STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–18.
Published online in International Academic Press (www.IAPress.org)

Maximum likelihood estimation of reflected fractional Ornstein-Uhlenbeck processes with mixed random effects

Hicham Chaouch, Hamid El Maroufy*, Jilali Stoutou

Laboratory of Mathematics and Applications, Faculty of Sciences and Techniques, Sultan Moulay Slimane University, Béni-Mellal, Morocco

Abstract We consider a stochastic reflected fractional Ornstein-Uhlenbeck process driven by a mixture of Gaussian random effects. The key parameters of the model are unknown and must be estimated. To do that, we formulate the likelihood function and derive its estimators. Hereafter, we establish the consistency of the maximum likelihood estimator, and we compute them by implementing the EM algorithm. When the number of mixture components is unknown, the Bayesian Information Criterion is used for model selection. To illustrate our theoretical results, we give a numerical simulation.

Keywords Fractional Brownian Motion, Reflected Ornstein-Uhlenbeck process, Stochastic Differential Equations, Girsanov Formula, Random Effects, Maximum Likelihood Estimation, EM algorithm.

AMS 2010 subject classifications 60G22, 62M09

DOI: 10.19139/soic-2310-5070-4363

1. Introduction

The reflected fractional Ornstein-Uhlenbeck (**rfOU**) process is widely used in fields such as mathematical finance, physics, biology, and queuing theory (see [37, 1, 4, 5]). The **rfOU** behaves similarly to a standard fractional Ornstein-Uhlenbeck (**fOU**) process. However, when it reaches zero, the process reflects with minimal force, ensuring non-negativity throughout its evolution. This model is particularly suitable for modeling queuing systems with reneging or balking customers, especially when inter-arrival or service times exhibit long-range dependence (see [39, 23, 41]). To describe the **rfOU** model, we consider a complete filtered probability space $\Lambda := (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\geq 0}$ is a filtration satisfying the usual conditions, on which we define the positive stochastic process **X**=$(X(t), t \geq 0)$, $i.e$, for $t \geq 0$, $X(t) \geq 0$ and satisfies the following fractional stochastic differential equation (**fSDE**) :

$$dX(t) = -\varphi X(t)dt + \sigma dB^H(t) + dL(t),\ X(0) \in \mathbb{R}^+_*. \tag{1}$$

Where $\mathbf{B}^H = \big(B^H(t), t \geq 0\big)$ the fractional Brownian motion (**fBm**) on Λ with Hurst index $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, supposed to be known, and $(\varphi, \sigma) \in \mathbb{R} \times \mathbb{R}^*_+$. The reflection constraint process **L**=$(L(t), t \geq 0)$, is non-decreasing, non-negative, and minimal with $L(0) = 0$, that increases only when **X** reaches the boundary 0. So we have:

$$\int_0^{+\infty} \mathbb{1}_{(X(t)>0)}(t)dL(t) = 0, \tag{2}$$

*Correspondence to: Email: Hamid El Maroufy (Email: h.elmaroufy@usms.ma). Address: Department of Mathematics, Faculty of Sciences and Techniques, Mghila C.P: 523, Béni-Mellal, 23000, Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 202x International Academic Press

where $\mathbb{1}_A$ is the indicator function of any set A. That is, $L(t)$ is the cumulative local time at the boundary zero up to time $t \geq 0$ of $\mathbf{X}$. The **rfOU** given by (1) is a Skorohod problem with linear drift for $\mathbf{B}^H$ with the solution given by the pair $(\mathbf{X}, \mathbf{L})$.

Let consider the classical Skorohod problem for $\mathbf{Z}$=$(Z(t),, t \geq 0) \in C(\mathbb{R}^+, \mathbb{R})$, where $C(\mathbb{R}^+, \mathbb{R})$ designates the space of continuous functions on $[0, \infty)$. That is, $X(t) = Z(t) + L(t)$, where for all $t \geq 0,\ X(t) \geq 0$ and $\mathbf{L}$ is non-decreasing with $L(0) = 0$, such that (2) is satisfy. By the classicla result of Skorohod (see, [38]), the unique solution $(\mathbf{X}, \mathbf{L}) = (\mathbf{\Psi}(\mathbf{Z}), \mathcal{A}(\mathbf{Z})) \in \mathbf{C}(\mathbb{R}^+, \mathbb{R}) \times \mathbf{C}(\mathbb{R}^+, \mathbb{R})$ is given by

$$X(t) = Z(t) + L(t) \quad \text{and} \quad L(t) = \sup_{0 \leq s \leq t} (-Z(s)) \vee 0. \tag{3}$$

To ensure the existence, uniqueness, and Lipschitz continuity of the map $(\Psi, \mathcal{A})$, the reader is referred to [9, Theorem 7.2]. The solution of the generalized drift Skorohod problem with a Lipschitz continuous drift f, $i.e$, $dX(t) = f(X(t))dt + dZ(t) + dL(t)$ for $\mathbf{Z} \in C(\mathbb{R}^+, \mathbb{R})$, as proved by [35, pp: 16-17], is given by $(X, L) = (\Psi(\mathfrak{M}(Z)), \mathcal{A}(\mathfrak{M}(Z)))$, where, $\mathfrak{M} : C(\mathbb{R}^+, \mathbb{R}) \to C(\mathbb{R}^+, \mathbb{R})$ is a mapping which gives $\mathfrak{M}(Z) = Y$, with Y solves the following equation:

$$dY(t) = f(\Psi(Y)(t))dt + dZ(t),$$

and its unique solution can be obtained by a standard Picard iteration, since f and Ψ are Lipschitz (see, e.g., [9, 15, 41, 36]. Accordingly, the Skorohod problem (1), has a unique pathwise solution $(X, L) \in C(\mathbb{R}^+, \mathbb{R}) \times C(\mathbb{R}^+, \mathbb{R})$. Additionally, by (3) it can be shown that the process $\mathbf{L}$ has the following explicit expression:

$$L(t) = \max\left(0, \sup_{s \in [0,t]} \left(-x + \varphi \int_0^s X(u)du - \sigma B^H(s)\right)\right), \quad \text{for all} \quad t \geq 0.$$

The present paper considers the **rfOU** model given in (1), with the parameter φ replaced by a random variable ϕ, called a random effect. This random effect is characterized by an unknown distribution, or by a known distribution with unknown parameters. Within this last context, the main goal of this work is situated. We assume that the distribution of ϕ is a Gaussian mixture depending on the parameter θ, which we plan to estimate; the details about the components of θ will be specified later. To do this, we consider N independent continuously observed processes over a time interval $[0, T],\ T > 0$; $(X_i(t), t \geq 0)\ ,\ i = 1, \cdots, N$, described by the reflected **SDE**'s given by (1) with random effect, henceforth abbreviated as **rSDEr**; for $i = 1, 2, ..., N$ and $t \geq 0$,

$$dX_i(t) = -\phi_i X_i(t)dt + \sigma dB_i^H(t) + dL_i(t), \quad X_i(t) \geq 0, \quad \text{and} \quad X_i(0) = x_i. \tag{4}$$

where $\left(\mathbf{B}_1^H, \cdots, \mathbf{B}_N^H\right)$ is a vector of independent Fractional Brownian Motion supposed to be independent of $(\phi_1, \cdots, \phi_N)$, a vector of independent and identically distributed (iid) random variables with common density $g(.\theta)d\nu(.)$ on $\mathbb{R}$, depending on θ, with θ_0 its true value. Here we suppose the diffusion coefficient σ is a known real constant. The aforementioned processes and random variables are all defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, where the natural filtration $\mathcal{F}_t = \sigma(\phi_i, B_i^H(s), s \leq t, i = 1, \cdots, N)$, and ϕ_i's are $\mathcal{F}_0$-measurable.

This estimation problem is naturally related to numerous important recent works focusing on statistical estimation of random effect distribution parameters in **SDE**'s as in (4), but without the reflected term and considering only standard Brownian motion as the driving process. Here we will cite only the most important works, including those in [12, 11, 33, 13]. In [12], the authors investigate a continuous observation of the considered process to derive a maximum likelihood estimator (**MLE**) of the parameters by exploiting the Girsanov formula, and study their asymptotic properties, where the drift depends linearly on a Gaussian random effect. In [11], the authors consider an SDE with linear random effects in drift and diffusion terms; they develop an approximate likelihood by exploiting discrete-time approximations and derive the estimators of the Gaussian random effect distribution parameters. In the third work, the authors in [33] establish approximate likelihood procedures by incorporating random effects with fixed parameters in the model. By using the same model as in [12], with a special extension provided by a Gaussian mixture as a distribution for the random effect, the authors [13] prove the consistency of the exact **MLE** of the distribution parameters, and compute them using the EM algorithm.

However, we find only a few studies that have addressed this problem of estimation when the **SDE** is driven by an **fBm**; we specifically mean the more recent works in [10, 31, 18, 19]. In [10], the authors construct the likelihood function via kernel transformation so that the Hurst index is known; they derive estimators for the Gaussian distribution of the random effects in the drift term. In [31], the authors treat the asymptotic normality and moment convergence of the MLE under small diffusion. In [18], after estimating the Hurst index, the authors establish the likelihood using the Girsanov theorem, estimate the mean and variance of the Gaussian random-effect distributions, and prove their consistency. In [19], the authors establish a rigorous simple method of moments for estimating the moments of generalized random-effect distributions, and subsequently they derive the parameter estimates for specific distributions.

Alternative studies propose a nonparametric estimation approach, for which the density of ϕ is estimated using kernel, histogram, Hermite or Bernstein polynomial approximation, or deconvolution methods, but only for non-reflected SDEs with fBm (see [16, 20, 30, 17, 8], and the references therein). Other works, as in [7], combine the nonparametric approach by Lagrange polynomial approximation with the parametric variation method.

Unlike the aforementioned works, particularly those in [12] and [13], which focused on the parametric approach and explored the standard Girsanov theorem based on Itô calculus to construct the likelihood, a task made more challenging by the presence of fBm as the driving noise and Skorokhod reflection, we address this difficulty by investigating fundamental martingale theory to establish the likelihood for (4). We do this with inspiration from the work in [10], where the authors do it but without reflection and without mixed random effects. After constructing the exact likelihood, we derive estimators of the random effects parameters, and we show their consistency, and we support the theoretical result with numerical implementations. Therefore, we try to estimate, numerically, the number of components of the mixed distribution.

