To this end, for any two points $t_1, t_2 \in [0,T]$, such that $0 < t_2 - t_1 < 1$,

$$\begin{aligned} \mathbf{E}\big(U_{t_2} - U_{t_1}\big)^2 =& \mathbf{E}\left(\int_{t_1}^{t_2} \psi(s,t_2)dW_s + \int_0^{t_1} \big(\psi(s,t_2) - \psi(s,t_1)\big)dW_s\right)^2 = \\ & \int_{t_1}^{t_2} \psi^2(s,t_2)ds + \int_0^{t_1} \big(\psi(s,t_2) - \psi(s,t_1)\big)^2 ds. \end{aligned} \tag{3.70}$$

To bound the first term, note that

$$\psi^2(s,t_2) \le \|g\|_\infty^2 n_0 \sum_{m=1}^{n_0-1} \left(\int_s^{t_2} \kappa^{(m)}(s,r)dr\right)^2 \le C_1 \sum_{m=1}^{n_0-1} (t_2-s)^{(4H-2)m} \le C_2(t_2-s)^{4H-2},$$

where $\|g\|_\infty = \sup_{r\le T} |g(r,r)| < \infty$, and consequently

$$\int_{t_1}^{t_2} \psi^2(s,t_2)ds \le C_3(t_2-t_1)^{4H-1}.$$

For the second term, we have

$$\begin{aligned} &\int_0^{t_1} \big(\psi(s,t_2) - \psi(s,t_1)\big)^2 ds = \int_0^{t_1} \left(\sum_{m=1}^{n_0-1} \int_{t_1}^{t_2} (-1)^m g(r,r)\kappa^{(m)}(s,r)dr\right)^2 ds = \\ &\sum_{m=1}^{n_0-1}\sum_{\ell=1}^{n_0-1} \int_0^{t_1}\int_{t_1}^{t_2}\int_{t_1}^{t_2} (-1)^{m+\ell} g(r,r)g(\tau,\tau)\kappa^{(m)}(s,r)\kappa^{(\ell)}(s,\tau)drd\tau ds. \end{aligned} \tag{3.71}$$

The dominating term in this sum corresponds to $m=1$, $\ell=1$:

$$\int_0^{t_1}\left(\int_{t_1}^{t_2} g(r,r)\kappa(r,s)dr\right)^2 ds.$$

We have

$$\begin{aligned} &\int_0^{t_1}\left(\int_{t_1}^{t_2} \kappa(r,s)dr\right)^2 ds = H^2 \int_0^{t_1} \Big((t_2-t_1+s)^{2H-1} - s^{2H-1}\Big)^2 ds = \\ &H^2(t_2-t_1)^{4H-1} \int_0^{\frac{t_1}{t_2-t_1}} \Big((1+u)^{2H-1} - u^{2H-1}\Big)^2 du. \end{aligned} \tag{3.72}$$

The increasing function

$$\gamma(y) := H^2 \int_0^y \Big((1+u)^{2H-1} - u^{2H-1}\Big)^2 du, \quad y \ge 0$$

satisfies

$$\lim_{y\to\infty} \gamma(y) = \gamma_H, \quad H \in \Big(\frac{1}{2}, \frac{3}{4}\Big)$$
$$\lim_{y\to\infty} \frac{\gamma(y)}{\log y} = \gamma_{3/4}, \quad H = \frac{3}{4},$$

with positive constants γ_H. By Lemma 3.3, $\inf_{r\le T} g(r,r) > 0$ and hence

$$c_4 \le \int_0^{t_1} \left(\int_{t_1}^{t_2} g(r,r)\kappa(s,r)dr\right)^2 \Big/ (t_2-t_1)^{4H-1}\gamma\Big(\frac{t_1}{t_2-t_1}\Big) \le C_4$$

with some positive constants c_4, C_4 for all sufficiently small $t_2 - t_1$. A similar calculation shows that the rest of the terms in (3.71) converge to zero as $t_2 - t_1 \to 0$ at a faster rate and assembling all parts together, we obtain

$$c_5 \le \mathbf{E}\big(U_{t_2} - U_{t_1}\big)^2 \Big/ (t_2-t_1)^{4H-1}\gamma\Big(\frac{t_1}{t_2-t_1}\Big) \le C_5. \tag{3.73}$$

Now let $0 = t_0 < t_1 < \ldots < t_n = T$ be an arbitrary partition, then for all $H \in (\frac{1}{2}, \frac{3}{4}]$

$$\mathbf{E}\sum_{i=1}^n \big(U_{t_i} - U_{t_{i-1}}\big)^2 \le C_5 \sum_{i=1}^n (t_i - t_{i-1})^{4H-1}\gamma\Big(\frac{T}{t_i - t_{i-1}}\Big) \le$$
$$C_6 \max_i (t_i - t_{i-1})^{4H-2} \log \frac{1}{t_i - t_{i-1}} \xrightarrow{n\to\infty} 0,$$

i.e., U has zero quadratic variation.