We organize the rest of the paper as follows: In Section 2, we formulate the likelihood function for **rfOU** model as in (4) when the random effect is distributed as a mixture of Gaussian distributions, and hereafter we derive the likelihood estimator (**MLE**), then we establish the theoretical consistency of the estimators when the components number is Known, in Section 3, we establish the joint estimation algorithm and the simulation methodology; specifically, we describe the iterative **EM** algorithm and **BIC** procedure for two cases when the components number is Known and unknown. in section 4, we illustrate our theoretical results by numerical simulation. We conclude by summarizing and discussing the obtained results.

2. Parameter Estimation of rfOU with Random Effect

In this section, we construct the likelihood function for the **rfOU** process $\mathbf{X} = (X_i(t),\ \ t \geq 0, i = 1, 2, \dots, N)$ ruled by (4). To achieve this, we first employ the classical Girsanov theorem, developed for the **rfOU** process without random effect [25], and then derive the maximum likelihood estimators (MLE) for the random effect parameters.

2.1. Girsanov theorem for rfOU

Let $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$ be a fixed constant. A fractional Brownian motion $\mathbf{B}^H$ with Hurst index H is a centered Gaussian process with covariance function

$$R_H(t,s) := \mathbb{E}\left(B^H(t)B^H(s)\right) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

Even if $\mathbf{B}^H$ is not a semi-martingale, then we can transform it into a martingale, and we have the natural filtration of fBm $\mathbf{B}^H$ and a fundamental Gaussian martingale (see [32]) defined by $M^H(t) = \int_0^t k_H(t,s)dB^H(s)$, where $k_H(t,s) = \kappa_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}$, $0 < s < t$, and $\kappa_H = 2H\Gamma(\frac{3}{2}-H)\Gamma(H+\frac{1}{2})$. For more details on **fBm**, the readers can consult the book by Mishura [29].

For convenience, we denote the quadratic variation $\langle \mathbf{M}^H \rangle(t)$ of the martingale $\mathbf{M}^H$ as $\omega^H(t)$ and we have .

$$\omega^H(t) = \lambda_H^{-1} t^{2-2H}, \text{ where } \lambda_H = 2H\Gamma(3-2H)\Gamma(H+\frac{1}{2})/\Gamma(\frac{3}{2}-H).$$

In the case, when $\phi_i = \varphi \in \mathbb{R}$, then the Skorohod problem

$$dX_i(t) = -\varphi X_i(s)dt + \sigma dB_i^H(t) + dL_i(t), \qquad X_i(0) = x^i, \tag{1}$$

has a unique solution (X_i, L_i) and the process $\mathbf{X}_i$ is adapted to the filtration $(\mathcal{F}_t, t \geq 0)$ (see [40, 9]). We can now use the fundamental martingale M^H to derive the Girsanov's theorem for our **rfOU** process $(X_i, \ \ i = 1, \ldots, N)$. For that, we consider the following processes:

$$\tilde{X}_i(t) = \int_0^t k_H(t,s)dX_i(s), \ \tilde{L}_i(t) = \int_0^t k_H(t,s)dL_i(s), \text{ and } \quad \chi_i(t) = \frac{d}{d\omega_i^H(t)} \int_0^t k_H(t,s)X_i(s)ds. \tag{2}$$

Thus

$$d\tilde{X}_i(t) = -\varphi \chi_i(t) d\omega_i^H(t) + \sigma dM_i^H(t) + d\tilde{L}_i(t). \tag{3}$$

Theorem 2.1 (Girsanov theorem for rfOU [25])
For $H \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$, let consider the sequence

$$\eta_i^{T,\varphi} = \exp\left(\frac{\varphi}{\sigma}\int_0^T \chi_i(s)dM_i^H(s) - \frac{\varphi^2}{2\sigma^2}\int_0^T \chi_i(s)^2 d\omega_i^H(s)\right), \ \ i = 1, \ldots, N.$$

Then, for $i = 1, 2, \ldots, N$, under the new probability, $d\tilde{\mathbb{P}}_i = \eta_i^{T,\varphi} d\mathbb{P}$, the process $\left(\sigma^{-1}X_i(s) : 0 \leq s \leq T\right)$ is a reflected fractional Brownian motion (rfBm).

We denote by $\mathbb{P}_i^{\varphi}$ and $\mathbb{P}_i^0$ the probability measures on the path space $(C([0,T]), \mathcal{B}_T)$ induced by X_i and X_i^0 respectively, where $(X_i(t), t \in [0,T])$ satisfies (1) and X_i^0 is a reflected fractional Brownian motions ($i.e.$ $\varphi = 0$). Following [27, Theorem 7.1] and using Theorem 2.1, we can show that, for $i = 1, \cdots, N$, $\mathbb{P}_i^{\varphi}$ and $\mathbb{P}_i^0$ are equivalent, and using the arguments in (3), we have:

$$\begin{aligned}
\frac{d\mathbb{P}_i^{\varphi}}{d\mathbb{P}_i^0} = \left(\eta_i^{T,\varphi}\right)^{-1} &= \exp\left(-\frac{\varphi}{\sigma}\int_0^T \chi_i(s)dM_i^H(s) + \frac{\varphi^2}{2\sigma^2}\int_0^T \chi_i^2(s)d\omega_i^H(s)\right), \\
&= \exp\left(-\frac{\varphi}{\sigma^2}\int_0^T \chi_s d\tilde{X}_s - \frac{\varphi^2}{2\sigma^2}\int_0^T \chi_s^2 d\omega_i^H(s) + \frac{\varphi}{\sigma^2}\int_0^T \chi(s)d\tilde{L}_s\right), \\
&= \exp\left(\frac{\varphi}{\sigma^2}U_i - \frac{\varphi^2}{2\sigma^2}V_i\right) := L^{x^i,T_i}(\mathbf{X_i}, \varphi),
\end{aligned} \tag{4}$$

where

$$U_i = -\int_0^T \chi_i(s)d\tilde{X}_i(s) + \int_0^T \chi_i(s)d\tilde{L}_i(s) \quad \textit{and} \quad V_i = \int_0^T \chi_i(s)^2 d\langle M_i^H\rangle(s). \tag{5}$$

We now extend this to construct the likelihood function for **rfOU** processes under random effects given by the equation (4). On the measurable space $(C_T, \mathcal{C}_T)$, we interpret $\mathbb{P}_i^{\varphi}$ as the conditional distribution of process $(X_i(t), t \in [0, T_i])$ given the random effect $\phi_i = \varphi$. This allows us to express the joint distribution of $(\phi_i, X_i(\cdot))$ on the product space $\mathbb{R} \times C_T$ by $\mathbb{P}_i^{\theta} = g(\varphi, \theta)d\nu(\varphi) \otimes \mathbb{P}_i^{\varphi}$. We define the marginal distribution of $(X_i(t), t \in [0, T_i])$ as $\mathbb{Q}_i^{\theta}$. Finally, to construct the likelihood function, we introduce the following assumption : For every φ and $\varphi^{'}$.

$$\text{A1} : \mathbb{P}^{\varphi}\left(\int_0^T \left(\frac{d}{d\omega_i^H(t)}\int_0^t k_H(t,s)X_i(s)\frac{\varphi^{'}}{\sigma}ds\right)^2 d\omega_i^H(t) < +\infty\right) = 1, \quad i = 1, \cdots, N.$$

It should be noted that the previous assumption, for rfOU, is always verified due to the following simple analysis: since $\chi_i(t) := \dfrac{d}{d\omega_i^H(t)} \displaystyle\int_0^t k_H(t,s)\, X_i(s)\, ds$, then $\int_0^T \left(\frac{\varphi'}{\sigma}\chi_i(t)\right)^2 d\omega_i^H(t) = \frac{(\varphi')^2}{\sigma^2} \int_0^T \chi_i(t)^2\, d\omega_i^H(t) = \frac{(\varphi')^2}{\sigma^2}\, V_i$, where $V_i := \int_0^T \chi_i(t)^2\, d\omega_i^H(t)$. Since $(\varphi')^2/\sigma^2$ is a real strictly positive, then $\frac{(\varphi')^2}{\sigma^2} V_i < +\infty,\ a.s \iff V_i < +\infty,\ a.s$. Thus, assumption A1 is equivalent to the following statement: $\mathbb{P}^{\varphi}\big(V_i < +\infty\big) = 1$. So, by [25, Lemma 4.4], $\mathbb{P}^{\varphi}\big(V_i < +\infty\big) = 1$, is therefore always verified.

Since $\mathbb{P}^{\theta} = \otimes_{i=1}^N \mathbb{P}_i^{\theta}$ represents the distribution of $((\varphi_i, X_i(\cdot)), i = 1, \cdots, N)$ on the space $\prod_{i=1}^N \mathbb{R} \times C_{T_i}$. On $\mathcal{C} = \prod_{i=1}^N \mathcal{C}_{T_i}$, the distribution of process $(X_i(t), t \in [0, T_i], i = 1, \cdots, N)$ is defined, through independence of the individuals, as $\mathbb{Q}^{\theta} = \otimes_{i=1}^N \mathbb{Q}_i^{\theta}$. Finally, the exact likelihood is the density of $\mathbb{Q}^{\theta}$ with respect to $\mathbb{Q} = \otimes_{i=1}^N \mathbb{Q}_i$, where $\mathbb{Q}_i$ is the distribution of $(X_i(t), t \in [0, T_i])$, as defined in (4) without drift term. Consequently, we can postulate the following theorem:

Theorem 2.2

Let consider the **rfOU** processes $(X_i(t), t \in [0,T])\ , i = 1, .., N)$, the solution of (4), and under the assumption A_1, the following results hold:

(i) The probability measure $\mathbb{Q}_i^{\theta}$ is characterized by the following Radon-Nikodym derivative w.r.t to $\mathbb{Q}_i$:

$$\frac{d\mathbb{Q}_i^{\theta}}{d\mathbb{Q}_i}(X_i) = \int_{\mathbb{R}} L^{x_i,T}(X_i,\varphi) g(\varphi,\theta) d\nu(\varphi) := \Lambda_i^H(X_i,\theta). \tag{6}$$

(ii) On the product space $\mathcal{C} = \prod_{i=1}^N \mathcal{C}_T$, we have:

$$\frac{d\mathbb{Q}^{\theta}}{d\mathbb{Q}}(X_1,\ldots,X_N) = \prod_{i=1}^N \Lambda_i^H(X_i,\theta).$$

(iii) The exact likelihood of $(X_i(t), t \in [0,T], i = 1, .., N)$ is expressed as follows:

$$L_N^H(X_1,\ldots,X_N,\theta) = \prod_{i=1}^N \Lambda_i^H(X_i,\theta). \tag{7}$$

Proof

Let consider a positive mesurable function F defined on $\mathcal{C}_T$, we have

$$\mathbb{E}_{\mathbb{Q}_i^{\theta}}(F(X_i)) = \mathbb{E}_{\mathbb{P}_i^{\theta}}(F(X_i)) = \mathbb{E}_{\mathbb{P}_i^{\theta}}[\mathbb{E}_{\mathbb{P}_i^{\theta}}(F(X_i)|\phi_i)].$$

Then, by conditioning on the random effect $\phi_i = \varphi$, we hav:

$$\mathbb{E}_{\mathbb{P}_i^{\theta}}((F(X_i)|\phi_i = \varphi)) = \mathbb{E}_{\mathbb{P}_i^{\varphi}}(F(X_i)) = \mathbb{E}_{\mathbb{Q}_i}(F(X_i)L^{x_i,T}(X_i,\varphi)).$$

The following equality can be established using the joint measurability of $L^{x_i,T}(X_i,\varphi)$ with two variables and the Fubini theorem,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_i^{\theta}}(F(X_i)) &= \int_{\mathbb{R}} g(\varphi,\theta) dv(\varphi) \mathbb{E}_{\mathbb{Q}_i}(F(X_i)L^{x_i,T}(X_i,\varphi)) \\ &= \mathbb{E}_{\mathbb{Q}_i}(F(X_i)) \int_{\mathbb{R}} g(\varphi,\theta) dv(\varphi)(L^{x_i,T}(X_i,\varphi)).\end{aligned}$$

Therefore, we can determine the Radon-Nikodym derivative, and we obtain the exact likelihood

$$\frac{d\mathbb{Q}_i^{\theta}}{d\mathbb{Q}_i}(X_i) = \int_{\mathbb{R}} L^{x_i,T}(X_i,\varphi) g(\varphi,\theta) d\nu(\varphi) := \Lambda_i^H(X_i,\theta).$$

□

We define the **MLE** for the parameter θ as $\hat{\theta}_N$, such that

$$\hat{\theta}_N = \arg\max_{\theta\in\Theta} L_N^H(X_1,\dots,X_N;\theta), \tag{8}$$

where $L_N^H(X_1,\dots,X_N;\theta)$ is the likelihood given by (7). The existence and uniqueness of the **MLE** is guaranteed under usual conditions on $\Lambda_i^H(.,\theta)$ for $i=1,...,N$, and the parameter space Θ (see, e.g.,. Prakasa Rao [34], and the references therein).

2.2. *Mixture Gaussian random effects*

In this section, we assume that the function $g(\varphi,\theta)$ is a Gaussian mixture distribution:

$$g(\varphi,\theta) = \sum_{\ell=1}^{M} \pi_\ell \varrho(\varphi;\mu_\ell,\sigma_\ell^2),$$

where $\varrho(\varphi;\mu_\ell,\sigma_\ell^2)$ is the Gaussian $\mathcal{N}(\mu_\ell,\sigma_\ell^2)$ density. For a fixed $\ell\in\{1,2,...,M\}$, let $\tau_\ell=(\mu_\ell,\sigma_\ell)$, we have $\mathbb{P}^{\tau_\ell}$ the joint distribution $n(\varphi;\mu_\ell,\sigma_\ell^2)\otimes\mathbb{Q}^{\tau_\ell}$ and $\mathbb{Q}^{\tau_\ell}$ the corresponding marginal distribution, when $\phi_\ell\sim\mathcal{N}(\mu_\ell,\sigma_\ell^2)$. Additionally, the mixture structure yields: $\mathbb{P}^\theta=\sum_{\ell=1}^M \pi_\ell\mathbb{P}^{\tau_\ell}, \quad \mathbb{Q}^\theta=\sum_{\ell=1}^M \pi_\ell\mathbb{Q}^{\tau_\ell}$. Its follows that the likelihood function $\Lambda_i^H(X_i,\theta)$ introduced in (6) can be rewritten as:

$$\Lambda_i^{H,M}(X_i,\theta) = \sum_{\ell=1}^{M} \pi_\ell \Lambda_i^H(X_i,\tau_\ell), \tag{9}$$

where $\Lambda_i^H(X_i,\tau_\ell)$ represents the likelihood function given in (6) when the random effect satisfies $\phi_i\sim\mathcal{N}(\mu_\ell,\sigma_\ell^2)$, which can be rewritten as:

$$\Lambda_i^H(X_i,\tau_\ell) = \frac{1}{(1+\sigma_\ell^2 V_i)^{\frac{1}{2}}}\exp\left[-\frac{V_i}{2(1+\sigma_\ell^2 V_i)}\left(\mu_\ell-\frac{U_i}{V_i}\right)^2\right]\exp\left(\frac{U_i^2}{2V_i}\right).$$

Under the probability measure $\mathbb{P}_i^{\tau_\ell}$ and given X_i, the random effect ϕ_i is characterized by the normal distribution $\mathcal{N}\left(\frac{\mu_\ell+\sigma_\ell^2 U_i}{1+\sigma_\ell^2 V_i},\frac{\sigma_\ell^2}{1+\sigma_\ell^2 V_i}\right)$ with mean $\frac{\mu_\ell+\sigma_\ell^2 U_i}{1+\sigma_\ell^2 V_i}$ and variance $\frac{\sigma_\ell^2}{1+\sigma_\ell^2 V_i}$, where U_i and V_i are given in (5).

2.3. *Consistency of the* **MLE** *when M is known.*

In the case where the number M of mixture is considered known. Under this assumption, we establish the asymptotic properties of the MLE of θ. To do so, we first specify the parameter space Θ as:

$$\Theta = \left\{ \begin{array}{r} \left((\mu_\ell,\sigma_\ell^2,\pi_\ell)\right)_{\ell=1}^M,:\ 0<\pi_\ell<1,\ \sum_{\ell=1}^M \pi_\ell = 1,\ \tau_\ell=(\mu_\ell,\sigma_\ell^2)\in\mathbb{R}\times(0,+\infty) \\ \ell\neq\ell' \Rightarrow \tau_\ell\neq\tau_{\ell'} \end{array} \right\}.$$

Since $\pi_M = 1-\sum_{\ell=1}^{M-1}\pi_\ell$, then there are $3M-1$ parameters that must be estimated. Let us denote this parameter via the vector $\theta=(\theta_1,\dots,\theta_{3M-1})$; two parameter values are identified as equivalent for the mixture distributions precisely when:

$$\theta\sim\theta_0 \iff \{(\pi_\ell,\tau_\ell),\ell=1,\dots,M\} = \{(\pi_{0,\ell},\tau_{0,\ell}),\ell=1,\dots,M\}.$$

Let now, denote by $I(\theta_0)$ the Fisher information matrix, $i.e$, for all $1\le k,j\le 3M-1$ that

$$I_{jk}(\theta_0) = \left(\mathbb{E}_{\mathbb{Q}_{\theta_0}}\left[\left.\frac{\partial\log\Lambda^{H,M}(X,\theta)}{\partial\theta_k}\right|_{\theta=\theta_0}\left.\frac{\partial\log\Lambda^{H,M}(X,\theta)}{\partial\theta_j}\right|_{\theta=\theta_0}\right]\right)_{jk}.$$

Proposition 2.3

(i) If $\mathbb{Q}^{\theta} = \mathbb{Q}^{\theta_0}$, then $\theta \sim \theta_0$. The mapping $\theta \to \log\left(\Lambda^{H,M}(X,\theta)\right)$ is C^{∞} on Θ, and we have

(ii) For $k = 1 \ldots, 3M-1$, we have

$$\mathbb{E}_{\mathbb{Q}_{\theta_0}}\left(\frac{\partial \log\left(\Lambda^{H,M}(X,\theta)\right)}{\partial \theta_k}\Bigg|_{\theta=\theta_0}\right) = 0 \quad \text{and} \quad \mathbb{E}_{\mathbb{Q}_{\theta_0}}\left(\left(\frac{\partial \log\left(\Lambda^{H,M}(X,\theta)\right)}{\partial \theta_k}\Bigg|_{\theta=\theta_0}\right)^2\right) < +\infty.$$

(iii) For $k, j = 1, \ldots, 3M-1$, we have

$$\mathbb{E}_{\mathbb{Q}_{\theta_0}}\left(\left|\frac{\partial^2 \log \Lambda^{H,M}(X,\theta)}{\partial \theta_k \partial \theta_j}\right|_{\theta=\theta_0}\right) < +\infty.$$

and

$$\mathbb{E}_{\mathbb{Q}_{\theta_0}}\left(\frac{\partial \log \Lambda^{H,M}(X,\theta)}{\partial \theta_k}\Bigg|_{\theta=\theta_0} \frac{\partial \log \Lambda^{H,M}(X,\theta)}{\partial \theta_j}\Bigg|_{\theta=\theta_0}\right) = -\mathbb{E}_{\mathbb{Q}_{\theta_0}}\left(\frac{\partial^2 \log \Lambda^{H,M}(X,\theta)}{\partial \theta_k \partial \theta_j}\Bigg|_{\theta=\theta_0}\right)$$

Proof

Let fix $i = 1, ..., N$ and two parameter values θ and θ_0. Our objectives is to show that if $\mathbb{Q}_i^{\theta} = \mathbb{Q}_i^{\theta_0}$, then the parameters θ and θ_0 must be identical.