On the other hand, since the process U is Gaussian

$$\mathbf{E}\sum_{i=1}^n \big|U_{t_i} - U_{t_{i-1}}\big| \ge \sqrt{\frac{2}{\pi}} c_5 \sum_{i: t_i \ge T/2} (t_i - t_{i-1})^{2H-\frac{1}{2}} \gamma^{1/2}\Big(\frac{T/2}{t_i - t_{i-1}}\Big) \ge$$
$$c_6 \min_i (t_i - t_{i-1})^{2H-\frac{3}{2}} \gamma^{1/2}\Big(\frac{T/2}{t_i - t_{i-1}}\Big) \xrightarrow{n\to\infty} \infty,$$

which implies that U has unbounded first variation (see, e.g., Theorem 4 Ch. 4 §9 in [48]).

Proof of (ii). For $0 < s < t \le T$

$$\dot\psi(s,t) := \frac{\partial}{\partial t}\psi(s,t) = -g(t,t)\sum_{m=1}^{n_0-1}(-1)^m \kappa^{(m)}(s,t)$$

and hence

$$\int_0^t \dot\psi(s,t)\kappa(r,s)dr = -\int_0^t \left(g(t,t)\sum_{m=1}^{n_0-1}(-1)^m\kappa^{(m)}(s,t)\right)\kappa(r,s)dr =$$
$$-g(t,t)\sum_{m=1}^{n_0-1}(-1)^m\kappa^{(m+1)}(s,t) = g(t,t)\sum_{m=2}^{n_0}(-1)^m\kappa^{(m)}(s,t) =$$
$$g(t,t)\kappa(s,t) - \dot\psi(s,t) + (-1)^{n_0} g(t,t)\kappa^{(n_0)}(s,t)$$

Adding this equality to (3.44), we get

$$\Big(\dot{g}(s,t)+\dot{\psi}(s,t)\Big)+\int_0^t\Big(\dot{g}(r,t)+\dot{\psi}(r,t)\Big)\kappa(r,s)dr=(-1)^{n_0}g(t,t)\kappa^{(n_0)}(s,t)$$

By the choice of n_0, the right hand side is square integrable and so is the function $\dot{g}(s,t)+\dot{\psi}(s,t)$, $s\in(0,t)$. Since $\psi(s,s)=0$,

$$V_t=\int_0^t\big(g(s,t)-g(s,s)+\psi(s,t)\big)dW_s=\int_0^t\int_s^t\big(\dot{g}(s,r)+\dot{\psi}(s,r)\big)drdW_s=\int_0^t\int_0^r\big(\dot{g}(s,r)+\dot{\psi}(s,r)\big)dW_sdr,$$

and hence V has bounded first variation. □

X is not a semimartingale for $H\in\left(\frac{1}{2},\frac{3}{4}\right]$

By Lemma 3.1, $X_t=\int_0^t G(s,t)dM_s$, where the function $G(s,t)$ satisfies (3.8). Hence

$$X_t=M_t-\int_0^t\frac{1}{g(s,s)}\int_0^t R(\tau,s)d\tau dM_s=M_t-\int_0^t\int_0^t R(\tau,s)d\tau dW_s=$$
$$M_t-\int_0^t\int_0^s R(\tau,s)d\tau dW_s-\int_0^t\int_s^t R(\tau,s)d\tau dW_s=:M_t-N_t-U_t$$

where W is $\mathcal{F}^X$-adapted Brownian motion, defined by (3.20). Since M is an $\mathcal{F}^X$-martingale, X will not be an $\mathcal{F}^X$-semimartingale if we show that

a. N is a martingale

b. U has zero quadratic variation, but unbounded first variation

Proof of (a). Let n_0 be the least integer greater than $\frac{1}{4H-2}$. Then it follows from (3.45) that the function