Since $\Lambda^{H,M}(X_i,\theta)$ and $\Lambda^{H}(X_i,\tau_\ell)$ depend on the observation $\mathbf{X_i}$ through the pair of statistics $U(\mathbf{X_i})$ = U_i and $V(\mathbf{X_i})$ =V_i. By this, we can simplify the notation by writing:

$$\Lambda^{H,M}(X_i,\theta) = \Lambda^{H,M}(U_i,V_i,\theta) \;\text{ and }\; \Lambda^{H}(X_i,\tau_\ell) = \Lambda^{H}(U_i,V_i,\tau_\ell).$$

Here $\Lambda^{H,M}(u,v,\theta)$ is the density of the pair (U_i,V_i) under $\mathbb{Q}_i^{\theta_0}$ relative to its density under $\mathbb{Q}_i^{0}$. Consequently, $\mathbb{Q}_i^{\theta} = \mathbb{Q}_i^{\theta_0}$ implies that: for all $u \in \mathbb{R}$ and $v > 0$, $\Lambda^{H,M}(u,v,\theta) = \Lambda^{H,M}(u,v,\theta_0.)$ This leads to the equality:

$$\sum_{\ell=1}^{M} \pi_\ell \frac{1}{\sqrt{1+\sigma_\ell^2 v}} \exp\left(-\frac{v(u/v-\mu_\ell)^2}{2(1+\sigma_\ell^2 v)}\right) = \sum_{\ell=1}^{M} \pi_{\ell,0} \frac{1}{\sqrt{1+\sigma_{\ell,0}^2 v}} \exp\left(-\frac{v(u/v-\mu_{\ell,0})^2}{2(1+\sigma_{\ell,0}^2 v)}\right). \tag{10}$$

To simplify (10), we introduce the following terms:

$$p(v) = \prod_{\ell=1}^{M} \sqrt{1+\sigma_\ell^2 v}, \qquad q_\ell(v) = \prod_{\ell' \neq \ell} \sqrt{1+\sigma_{\ell'}^2 v},$$

and $p_0(v)$ as well as $q_{\ell,0}(v)$ replacing σ_ℓ by $\sigma_{\ell,0}$, respectively in $p(v)$ and $q_\ell(v)$ These terms do not depend on ℓ, and by (10) we have:

$$\frac{p_0(v)}{p(v)} = \frac{\sum\limits_{\ell=1}^{M} \pi_{\ell,0} q_{\ell,0}(v) \exp\left(-\frac{v(u/v-\mu_{\ell,0})^2}{2(1+\sigma_{\ell,0}^2 v)}\right)}{\sum\limits_{\ell=1}^{M} \pi_\ell q_\ell(v) \exp\left(-\frac{v(u/v-\mu_\ell)^2}{2(1+\sigma_\ell^2 v)}\right)}.$$

The left-hand member depends only on v, while the right-hand member depends on both u and v. This can only be true if $p_0(v) = p(v)$ for all $v > 0$, hence, we conclude that $\{\sigma_1^2, \ldots, \sigma_M^2\} = \{\sigma_{1,0}^2, \ldots, \sigma_{M,0}^2\}$. Up to a permutation of terms, this confirms the equality of variances. Fixing v and setting $\sigma_\ell^2(v) = (\sigma_\ell^2 v + 1)v$. Then the following equality holds:

$$\sum_{\ell=1}^{M} \pi_\ell q_\ell(v) \exp\left(-\frac{v(u/v-\mu_\ell)^2}{2(1+\sigma_\ell^2 v)}\right) = \sqrt{2\pi v}p(v) \sum_{\ell=1}^{M} \pi_\ell \varrho(u, (\mu_\ell v, \sigma_\ell^2(v))), \tag{11}$$

where $\varrho(u, (m, \sigma^2))$ is Gaussian density at u with usual parameters m and σ^2. Similarly to (11), we have

$$\sum_{\ell=1}^{M} \pi_{\ell,0} q_{\ell,0}(v) \exp\left(-\frac{v(u/v-\mu_{\ell,0})^2}{2(1+\sigma_{\ell,0}^2 v)}\right) = p(v)\sqrt{2\pi v} \sum_{\ell=1}^{M} \pi_{\ell,0} \varrho(u, (\mu_{\ell,0} v, \sigma_\ell^2(v))).$$

For all $v > 0$, we now conclude the equality by using :

$$\sum_{\ell=1}^{M} \pi_\ell \varrho(u, (\mu_\ell v, \sigma_\ell^2(v))) = \sum_{\ell=1}^{M} \pi_{\ell,0} \varrho(u, (\mu_{\ell,0} v, \sigma_\ell^2(v))).$$

This proves that two Gaussian mixture distributions with the same variances $v(1 + \sigma_\ell^2 v)$ and means $\mu_\ell v$ and $\mu_{\ell,0} v$, respectively, in addition to π_ℓ and $\pi_{\ell,0}$. So, by identifiability of Gaussian mixtures, we have that:

$$\{(\mu_\ell, \pi_\ell), \ell = 1, \ldots, M\} = \{(\mu_{\ell,0}, \pi_{\ell,0}), \ell = 1, \ldots, M\},$$

which implies that $\theta \sim \theta_0$.
For the outer point in the proposition, we rely on the result from Proposition 3.1 in [12], and we have. The relevant findings are as follows: For any $\tau = (\mu, \sigma^2)$ and $u \in \mathbb{R}$, for $i = 1, 2..., N$, recall that $\mathbb{Q}_i^\tau$ is the distribution of X_i when Φ's is gaussian with mean μ and variance σ then we have:

$$\mathbb{E}_{\mathbb{Q}_i^\tau}\left(\exp\left(u\frac{U_i}{1+\sigma^2 V_i}\right)\right) < +\infty.$$

This implies for all integer $m > 0$ $\quad \mathbb{E}_{\mathbb{Q}_i^\tau}\left|\frac{U_i}{1+\sigma^2 V_i}\right|^m < +\infty$. Setting $\gamma(\tau) = \frac{U_i - \mu V_i}{1+\sigma^2 V}$, $\quad I(\sigma^2) = \frac{V_i}{1+\sigma^2 V_i}$, $\gamma(\tau)$, under $\mathbb{Q}_i^\tau$, and $I(\sigma^2)$ is bounded, has moments of all orders. So the following expectation exists

$$\mathbb{E}_{\mathbb{Q}_i^\tau}(\gamma(\tau)) = 0, \quad \mathbb{E}_{\mathbb{Q}_i^\tau}\left(\gamma^2(\tau) - I(\sigma^2)\right) = 0,$$

$$\mathbb{E}_{\mathbb{Q}_i^\tau}\left(\frac{1}{2}\left(\gamma^2(\tau) - I(\sigma^2)\right)^2 - \gamma^2(\tau)I(\sigma^2) - \frac{1}{2}I^2(\sigma^2)\right) = 0,$$

$$\mathbb{E}_{\mathbb{Q}_i^\tau}\left(\frac{1}{2}\gamma^3(\tau) - \frac{3}{2}\gamma(\tau)I(\sigma^2)\right) = 0.$$

For $\ell = 1, \ldots, M-1$, the derivative of $\Lambda^{H,M}(X_i, \theta)$ w.r.t π_ℓ is given by $\frac{\partial \Lambda^{H,M}(X,\theta)}{\partial \pi_\ell} = \Lambda^H(X_i, \tau_\ell) - \Lambda^H(X, \tau_M)$. Since $\mathbb{Q}_i^\tau$ has density $\Lambda^H(X_i, \tau)$ w.r.t $\mathbb{Q}$, and $\int_{C_T} \frac{\partial \Lambda^{H,M}(X,\theta)}{\partial \pi_\ell} d\mathbb{Q} = \int_{C_T} \frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \pi_\ell} d\mathbb{Q}_i^\theta = 0$.

Furthermore, because $\frac{\Lambda^H(X,\tau_\ell)}{\Lambda^{H,M}(X,\theta)} \leq \pi_\ell^{-1}$, we have:

$$\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \pi_\ell}\right)^2 \Lambda^{H,M}(X,\theta) = \frac{\left(\frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \pi_\ell}\right)^2}{\Lambda^{H,M}(X_i,\theta)} \leq \frac{2}{\pi_\ell}\Lambda^H(X,\tau_\ell) + \frac{2}{\pi_M}\Lambda^H(X,\tau_M).$$

Thus: $\mathbb{E}_{\mathbb{Q}_i^\theta}\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \pi_\ell}\right)^2 \leq \frac{2}{\pi_\ell} + \frac{2}{\pi_M}$. Higher-order derivatives with respect to π_ℓ are nulls: $\frac{\partial^2 \Lambda^{H,M}(X_i,\theta)}{\partial \pi_\ell \partial \pi_{\ell'}} = 0$.
Next, we study the partial derivatives w.r.t. the parameters μ_ℓ and σ_ℓ for a fixed $l = 1, 2, ..., M$,

$$\frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell} = \pi_\ell \frac{\partial \Lambda^H(X_i,\tau_\ell)}{\partial \mu_\ell} = \pi_\ell \gamma(\tau_\ell)\Lambda^H(X,\tau_\ell).$$

From the integrability properties, we obtain: $\mathbb{E}_{\mathbb{Q}_i^{\tau_\ell}}|\gamma(\tau_\ell)| < +\infty, \quad \mathbb{E}_{\mathbb{Q}_i^{\tau_\ell}}\gamma(\tau_\ell) = 0$, it follows that:

$$\int_{C_T} \left| \frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell} \right| d\mathbb{Q} < +\infty, \quad \int_{C_T} \frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell} d\mathbb{Q} = 0.$$

Furthermore, we obtain: $\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell}\right)^2 \Lambda^{H,M}(X_i,\theta) = \frac{\left(\frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell}\right)^2}{\Lambda^{H,M}(X_i,\theta)} \le \pi_\ell \gamma^2(\tau_\ell)\Lambda^H(X_i,\tau_\ell)$. First, notice that: $\mathbb{E}_{\mathbb{Q}^\theta}\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell}\right)^2 \le \pi_\ell \mathbb{E}_{\mathbb{Q}_i^{\tau_\ell}}(\gamma^2(\tau_\ell)) = \pi_\ell I(\omega_\ell^2)$. Next, we compute the derivative w.r.t σ_ℓ and we have $\frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \sigma_\ell^2} = \pi_\ell \frac{\partial \Lambda^H(X_i,\tau_\ell)}{\partial \sigma_\ell^2} = \pi_\ell \frac{1}{2}(\gamma^2(\tau_\ell) - I(\sigma_\ell^2))\Lambda^H(X,\tau_\ell)$.
Therefore, we obtain $\mathbb{E}_{\mathbb{Q}_i^\theta}\left|\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \sigma_\ell^2}\right| < \infty, \quad \mathbb{E}_{\mathbb{Q}_i^\theta}\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \sigma_\ell^2} = 0$.
Moreover, $\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \sigma_\ell^2}\right)^2 \Lambda^{H,M}(X_i,\theta) \le \pi_\ell \left[\frac{1}{2}(\gamma^2(\tau_\ell) - I(\sigma_\ell^2))\right]^2 \Lambda^H(X_i,\tau_\ell)$.
This implies $\mathbb{E}_{\mathbb{Q}^\theta}\left(\frac{\partial \log \Lambda^{H,M}(X_i,\theta)}{\partial \sigma_\ell^2}\right)^2 < \infty$.