$$Q(s,t):=\int_0^t R(r,t)\kappa^{(n_0-1)}(r,s)dr.$$

satisfies

$$Q(s,t)+\int_0^t Q(r,t)\kappa(r,s)dr=-\kappa^{(n_0)}(s,t),$$

and hence $Q(\cdot,t)\in L^2([0,t])$. Iterating the equation (3.45) we get

$$R(s,t)=\sum_{m=1}^{n_0-1}(-1)^m\kappa^{(m)}(s,t)+(-1)^{(n_0-1)}Q(s,t),\tag{3.74}$$

and

$$\Big|\int_0^s R(\tau,s)d\tau\Big|\le\sum_{m=1}^{n_0-1}\int_0^s\kappa^{(m)}(\tau,s)d\tau+\int_0^s Q(\tau,s)d\tau\le C_1s^{2H-1}.$$

Hence the function $s \mapsto \int_0^s R(\tau,s)d\tau$ is square integrable for all $H \in \left(\frac{1}{2},1\right)$ and so N is a martingale.

Proof of (b). Define $\phi(s,t) := \int_s^t R(\tau,s)d\tau$, then similarly to (3.70),

$$\mathbf{E}\big(U_{t_2}-U_{t_1}\big)^2 = \int_{t_1}^{t_2} \phi^2(s,t_2)ds + \int_0^{t_1} \big(\phi(s,t_2)-\phi(s,t_1)\big)^2 ds. \tag{3.75}$$

By (3.74)

$$\phi^2(s,t) \le C_1 \sum_{m=1}^{n_0-1} \left(\int_s^t \kappa^{(m)}(\tau,s)d\tau\right)^2 + C_1\left(\int_s^t Q(\tau,s)d\tau\right)^2 \le C_2|t-s|^{4H-2}$$

and hence the first term in (3.75) is bounded by

$$\int_{t_1}^{t_2} \phi^2(s,t_2)ds \le \int_{t_1}^{t_2} C_2(t_2-s)^{4H-2}ds \le C_3(t_2-t_1)^{4H-1}.$$

Further,

$$\int_0^{t_1} \big(\phi(s,t_2)-\phi(s,t_1)\big)^2 ds = \int_0^{t_1}\left(\int_{t_1}^{t_2} R(\tau,s)d\tau\right)^2 ds = \int_0^{t_1}\int_{t_1}^{t_2}\int_{t_1}^{t_2} R(\tau,s)R(r,s)d\tau dr ds.$$

Plugging in the expression (3.74), the dominating term is readily seen to be given by (3.72) and hence as in the previous section the bound (3.73) holds. The claim (b) now follows by the same argument.

6.2 Proof of (2)

Equivalence for $H < \frac{1}{4}$

By calculations as in Section 6.1,

$$\widetilde{W}_t = \widetilde{X}_t + \int_0^t \widetilde{\varphi}_s(\widetilde{X})ds,$$

where $\widetilde{W}$ is an $\mathcal{F}^{\widetilde{X}}$-Brownian motion and $\int_0^T \mathbf{E}\widetilde{\varphi}_t^2(\widetilde{X})dt < \infty$. Hence the measures $\mu^{\widetilde{X}}$ and $\mu^{\widetilde{W}}$ are equivalent and the derivative $\dfrac{d\mu^{\widetilde{X}}}{d\mu^{\widetilde{W}}}(\widetilde{X})$ equals the expression in the right hand side of (3.21). Then under the probability $\mathbf{Q}$, defined by

$$\frac{d\mathbf{Q}}{d\mathbf{P}} := \frac{d\mu^{\widetilde{W}}}{d\mu^{\widetilde{X}}}(\widetilde{X}), \tag{3.76}$$

the process $\widetilde{X}$ is a Brownian motion. By the inversion formula (3.61), the process X is an fBm with the Hurst exponent H. This proves the claimed equivalence and verifies the formula (3.21), since $\mathcal{F}_T^{\widetilde{X}} = \mathcal{F}_T^X$ $\mathbf{P}$-a.s. and therefore the random variable in (3.76) is $\mathcal{F}_T^X$-measurable.

Singularity for $H \ge \frac{1}{4}$

The claim is obvious for $H = \frac{1}{2}$. For $H > \frac{1}{2}$ the process X has positive quadratic variation and hence can not be equivalent to fBm with $H > \frac{1}{2}$, which has zero quadratic variation.

To prove singularity for $H \in [\frac{1}{4}, \frac{1}{2})$, suppose there is a probability $\mathbf{Q}$, equivalent to $\mathbf{P}$, under which X is an fBm with the Hurst exponent H in its own filtration. Then $\widetilde{X}_t = \int_0^t \widetilde{\rho}(s,t)dX_s$, with $\widetilde{\rho}(s,t)$ defined in (3.59), is a Brownian motion under $\mathbf{Q}$. By calculations as in Subsection 6.1, one can show that $\widetilde{X}$ is not a semimartingale for $H \in [\frac{1}{4}, \frac{1}{2})$, thus obtaining a contradiction.

7 Properties of Drift Estimation: proof

Here we will prove the theorem 3.4. Since μ^X is independent of θ, the likelihood function is given by (3.15) with $f(t) \equiv \theta$. In this case by Corollary 3.8

$$\Phi(t) = \frac{d}{d\langle M\rangle_t} \int_0^t g(s,t)\theta ds = \theta$$

and hence

$$L(Y;\theta) := \frac{d\mu^Y}{d\mu^X}(Y) = \exp\left\{\theta Z_T - \frac{\theta^2}{2}\langle M\rangle_T\right\}.$$

The unique maximizer is $\hat{\theta}_T = Z_T/\langle M\rangle_T$, which is the expression claimed in (3.23).

Notice that

$$\hat{\theta}_T = \frac{M_T + \theta\int_0^T g(s,T)ds}{\langle M\rangle_T} = \frac{M_T}{\langle M\rangle_T} + \theta$$

and thus $\hat{\theta}_T$ is normal and unbiased with the variance

$$\mathbf{E}\big(\hat{\theta}_T - \theta\big)^2 = \mathbf{E}\left(\frac{M_T}{\langle M\rangle_T}\right)^2 = \frac{1}{\langle M\rangle_T}, \tag{3.77}$$

which is the formula (3.24).

The asymptotic variance is calculated as follows.

7.1 Proof of (3.25)

Let $\varepsilon := T^{1-2H}$ and define $g_\varepsilon(u) := T^{2H-1}g(uT,T)$, $u \in [0,1]$. The function g_ε solves the equation (3.49) with $\varphi \equiv 1$ and

$$\langle M\rangle_T = \int_0^T g(s,T)ds = T^{2-2H}\int_0^1 g_\varepsilon(u)du. \tag{3.78}$$

Applying Lemma 3.6 with $\varphi = \psi \equiv 1$ and using the formulas (3.77) and (3.78), we obtain

$$T^{2-2H}\mathbf{E}\big(\hat{\theta}_T - \theta\big)^2 = T^{2-2H}\frac{1}{\langle M\rangle_T} = \frac{1}{\int_0^1 g_\varepsilon(u)du} \xrightarrow{T\to\infty} \frac{1}{\int_0^1 g(u)du},$$

where g is the solution of the limit equation $\int_0^1 g(u)\kappa(u,v)dv = 1$ and we used Lemma 3.6. The constant (3.25) is obtained by plugging the explicit expression for g, given by (3.51).

7.2 Proof of (3.26)

Let $\varepsilon := T^{2H-1}$ and define $g_\varepsilon(u) := g(uT, T)$, where now g is the unique solution of equation (3.36). The function g_ε solves:

$$\varepsilon g_\varepsilon(u) + \beta_H \int_0^1 g_\varepsilon(v)\bar{\kappa}(u,v)dv = c_H u^{1/2-H}(1-u)^{1/2-H} \quad u \in [0,1].$$

The same arguments as in Lemma 3.6 (see Remark 11) with $\varphi(u) = c_H u^{1/2-H}(1-u)^{1/2-H}$ and $\psi \equiv 1$ imply

$$T\mathbf{E}\big(\hat{\theta}_T - \theta\big)^2 = T\frac{1}{\langle M\rangle_T} = \frac{1}{\int_0^1 g_\varepsilon(u)du} \xrightarrow{T\to\infty} \frac{1}{\int_0^1 g(u)du},$$

where g is the solution of the limit equation $\beta_H \int_0^1 g(u)\bar{\kappa}(u,v)dv = c_H u^{1/2-H}(1-u)^{1/2-H}$. A direct calculation shows that $g(u) \equiv 1$, which confirms the constant in the right hand side of (3.26).