Now, we examine the second-order derivatives. The successive derivatives with respect to μ_ℓ, $\mu_{\ell'}$, σ_ℓ, and $\sigma_{\ell'}$ are zero for all $\ell \neq \ell'$. We compute: $\frac{\partial^2 \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell^2} = \pi_\ell \frac{\partial^2 \Lambda^H(X_i,\tau_\ell)}{\partial \mu_\ell^2} = \pi_\ell(\gamma^2(\tau_\ell) - I(\sigma_\ell^2))\Lambda^H(X_i,\tau_\ell)$. Consequently, we have: $\frac{\partial^2 \log \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell^2}\Lambda^{H,M}(X_i,\theta) = \frac{\partial^2 \Lambda^{H,M}(X,\theta)}{\partial \mu_\ell^2} - \left(\frac{\partial \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell}\right)^2 \Lambda^{H,M}(X_i,\theta)$, Then we find $\mathbb{E}_{\mathbb{Q}_i^\theta}\frac{\partial^2 \log \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell^2} = -\mathbb{E}_{\mathbb{Q}_i^\theta}\left(\frac{\partial \log \Lambda^H(X,\theta)}{\partial \mu_\ell}\right)^2$. Next, we compute the mixed second-order derivatives:

$$\frac{\partial^2 \Lambda^{H,M}(X_i,\theta)}{\partial \mu_\ell \partial \sigma_\ell^2} = \pi_\ell \frac{\partial^2 \Lambda^H(X,\tau_\ell)}{\partial \mu_\ell \partial \sigma_\ell^2} = \pi_\ell \left(\frac{1}{2}\gamma^3(\tau_\ell) - \frac{3}{2}\gamma(\tau_\ell)I(\sigma_\ell^2)\right)\Lambda^H(X_i,\tau_\ell).$$

Finally, we have the second derivative with respect to σ_ℓ^2:

$$\frac{\partial^2 \Lambda^{H,M}(X_i,\theta)}{\partial (\sigma_\ell^2)^2} = \pi_\ell \left[\left(\frac{1}{2}(\gamma^2(\tau_\ell) - I(\sigma_\ell^2))\right)^2 - \gamma^2(\tau_\ell)I(\sigma_\ell^2) - \frac{1}{2}I^2(\sigma_\ell^2)\right]\Lambda^H(X_i,\tau_\ell).$$

□

Theorem 2.4
Assuming that $I(\theta_0)$ is non-singular. Then, there exists an estimator $\hat{\theta}_N$ that satisfies, with probability approaching 1 as $N \to \infty$, $\frac{\partial L^H(X_1,\cdots,X_N,\theta)}{\partial \theta} = 0$. Furthermore, $\hat{\theta}_N$ converges in probability to θ_0, as $N \to \infty$.

Proof
To establish weak consistency of our estimator, we verify the classical uniformity condition over the parametre space. Specifically, we show that there exists an open convex set $S \subseteq \Theta$ around the true parametre vector θ_0, and an integrable functions $G_{k,j,\ell}(X)$ that satisfy the following inequlity for all $\theta \in S$:

$$\left|\frac{\partial^3 \log \Lambda^{H,M}(X_i,\theta)}{\partial \theta_k \partial \theta_j \partial \theta_\ell}\right| \le G_{k,j,\ell}(X_i),$$

where $\mathbb{E}_{\mathbb{Q}_i^{\theta_0}}(G_{k,j,\ell}(X_i)) < +\infty$ for all indices $1 \le k,j,\ell \le 3M-1$. Furthermore, let $K, \alpha, \beta, c_0, c_1$ be positive constants satisfying $0 < \alpha < \beta < 1$ and $0 < c_0 < c_1$. Assume that θ_0 lies within the following set:

$$S = \{(\pi_\ell, \mu_\ell, \sigma_\ell^2)_{1 \le \ell \le M} : \alpha < \pi_\ell < \beta, \, |\mu_\ell| < K, \, c_0 < \sigma_\ell^2 < c_1, \, 1 \le \ell \le M\}.$$

whith $\pi_M = 1 - \sum_{\ell=1}^{M-1} \pi_\ell$ as before. A direct computation gives the third derivative of the log-likelihood ratio as follow:

$$\begin{aligned}\frac{\partial^3 \log \Lambda^{H,M}(X_i,\theta)}{\partial\theta_j\partial\theta_k\partial\theta_r} &= \frac{1}{\Lambda^{H,M}(X_i\theta)}\frac{\partial^3\Lambda^{H,M}(X_i,\theta)}{\partial\theta_j\partial\theta_k\partial\theta_r} - \frac{1}{(\Lambda^{H,M}(X_i,\theta))^2}\frac{\partial\Lambda^{H,M}(X_i,\theta)}{\partial\theta_r}\frac{\partial^2\Lambda^{H,M}(X_i,\theta)}{\partial\theta_j\partial\theta_k}\\ &- \frac{1}{(\Lambda^{H,M}(X_i,\theta))^2}\left(\frac{\partial\Lambda^{H,M}(X_i,\theta)}{\partial\theta_k}\frac{\partial^2\Lambda^{H,M}(X_i,\theta)}{\partial\theta_j\partial\theta_r}\right.\\ &- \left.\frac{\partial\Lambda^{H,M}(X_i,\theta)}{\partial\theta_j}\frac{\partial^2\Lambda^{H,M}(X_i,\theta)}{\partial\theta_r\partial\theta_k}\right)\\ &+ \frac{2}{(\Lambda^{H,M}(X_i,\theta))^3}\frac{\partial\Lambda^{H,M}(X_i,\theta)}{\partial\theta_j}\frac{\partial\Lambda^{H,M}(X_i\theta)}{\partial\theta_k}\frac{\partial\Lambda^{H,M}(X_i,\theta)}{\partial\theta_r}.\end{aligned}$$

For a different j,k,r among the mixture weights, we have

$$\begin{aligned}\frac{\partial^3 \log \Lambda^{H,M}(X_i,\theta)}{\partial\pi_j\partial\pi_k\partial\pi_r} = &\frac{2}{(\Lambda^{H,M}(X_i,\theta))^3}\left(\Lambda^H(X_i,\tau_j) - \Lambda^H(X_i,\tau_M)\right)\left(\Lambda^H(X_i,\tau_k) - \Lambda^H(X_i,\tau_M)\right)\\ &\left(\Lambda^H(X_i,\tau_r) - \Lambda^H(X_i,\tau_M)\right).\end{aligned}$$

Since $\frac{\Lambda^H(X_i,\tau_j)}{\Lambda^{H,M}(X_i,\theta)} \le \pi_j^{-1} < \alpha^{-1}$, it follows that: $\frac{\partial^3 \log \Lambda^{H,M}(X_i,\theta)}{\partial\pi_j\partial\pi_k\partial\pi_r} \le \frac{2^4}{\alpha^3}$. To find upper bounds for the other third-order derivatives, we remark the appearance of random variables: $\gamma^m(\tau) = \left(\frac{U_i-\mu V_i}{1+\sigma^2 V_i}\right)^m$, for various values of m. We now bound $\gamma(\tau)$ by a random variable independent of τ with moments of any order w.r.t $\mathbb{Q}^{\theta_0}$. Note that: $\frac{U_i}{1+\sigma^2V_i} = \frac{U_i}{1+c_1V_i}\left(1+\frac{(c_1-\sigma^2)V_i}{1+\sigma^2V_i}\right)$. Thus $|\gamma(\tau)| \le \frac{c_1}{c_0}\left|\frac{U_i}{1+c_1V_i}\right| + \frac{K}{c_0}$. Similarly:

$$\left|\frac{U_i}{1+c_1V_i}\right| = \left|\frac{U_i}{1+\sigma^2V_i}\left(\frac{1+(\sigma^2-c_1)V_i}{1+c_1V_i}\right)\right| \le 3\left|\frac{U_i}{1+\sigma^2V_i}\right|.$$

Therefore, for all τ: $\mathbb{E}_{\mathbb{Q}_i^\tau}\left|\frac{U_i)}{1+c_1V_i}\right|^m < +\infty$. Consequently: $\mathbb{E}_{\mathbb{Q}_i^{\theta_0}}\left|\frac{U_i}{1+c_1V_i}\right|^m = \sum_{\ell=1}^M \pi_{\ell,0}\mathbb{E}_{\mathbb{Q}_i^{\tau_0}}\left|\frac{U_i}{1+c_1V_i}\right|^m < +\infty$. By this, the proof of Theorem 2.4 is now complete. □