Strong consistency for $H \in (0,1)$ follows from the law of large numbers for martingales.

Bibliography

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

[2] Anděl J. (1986) Long memory time series models, *Kybernetika*, 22(2), 105-123.

[3] Anderson T. W. (1959) On asymptotic distribution of estimates of parameters of stochastic difference equations, *Annals of Mathematical Statistics*, 30, 676-687.

[4] Aoki H. and Staley R. (1970) On input signal synthesis in parameter identification, *Automatica*, 6, 431-440.

[5] Azmooden E. and Valkeila E. (2013) Spectral characterization of the quadratic variation of mixed fractional brownian motion. *Stat. Inference Stoch. Process*, 16, 97-112.

[6] Baudoin F. and Nualart D. (2003) Equivalence of Volterra processes, *Stochastic Process. Appl*, 107(3), 2003.

[7] Bender C. , Sottinen T. and Valkeila E. (2011) Fractional processes as models in stochastic finance, *Adavanced mathematical methods for finance*,75-103, Springer, Heidelberg.

[8] Bercu B. , Coutin L. and Savy N. (2011) Sharp large deviations for the fractional Ornstein-Uhlenbeck process, *SIAM Theory of Probability and its Applications*, 55, 575-610.

[9] Biagini F. , Hu Y. , Øksendal B. and Zhang T. (2008) Stochastic calculus for fractional Brownian motion and application., *Probability and its Applications(New York)*, Springer-Verlag London.

[10] Brouste A. and Kleptsyna M. (2010) Asymptotic properties of MLE for partially observed fractional diffusion system, *Statistical Inference for Stochastic Processes*, 13(1), 1-13.

[11] Brouste A., Kleptsyna M. and Popier. A. (2011) Fractional diffusion with partial observations, *Communications in Statistics - Theory and Methods*, 19-20(40), 3479-3491

[12] Brouste A., Kleptsyna M. and Popier A. (2012) Design for estimation of drift parameter in fractional diffusion system, *Statistical Inference for Stochastic Process*, 15(2), 133-149.

[13] Brouste A. and Kleptsyna M. (2012) Kalman's type filter under stationary noises, *Systems and Control Letters*, 1, 517-532.

[14] Brouste A. and Cai C. (2012) Controlled drift estimation in fractional diffusion linear systems, *Stochastic and Dynamics*, 13(3).

[15] Chan N. H. and Wei C. Z. (1987) Asymptotic inference for nearly non statinary AR(1) process *Annals of Statistics*, 15(3), 1050-1063.

[16] Cheridito P. (2001) Mixed fractional Brownian motion, *Bernoulli*, 7(6), 913-934.

[17] Cheridito P. (2003) Arbitrage in fractional Brownian motion models, *Finance Stoch.*, 7(4), 533-553.

[18] Cheridito P. , Kawaguchi H. and Maejima M. (2003) Fractional Ornstein-Uhlenbeck processes, *Electronic Journal of Probability*, 8(3), 1–14.

[19] Cialenco I. and Lototsky S. and Pospisil J. (2009) Asymptotic properties of the Maximum Likelihood Estimator for stochastic parabolic equations with additive fractional Brownian motion, *Stochastics and Dynamics*, 9(2), 169-185.

[20] Duflo M. , Senoussi R. , and Touati A. (1991) *Propriété asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autorégressif vectoriel*, *Annals de l'institut Henri Poincaré*, 27(1), 1-25.

[21] Durbin J. (1960) The fitting of time series models, *Rev. Inst. Int. Stat*, 28, 233-243.

[22] Edwards R. E. (1965) Function analyse Theorem and Application. *Holt, Rinehart and Winston, New York.*

[23] Fox R. and Taqqu M. (1986) Large-Sample properties of parameter estimates for Strongly dependent stationary Gaussian time series, *The Annals of Statistics*, 14(2), 517-532.

[24] Gevers M. (2005) From the early achievement to the revival of experiment design, *European Journal of Control*, 11, 1-18.

[25] Goodwin G. and Payne R. (1977) Dynamic System Identification: Experiment Design and Data Analysis, *Academic Press*, New York.

[26] Goodwin G. , Rojas C. , Welsh J. and Feuer A. (2007) Robust optimal experiment design for system identification, *Automatica*, 43, 993-1008.