3. Simulation and algorithm of estimation

In this section, we will do some numerical simulations of the model. Since we can only get discrete data from the simulation, let us start with a simple discussion of the discrete case. Suppose that we obtain the process $X_i(t)$ by observing at times $t_k^n = t_k = k\frac{T}{n}, k = 1,2,\ldots,n+1$ simultaneously. We give the calculation of $\chi_i(t)$ and $Z_i(t)$ to establish estimators $\hat{\theta}_N^{(n)}$. For $i = 1,2,...,N$, we have

$$\chi_i(t_k) = \frac{\int_0^{t_k} k_H(t_k,s)\frac{X_i(s)}{\sigma}ds - \int_0^{t_{k-1}} k_H(t_{k-1},s)\frac{X_i(s)}{\sigma}ds}{\omega_i^H(t_k) - \omega_i^H(t_{k-1})},$$

$$\tilde{X}_i(t_k) \approx \sum_{j=2}^{k-1} k_H^{-1} t_j^{\frac{1}{2}-H}(t_k-t_j)^{\frac{1}{2}-H}\sigma^{-1}(X_i(t_{j+1}) - X_i(t_j)),$$

where

$$\int_0^{t_k} k_H(t_k,s)\frac{X_i(s)}{\sigma}ds \approx \sigma^{-1}\sum_{j=2}^{k-1} k_H^{-1} t_j^{\frac{1}{2}-H}(t_k-t_j)^{\frac{1}{2}-H}(t_{j+1}-t_j)X_i(t_j).$$

Then the random variables U_i and V_i, are, respectively, replaced by

$$U_i^n \approx \sum_{k=1}^{n} \chi_i(t_k)(\tilde{X}_i(t_{k+1}) - \tilde{X}_i(t_k)) - \sum_{k=1}^{n} \chi_i(t_k)(\tilde{L}_i(t_{k+1}) - \tilde{L}_i(t_k)), \tag{12}$$

and

$$V_i^n \approx \sum_{k=1}^{n} \chi_i^2(t_k)(\omega_i^H(t_{k+1}) - \omega_i^H(t_k)). \tag{13}$$

3.1. Algorithm of estimation

Considering ϕ_i as a mixture distribution and assuming that the population of individuals is partitioned into M distinct sub-groups. To formalize this framework, we introduce a discrete random variable $Y_i \in \{1, \dots, M\}$ for each i, we have

$$\mathbb{P}_i^\theta(\phi_i \in d\phi | Y_i = \ell) = \mathcal{N}(\mu_\ell, \sigma_\ell^2) \quad \text{and} \quad \mathbb{P}_i^\theta(Y_i = \ell) = \pi_\ell.$$

We assume $((\phi_1, Y_1), (\phi_2, Y_2), ..., (\phi_N, Y_N))$ is a vector of independent bivariate random variables which is independent of the vector $(B_1^H, \dots, B_N^H)$. Since the variables $Y_1, \dots, Y_N$ are unobserved, this is exaclty the missing data setting of the **EM** algorithm described in [14] and [13], where $(X_i)_{i=1,2,\dots,N}$ represents the incomplete data. For that, we define the indicator vectors $Z_i = (Z_{i,1}, \dots, Z_{i,M}) \in \{0,1\}^M$ such that $Z_{i\ell} = 1$ if $Y_i = \ell, \ell = 1, \dots, M$. Under this parametrization the log-likelihood of the complete data $(\mathbf{X_i}, \mathbf{Z_i})$ takes the form:

$$\mathcal{L}_N((X_1, Z_1), \dots, (X_N, Z_N); \theta) = \sum_{i=1}^{N} \sum_{\ell=1}^{M} Z_{i\ell} \log(\pi_\ell \Lambda_i^H(X_i, \tau_\ell)). \tag{14}$$

The **EM** algorithm is an iterative procedure that alternates between two key steps: the Expectation step (E-step) and the Maximization step (M-step). For E-step, we compute $G^H(\theta, \theta')$, which is the expected value of the complete data likelihood function, given by

$$G^H(\theta, \theta') = \mathbb{E}\left(\mathcal{L}_N((X_i, Z_i); \theta) \mid X_i; \theta'\right),$$

where $\mathbb{E}(\cdot \mid X_i; \theta')$, given the observed data X_i, is the conditional expectation, computed using the distribution of all data, under the parameter value θ'. For the E-step, we specifically compute:

$$G^H(\theta, \theta') = \sum_{i=1}^{N} \sum_{\ell=1}^{M} \tilde{\pi}_\ell(X_i, \theta') \log\left(\pi_\ell \Lambda_i^H(X_i, \tau_\ell)\right),$$

where $\tilde{\pi}_\ell(X_i, \theta')$ is the posterior probability, and is given by the following equation:

$$\tilde{\pi}_\ell(X_i, \theta') := \mathbb{P}(Z_{i\ell} = 1 \mid X_i, \theta') = \frac{\pi'_\ell \Lambda_i^H(X_i, \tau'_\ell)}{\Lambda_i^{H,M}(X_i, \theta')} = \frac{\pi'_\ell \Lambda_i^H(X_i, \tau'_\ell)}{\sum\limits_{r=1}^{M} \Lambda_i^H(X_i, \tau'_r)}. \tag{15}$$

At iteration m of the **EM** algorithm, the E-step updates the parameter vector θ by maximizing the objective function $G^H(\theta, \theta^{(m)})$, where $\theta^{(m)}$ is the the current parameter estimate and we have:

$$\theta^{(m+1)} = \arg\max_{\theta \in \Theta} G^H(\theta, \theta^{(m)})$$

This maximization can be divided into two independent optimisation problems: one optimizing over the mixing weight π_ℓ, and the second optimizing over the parameters $\tau_\ell = (\mu_\ell, \sigma_\ell^2)$. For the π_ℓ and the contraint $\sum_{\ell=1}^{M} \pi_\ell = 1$

via a lagrange multiple α leads to find the stationary point of the equation:

$$\frac{\partial}{\partial \pi_\ell}\left[\sum_{i=1}^{N}\sum_{\ell=1}^{M}\tilde{\pi}_\ell(X_i,\theta^{(m)})\log(\pi_\ell)+\alpha\left(\sum_{\ell=1}^{M}\pi_\ell-1\right)\right]=0.$$

which yields the standard explicit solution $\pi_\ell^{(m+1)}=\frac{1}{N}\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})$. When the variances σ_ℓ^2 for $\ell=1,\ldots,M$ are treated as fixed and known, the coordinate maximizer for μ_ℓ admits the solution:

$$\hat{\mu}_\ell^{(m+1)}=\left(\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})(1+\sigma_\ell^2V_i)^{-1}\frac{U_i}{V_i}\right)\left(\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})(1+\sigma_\ell^2V_i)^{-1}\right)^{-1}.$$

Conversely, if σ_ℓ is not known, μ_ℓ and σ_ℓ^2 do not decouple. In this case, they must be updated jointly by solving the following system, using the function *scipy.optimize.minimize* in Python:

$$\begin{cases}\hat{\mu}_\ell^{(m+1)}=\left(\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})(1+\sigma_\ell^{2\,(m+1)}V_i)^{-1}\frac{U_i}{V_i}\right)\left(\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})(1+\sigma_\ell^{2\,(m+1)}V_i)^{-1}\right)^{-1};\\[2ex] \sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})\frac{V_i}{1+\sigma_\ell^{2\,(m+1)}V_i}=\sum_{i=1}^{N}\tilde{\pi}_\ell(X_i,\theta^{(m)})\frac{V_i^2}{(1+\sigma_\ell^{2\,(m+1)}V_i)^2}\left(\mu_\ell^{(m+1)}-\frac{U_i}{V_i}\right)^2.\end{cases}$$

Since $\Theta\subset\mathbb{R}^{3M-1}$, $\Lambda^{H,M}(X_i,\theta)$ is continuous on Θ and differentiable on the interior of Θ, $G^H(\theta,\theta')$ is continuous with respect to both θ and θ' $\left.\dfrac{\partial G^H(\theta,\hat{\theta}^{(m)})}{\partial\theta}\right|_{\theta=\hat{\theta}^{(m+1)}}=0$, and $\dfrac{\partial G^H(\theta,\theta')}{\partial\theta}$ is continuous in both θ and θ' as well as $\Theta_{\theta_0}=\{\theta\in\Theta, L_N^H(X_1,X_2,...,X_N,\theta)\geq L_N^H(X_1,X_2,...,X_N,\theta_0)\}$ is a compact set if $L_N^H(X_1,X_2,...,X_N,\theta_0)>-\infty$; we obtain the statement below, which confirms the stationarity of the EM sequence. The first five previous conditions are directly deduced from the regularity of the likelihood (see Proposition 2.3); for the last condition, the reader can see [13, section 3] (see also [28, 42, 6] for more details).

Proposition 3.1
The sequence $(\hat{\theta}^{(m)})_{m\in\mathbb{N}^*}$ generated by the **EM** algorithm converges to a point that maximizing the likelihood.

Let $\widehat{\theta}$ denote the parameter estimator returned by the **EM** algorithm. For each subject i, it is useful to determine the cluster to which they belong. This is done by estimating the posterior probabilities $\tilde{\pi}_\ell(X_i,\theta)$ (as definet in (15)). Generally, the subject i is assigned to the cluster ℓ that corresponds to the highest posterior probability,

$$\tilde{\pi}_\ell(X_i,\tilde{\theta})=\max_{1\leq\ell'\leq M}\tilde{\pi}_{\ell'}(X_i,\tilde{\theta}). \tag{16}$$

When M is unknown, it cannot be estimated simultaneously with θ; we first compute the estimator $\theta=\theta(M)$ across a range of condidate values for M, and perform model selection using the Bayesian Information Criterion (**BIC**), and we have: **BIC**$(M,\theta)=-2\log(\mathcal{L}_N(\theta))+(3M-1)\log(N)$, and choose the value of M that minimizing the **BIC**(M,θ), then

$$\hat{M}=\arg\min_{M\in\{1,2,\ldots,M_N\}}\mathbf{BIC}(M,\hat{\theta}(M)), \tag{17}$$

where $\hat{\theta}(M)$ is the final iteration of the **EM** algorithm for M components; this procedure is justified by the known asymptotic properties of the **BIC** in finite mixture models: as $N\longrightarrow+\infty$, the **BIC** is almost surely does not overestimate the number of components, and the selected $\hat{\theta}(M)$ almost surely converges to the true number of mixture components M (see [26, 22]).