[27] Granger C.W. and Joyerx R. (1980) An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1(1), 15-29.

[28] Greene M. T. and Fielitz B. D. (1977) Long-term dependence in common stock returns, *Journal of Financial Economics*, 4, 339-349.

[29] Hida T. and Hitsuda M. (1976) Gaussian processes, *American Mathematical Society, Provindence, RI, 1993* volume 120 of *Translations of Mathematical Monographs.*

[30] Hosking J. R. M. (1981) Fractional differencing, *Biometrika*, 68(1), 165-176.

[31] Ibragimov I. and Khasminskii R. (1981) Statistical Estimation. Asymptotic Theory, *Springer*.

[32] Istas J. and Lang G. (1997) Quadratic variations and estimation of the local Hölder index of a Gaussian process, *Ann. Inst. Henri Poincaré*, 33(4), 407-436.

[33] Jost C. (1999) Transformation formulas for fractional brownian motion, *Stochastic Process. Appl*, 116(10), 1341-1357.

[34] Kiefer J. (1974) On the efficient design of statistical investigation, *The Annals of Statistics*, 2, 849-879.

[35] Kleptsyna M. , Le Breton A. and Roubaud M. -C. (2000) Parameter estimation and optimal filtering for fractional type stochastic systems, *Statistic inference for Stochastic Processes*, 3(1-2), 173-182.

[36] Kleptsyna M. and Le Breton A. (2001) Optimal linear filtering of general multidimensinal Gaussian process - application to Laplace transforms of quadratic functionals, *Journal of Applied Mathematics and Stochastic Analysis*, 14(3), 215-226.

[37] Kleptsyna M. and Le Breton A. (2002) Statistical Analysis of the Fractional Ornstein-Uhlenbeck type Process, *Statistical Inference for Stochastic Processes*, 5, 229–241.

[38] Kleptsyna M. and Le Breton A. (2002) Extension of the Kalman-Bucy filter to elementary linear systems with fractional Brownian noises. *Statistical Inference for Stochastic Process*, 5, 249-271.

[39] Kleptsyna M. , Le Breton A. and Viot M. (2002) New formulas concerning Laplace transforms of quadratic form for general Gaussian sequences, *Journal of Applied Mathematics and Stochastic Analysis*, 1.

[40] Kunita M. (1984) Equations and stochatic flow of diffeomorphisms, *Lectures Notes on Mathematics*, 1097, 143-303.

[41] Kutoyants Y. A. (2002) Statistical Inference For Ergodic Diffusion Process , *Springer*.

[42] Lai T. L. and Wei C. Z. (1983) Asymptotic properties of general autoregressive models and strong consistency of least square estimate and their parameters, *Journal of Multivariate Analysis*, 13, 1-23.

[43] Lawrance A. J. and Kottegoda N. T. (1977) Stochastic modeling of river flow time series, *Journal of the Royal Statistical Society*, 140, 1-27.

[44] Le Broton A. and Pham D. T. (1989) On the bias of the least squares estimator for the first order autoregrssive process, *Ann. Inst. Statist. Math* 41(3), 1-27.

[45] Le Breton A. (1998) Filtering and parameter estimation in a simple linear system driven by a fracional Brownian motion , *Statist. Probab. Lett*, 38(3), 263-274.

[46] Levadi V. (1966) Design of input signals for parameter estimation, *IEEE Trans. Automat. Control* , 11, 205-211.

[47] Levin M. (1960) Estimation of impulse response in the presence of noise, *IRE Trans. Circuit Theory*, 7, 50-56.

[48] Liptser R. and Shiryaev A. (1989) Theory of Martingales, *Kluwer Academic.*

[49] Liptser R. and Shiryaev A. (2001) Statistics of Random Processes, *Springer.*

[50] Mandelbrot B. and van Ness J. W. (1968) Fractional brownian motion, fractional noises and application, *SIAM Review*, 10(4), 422-437.

[51] Mann H. and Wald A. On the statistical treatment of linear stochastic difference equation, *Econometrica*, 11(3 & 4), 173-220.

[52] Maxim V. , SendurL. , Fadili J. , Suckling J. , Gould R. , Howard R. and Bullmore E. (2005) Fractional gaussian noise, function MRI and Alzheimer's disease, *NeuroImage*, 25, 141-158.

[53] Mehra R. (1974) Optimal input for linear system identification, *IEEE Trans. Automat. Control*, 19, 192-200.

[54] Mehra R. (1974) Optimal inputs signal for parameter estimation in dynamic systems-survey and new results, *IEEE Trans. Automat. Control*, 19, 753-768.

[55] Mishura Y. S. (2008) Stochastic calculus for fractional Brownian motion and related processes, *Lecture Notes in Mathematics*, Volume 1929, Springer-verlag, Berlin.

[56] Norros I. , Valkeila E. and Virtamo J. (1999) An elementary approach to a Girsanov formula and other analytical results on fractional Brownian motion, *Bernoulli*, 5, 571-587.

[57] Ovseevich A., Khasminskii R. and Chow P. (2000) Adaptative design for estimation of unknown parameters in linear systems, *Problems of Information Transmission*, 36(2), 125-153.

[58] Pal'tsev B. V. (2003) Asymptotics of the spectral of integral convolution operators on a finite interval with homogeneous polar kernel, *Izvestiya: mathematics*, 67(4), 695.

[59] Pipiras V. and Taqqu M. S. (2001) Are classed of deterministic integrands for fractional Brownian motion on an interval complete, *Bernoulli*, 7(6), 873-897.

[60] Polyanin A. D. and Manwzhirov A. V. (2008) Handbook of integral equation, *Chapman & Hall/CRC, Boca Raton, FL*, second edition

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

[62] Rao M. M. (1961) Consistency and limit distributions of estimators of parameters in explosive stochastic difference equations, *Annals of Mathematical Statistics*, 32(1), 195-218.

[63] Rao M. M. (1978) Asymptotic distribution of an estimator of the boundary parameter of an unstable process, *Annals of Statistics*, 6(1), 185-190.

[64] Riesz F. and Sz.-Nagy B. (1990) Functional analysis, *Dover books on Advanced Mathematics*. Dover Publication Inc, New York. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.

[65] Rozovskii B. (1990) Stochastic Evolution System, *Kluwer*.

[66] Rubin H. (1950) Consistency of maximum-likelihood estimates in the explosive case, *Statistical Inference in Dynamic Economic Models*, 356-364.

[67] Samko S. G., Kibas A. A. and Marichev O. I. (1993) Fractional integrals and derivatives, Theorem and applications, *Gordon and Breach Science Publishers*, Translated from the 1987 Russian original.

[68] Sethuraman S. and Basawa I. V. (1997) The asymptotic distribution of the maximum likelihood estimator for a vector time series model with long memory dependence, *Statistics and Probability Letters*, 31, 285-293.

[69] Shubin C. (2006) Singularly perturbed integral equation, *J. Math. Anal. Appl.*, 313(1), 234-250.

[70] Sinai Y. G. (1976) Self-similar probability distribution, *Theory of probability and Its application*, 21(1).

[71] Sweeting T. J. (1980) Uniform asymptotic normality of the maximum likelihoos estimator, *The Annals of Statistics*, 8, 1375-1381.

[72] Vainikko G. and Pedas A. (1980) The properties of solutions of weakly singular integral equations, *J. Austral. Math. Soc. Ser*, B 22(4), 419-430.

[73] Vainikko G. (1983) Multidimensional weakly singular integral equation. *Lecture notes in Mathematics*, Vol 1549, Springer-Verlag, Berlin.

[74] Van Zanten H. (2007) When is a linear combination of independent fBm's equivalent to a single fBm, *Stochastic Process. Appl*, 117(1), 57-70.

[75] Van Zanten H. (2008) A remark on the equivalence of Gaussian processes, *Electron. Commun. Probab.*, 13, 54-59.

[76] Wald A. (1943) On the efficient design of statistical investigation, *Annals Math. Stat*, 14, 134-140.

[77] White J. (1958) The limiting distribution of the serial correlation coefficient in the explosive case, *Annals of Mathematical Statistics*, 29, 1188-1197.

[78] White J. (1959) The limiting distribution of the serial correlation coefficient in the explosive case ii, *Annals of Mathematical Statistics*, 30, 831-834.

[79] Whittele P. (1973) Some general points in the theory of optimal experimental design, *J. R. Stat. Stoc*, 1, 123-140.

[80] Wood A. and Chan G. (1994) Simulation of stationary gaussian processes, *Journal of computational and graphical statistics*, 3(4), 409-432.

[81] Yajima Y. (1985) On estimation of long-memory time series models, *Australian Journal of Statistics*, 27(3), 303-320.