3.2. Numerical simulation

In this section, we apply the proposed methods to simulated data. To achieve this, we consider two cases: the first assumes the number of components M is known, and the second assumes M is unknown; so we treat it as a parameter to be carefully estimated. We simulate paths on 1000 observation time points on the interval $[0, T]$, with $T = 1$, $T = 5$ and $T = 10$, by varying the number of subjects N; $N = 100$, $N = 500$ and $N = 1000$. We do this for two values of H: $H = 0.3$ (short memory) and $H = 0.7$ (long memory). For the first case, we assume that the random effects follow two examples of mixed Gaussian distribution with three components ($M = 3$): $\Phi_i \hookrightarrow 0.3\mathcal{N}(-1.3, 0.25) + 0.4\mathcal{N}(1.2, 0.16) + 0.3\mathcal{N}(3.7, 0.25)$ and $\Phi_i \hookrightarrow 0.3\mathcal{N}(-1.8, 1) + 0.4\mathcal{N}(1.2, 0.16) + 0.3\mathcal{N}(3.7, 0.09)$. We simulated a dataset of 200 repetitions for the model (4) with $\sigma = 1$ and computed the classification rate **CR**, defined as the proportion of trajectories correctly assigned to their subjects of origin.Tables 1 and 2 show the means and standard deviations of the 200 estimators, respectively, for the two above examples of a Gaussian mixed distribution, in the case of well-separated components. The tables illustrate that the estimates are close to the true values. For the second case, we simulated the model (4) with the true number of components $M = 4$ and for distribution of Φ_i's: $\Phi_i \hookrightarrow 0,2\mathcal{N}(-1.8, 0.25) + 0,3\mathcal{N}(1.5, 0.04) + 0,3\mathcal{N}(2.7, 0.25) + 0,2\mathcal{N}(5.8, 0.09)$. According to (17), we select the estimators of M across a hundred tests. Figure 1 and 2 presents the model selection performance simulation scenarios for $H = 0.3$ and $H = 0.7$. Each figure presents both the BIC evolution curves and the corresponding selection frequency histograms over a candidate set $M \in \{1, 2, \ldots, 6\}$. For each candidate, the **EM** algorithm was used to maximize the log-likelihood, and the procedure was repeated a hundred times. The histograms provide empirical evidence of the consistency of the **BIC** method, showing the convergence to the true number of components M. Specifically, the global minimum of the **BIC** curves and the dominant mode of the frequency histograms consistently identify the global minimum of the **BIC** curves coupled with the dominant mode of the frequency histograms, $M = 4$ as the most parsimonious structure for both $H = 0.3$ and $H = 0.7$. Tables 5 and 6 illustrate the final parameter estimates; the results demonstrate that the joint **EM** and **BIC** effectively estimate the parameters where the number of components M was unknown a priori. As illustrated by these tables, the parameter estimates are very close to the true values, with small standard deviations. Furthermore, regarding the number of components M, the number selected by the **BIC** method matched the true value with a high frequency, as also illustrated in Figures 1 and 2. Tables 3 and 4 illustrate the estimation results for the not well-separated components: $\Phi_i \hookrightarrow 0.7\mathcal{N}\left(-0.5, 0.5^2\right) + 0.3\mathcal{N}\left(-1.8, 0.5^2\right)$; these results do not differ from those obtained in the well-separated scenarios, even though the classification rate (CR) is not sufficiently high. A comparison with the estimation values for the **OU** model, given in [13, Table 1-3 bottom], reveals no significant difference in the estimation results; this indicates that the reflection term has little impact on the quality of the estimation, even though this term plays a crucial role in ensuring the positivity of the process.

Table 1. Estimation EM algorithm of component weights, means, and variances for $M = 3$ and $H = 0.3$ for different values of N and T, with standard deviation (sd), for 200 repetitions.

True value	N=500, T=1 Mean (sd)	N=1000, T=1 Mean (sd)	N=500, T=10 Mean (sd)	N=1000, T=10 Mean (sd)
$\pi_1 = 0.3$	0.2997 (0.0223)	0.3011 (0.0167)	0.3001 (0.0220)	0.3008 (0.0147)
$\mu_1 = -1.8$	-1.7987 (0.0974)	-1.7922 (0.0781)	-1.7954 (0.1074)	-1.7977 (0.0669)
$\sigma_1^2 = 1$	0.9893 (0.1612)	1.0051 (0.1070)	0.9923 (0.1425)	1.0125 (0.1067)
$\pi_2 = 0.4$	0.4007 (0.0232)	0.3987 (0.0177)	0.3992 (0.0211)	0.3980 (0.0159)
$\mu_2 = 1.2$	1.2087 (0.0467)	1.2083 (0.0295)	1.2279 (0.0355)	1.2254 (0.0265)
$\sigma_2^2 = 0.16$	0.1650 (0.0398)	0.1608 (0.0304)	0.1680 (0.0264)	0.1655 (0.0182)
$\pi_3 = 0.3$	0.2996 (0.0199)	0.3002 (0.0137)	0.3006 (0.0198)	0.3012 (0.0145)
$\mu_3 = 3.7$	3.7281 (0.0661)	3.7285 (0.0429)	3.7702 (0.0382)	3.7697 (0.0273)
$\sigma_3^2 = 0.09$	0.0941 (0.0614)	0.0965 (0.0495)	0.0946 (0.0263)	0.0938 (0.0166)
CR	0.9738%	0.9750%	0.9876%	0.9883%

Table 2. Estimation EM algorithm of component weights, means, and variances for $M = 3$ and $H = 0.7$ for different values of N and T, with standard deviation (sd), for 200 repetitions.

True	N=100, T=1	N=1000, T=1	N=100, T=5	N=1000, T=5
value	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
$\pi_1 = 0.3$	0.2974 (0.0443)	0.2995 (0.0149)	0.2970 (0.0460)	0.2993 (0.0142)
$\mu_1 = -1.3$	-1.3147 (0.0937)	-1.3013 (0.0344)	-1.2899 (0.0989)	-1.2988 (0.0300)
$\sigma_1^2 = 0.25$	0.2358 (0.0723)	0.2490 (0.0234)	0.2377 (0.0716)	0.2507 (0.0226)
$\pi_2 = 0.4$	0.4050 (0.0551)	0.3991 (0.0173)	0.4025 (0.0469)	0.4001 (0.0145)
$\mu_2 = 1.2$	1.1860 (0.0958)	1.1935 (0.0273)	1.1739 (0.0978)	1.1721 (0.0279)
$\sigma_2^2 = 0.16$	0.1483 (0.0807)	0.1582 (0.0216)	0.1560 (0.0826)	0.1506 (0.0212)
$\pi_3 = 0.3$	0.2976 (0.0532)	0.3014 (0.0149)	0.3004 (0.0439)	0.3007 (0.0147)
$\mu_3 = 3.7$	3.6865 (0.1460)	3.6815 (0.0452)	3.6597 (0.1425)	3.6443 (0.0444)
$\sigma_3^2 = 0.25$	0.2552 (0.1956)	0.2450 (0.0429)	0.2394 (0.1613)	0.2369 (0.0458)
CR	98.04 %	98.44%	98.22%	98.70 %

Table 3. Estimation EM algorithm with not well-separated components: weights, means, and variances for $M = 2$ and $H = 0.3$ for different values of N and T, with standard deviation (sd), for 200 repetitions (not well-separated components).

True	N=100, T=1	N=1000, T=1	N=100, T=10	N=1000, T=10
value	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
$\pi_1 = 0.3$	0.3846 (0.1715)	0.3130 (0.0643)	0.3412 (0.1508)	0.3028 (0.0556)
$\mu_1 = -1.8$	-1.6859 (0.3424)	-1.7779 (0.1336)	-1.7334 (0.3127)	-1.7956 (0.1163)
$\sigma_1^2 = 0.25$	0.3049 (0.1902)	0.2601 (0.0677)	0.2811 (0.1760)	0.2519 (0.0568)
$\pi_2 = 0.7$	0.6154 (0.1715)	0.6870 (0.0643)	0.6588 (0.1508)	0.6972 (0.0556)
$\mu_2 = -0.5$	-0.4540 (0.1631)	-0.4875 (0.0604)	-0.4820 (0.1427)	-0.4962 (0.0529)
$\sigma_2^2 = 0.25$	0.2313 (0.1138)	0.2502 (0.0357)	0.2456 (0.0928)	0.2534 (0.0302)
CR	84.56 %	90.08%	86.87%	90.68%

Table 4. Estimation EM algorithm with not well-separated components, of weights, means, and variances for $M = 2$ and $H = 0.7$ for different values of N and T, with standard deviation (sd), for 200 repetitions.

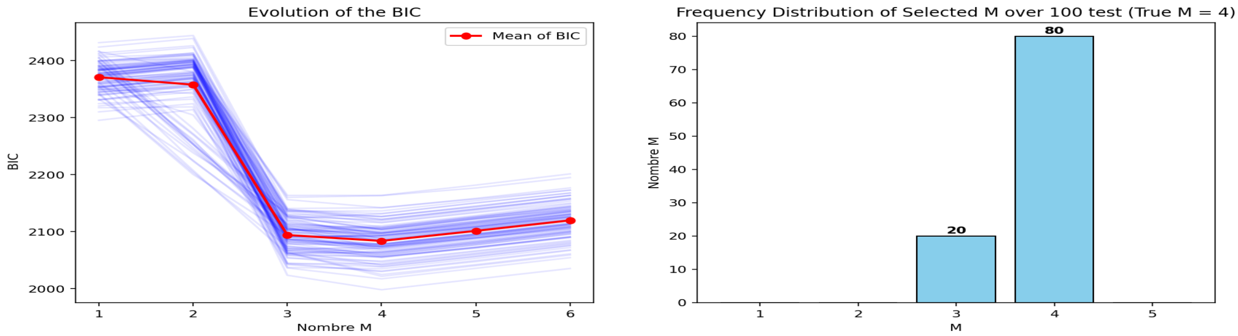
True	N=100, T=1	N=1000, T=1	N=100, T=5	N=1000, T=5
value	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
$\pi_1 = 0.3$	0.3554 (0.1518)	0.3182 (0.0579)	0.3574 (0.1642)	0.3227 (0.0592)
$\mu_1 = -1.8$	-1.7103 (0.2969)	-1.7658 (0.1233)	-1.7330 (0.3219)	-1.7551 (0.1253)
$\sigma_1^2 = 0.25$	0.2832 (0.1701)	0.2616 (0.0596)	0.2737 (0.1730)	0.2710 (0.0653)
$\pi_2 = 0.7$	0.6446 (0.1518)	0.6818 (0.0579)	0.6426 (0.1642)	0.6773 (0.0592)
$\mu_2 = -0.5$	-0.4674 (0.1367)	-0.4870 (0.0550)	-0.4710 (0.1475)	-0.4824 (0.0525)
$\sigma_2^2 = 0.25$	0.2293 (0.0959)	0.2427 (0.0333)	0.2202 (0.0911)	0.2339 (0.0309)
CR	86.72%	90.27 %	86.37%	90.53%

Table 5. Estimation EM algorithm of component weights, means, and variances with standard deviation (sd), and M with the selection percentage. For $H = 0.3$ and $M = 4$, and different for values of N and T, for 100 repetitions.