[82] Brouste A. , Cai C. and Kleptsyna M. (2013) Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise submitted.

[83] Cai C. , Chigansky P. and Kleptsyna M. (2013) Mixed fractional brownian motion: The filtering perspective Submitted.

Thèse de Doctorat

Chunhao CAI

Analyse Statistique de quelques Modèles de Processus de Type Fractionnaire

Résumé

Cette thèse porte sur l'analyse statistique de quelques modèles de processus stochastiques gouvernés par des bruits de type fractionnaire, en temps discret ou continu.

Dans le Chapitre 1, nous étudions le problème d'estimation par maximum de vraisemblance (EMV) des paramètres d'un processus autorégressif d'ordre p (AR(p)) dirigé par un bruit gaussien stationnaire, qui peut être à longue mémoire comme le bruit gaussien fractionnaire. Nous donnons une formule explicite pour l'EMV et nous analysons ses propriétés asymptotiques. En fait, dans notre modèle la fonction de covariance du bruit est supposée connue, mais le comportement asymptotique de l'estimateur (vitesse de convergence, information de Fisher) n'en dépend pas.

Le Chapitre 2 est consacré à la détermination de l'entrée optimale (d'un point de vue asymptotique) pour l'estimation du paramètre de dérive dans un processus d'Ornstein-Uhlenbeck fractionnaire partiellement observé mais contrôlé. Nous exposons un principe de séparation qui nous permet d'atteindre cet objectif.

Dans le Chapitre 3, nous présentons une nouvelle approche pour étudier les propriétés du mouvement brownien fractionnaire mélangé et de modèles connexes, basée sur la théorie du filtrage des processus gaussiens. Les résultats mettent en lumière la structure de semimartingale et mènent à un certain nombre de propriétés d'absolue continuité utiles. Nous établissons l'équivalence des mesures induites par le mouvement brownien fractionnaire mélangé avec une dérive stochastique, et en déduisons l'expression correspondante de la dérivée de Radon-Nikodym. Pour un indice de Hurst H> 3/4, nous obtenons une représentation du mouvement brownien fractionnaire mélangé comme processus de type diffusion dans sa filtration naturelle et en déduisons une formule de la dérivée de Radon-Nikodym par rapport à la mesure de Wiener. Pour H <1/4 , nous montrons l'équivalence da la mesure avec celle la composante fractionnaire et obtenons une formule pour la densité correspondante.

Mots clés : processus fractionnaire, mouvement brownien fractionnaire, mouvement brownien fractionnaire mélangé, estimateur de maximum de vraisemblance.

Abstract

This thesis focuses on the statistical analysis of some models of stochastic processes generated by fractional noise in discrete or continuous time.

In Chapter 1, we study the problem of parameter estimation by maximum likelihood (MLE) for an autoregressive process of order p (AR (p)) generated by a stationary Gaussian noise, which can have long memory as the fractional Gaussian noise. We exhibit an explicit formula for the MLE and we analyze its asymptotic properties. Actually in our model the covariance function of the noise is assumed to be known but the asymptotic behavior of the estimator (rate of convergence, Fisher information) does not depend on it.

Chapter 2 is devoted to the determination of the asymptotical optimal input for the estimation of the drift parameter in a partially observed but controlled fractional Ornstein-Uhlenbeck process. We expose a separation principle that allows us to reach this goal. Large sample asymptotical properties of the MLE are deduced using the Ibragimov-Khasminskii program and Laplace transform computations for quadratic functionals of the process.

In Chapter 3, we present a new approach to study the properties of mixed fractional Brownian motion (fBm) and related models, based on the filtering theory of Gaussian processes. The results shed light on the semimartingale structure and properties lead to a number of useful absolute continuity relations. We establish equivalence of the measures, induced by the mixed fBm with stochastic drifts, and derive the corresponding expression for the Radon-Nikodym derivative. For the Hurst index H>3/4 we obtain a representation of the mixed fBm as a diffusion type process in its own filtration and derive a formula for the Radon-Nikodym derivative with respect to the Wiener measure. For H<1/4, we prove equivalence to the fractional component and obtain a formula for the corresponding derivative. An area of potential applications is statistical analysis of models, driven by mixed fractional noises.

Key Words: fractional Brownian motion, mixed fractional Brownian motion, maximum likelihood estimator