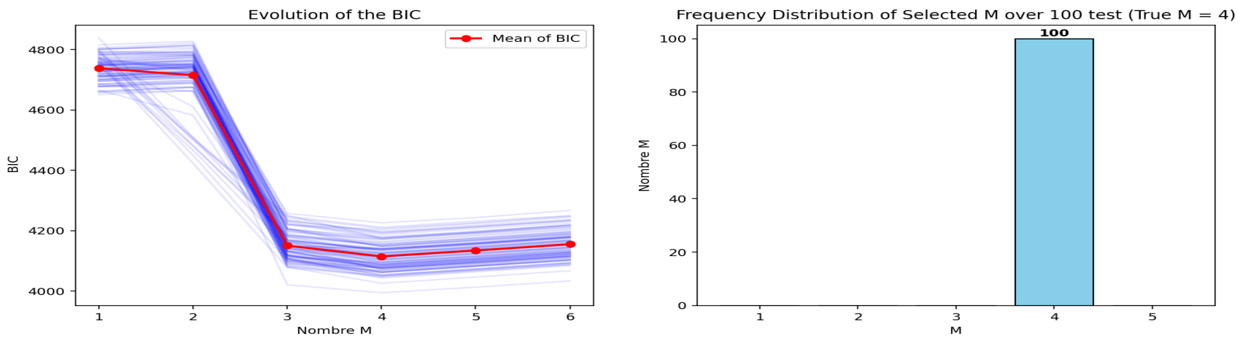
True value	N=500, T=1	N=1000, T=1	N=500, T=10	N=1000, T=10
	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
$\pi_1 = 0.2$	0.1991 (0.0182)	0.1999 (0.0118)	0.2014 (0.0177)	0.2003 (0.0114)
$\mu_1 = -1.8$	-1.8062 (0.0538)	-1.7944 (0.0342)	-1.7996 (0.0525)	-1.8032 (0.0316)
$\sigma_1^2 = 0.25$	0.2452 (0.0396)	0.2509 (0.0248)	0.2470 (0.0334)	0.2500 (0.0246)
$\pi_2 = 0.3$	0.3070 (0.0360)	0.3114 (0.0308)	0.2943 (0.0389)	0.2955 (0.0278)
$\mu_2 = 1.5$	1.5161 (0.0652)	1.5244 (0.0502)	1.5304 (0.0477)	1.5283 (0.0335)
$\sigma_2^2 = 0.04$	0.0404 (0.0262)	0.0431 (0.0206)	0.0405 (0.0178)	0.0402 (0.0116)
$\pi_3 = 0.3$	0.2889 (0.0352)	0.2874 (0.0314)	0.3048 (0.0384)	0.3011 (0.0275)
$\mu_3 = 2.7$	2.7416 (0.1158)	2.7467 (0.1022)	2.7385 (0.1051)	2.7433 (0.0757)
$\sigma_3^2 = 0.25$	0.2221 (0.1229)	0.2259 (0.0999)	0.2663 (0.0840)	0.2654 (0.0634)
$\pi_4 = 0.2$	0.2050 (0.0193)	0.2013 (0.0135)	0.1995 (0.0186)	0.2031 (0.0123)
$\mu_4 = 5.8$	5.8223 (0.1001)	5.8325 (0.0751)	5.9053 (0.0545)	5.8984 (0.0366)
$\sigma_4^2 = 0.09$	0.1270 (0.1298)	0.1123 (0.0956)	0.1022 (0.0452)	0.0980 (0.0308)
$M = 4$	$\hat{M} = 4\,(80\%)$	$\hat{M} = 4\,(100\%)$	$\hat{M} = 4\,(100\%)$	$\hat{M} = 4\,(100\%)$

Table 6. Estimation EM algorithm of component weights, means, and variances with standard deviation (sd), and M with the selection percentage. For $H = 0.7$ and $M = 4$, and different for values of N and T, for 100 repetitions.

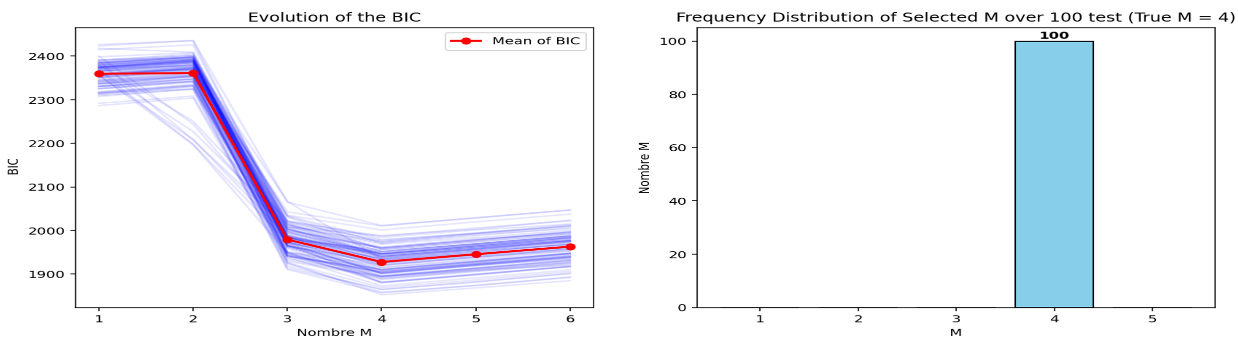
True value	N=500, T=1	N=1000, T=1	N=500, T=10	N=1000, T=10
	Mean (sd)	Mean (sd)	Mean (sd)	Mean (sd)
$\pi_1 = 0.2$	0.2002 (0.0164)	0.1996 (0.0126)	0.2004 (0.0187)	0.2004 (0.0111)
$\mu_1 = -1.8$	-1.8038 (0.0479)	-1.7980 (0.0379)	-1.8001 (0.0520)	-1.8006 (0.0366)
$\sigma_1^2 = 0.25$	0.2470 (0.0357)	0.2490 (0.0258)	0.2415 (0.0331)	0.2480 (0.0264)
$\pi_2 = 0.3$	0.3060 (0.0420)	0.3040 (0.0348)	0.3043 (0.0417)	0.3019 (0.0304)
$\mu_2 = 1.5$	1.5062 (0.0618)	1.4995 (0.0445)	1.4556 (0.0591)	1.4468 (0.0434)
$\sigma_2^2 = 0.04$	0.0391 (0.0272)	0.0392 (0.0201)	0.0359 (0.0213)	0.0339 (0.0156)
$\pi_3 = 0.3$	0.2965 (0.0429)	0.2950 (0.0354)	0.2959 (0.0440)	0.2971 (0.0316)
$\mu_3 = 2.7$	2.7087 (0.1274)	2.7094 (0.0904)	2.6359 (0.1192)	2.6395 (0.0907)
$\sigma_3^2 = 0.25$	0.2309 (0.1044)	0.2280 (0.0738)	0.2284 (0.1013)	0.2247 (0.0681)
$\pi_4 = 0.2$	0.1973 (0.0185)	0.2013 (0.0140)	0.1995 (0.0176)	0.2005 (0.0138)
$\mu_4 = 5.8$	5.7830 (0.0601)	5.7806 (0.0452)	5.6854 (0.0585)	5.6766 (0.0413)
$\sigma_4^2 = 0.09$	0.0833 (0.0514)	0.0888 (0.0390)	0.0681 (0.0438)	0.0697 (0.0326)
$M = 4$	$\hat{M} = 4\,(95\%)$	$\hat{M} = 4\,(100\%)$	$\hat{M} = 4\,(100\%)$	$\hat{M} = 4\,(100\%)$



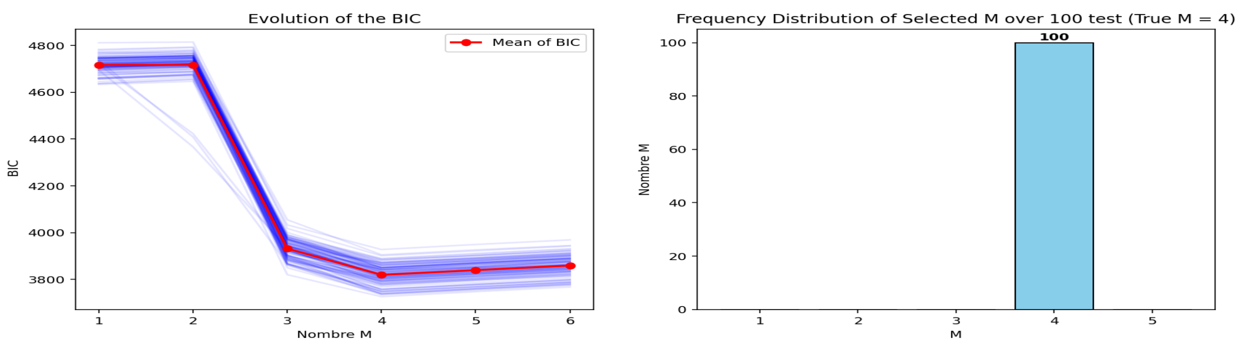
(a) $N = 500$ and $T = 1$



(b) $N = 1000$ and $T = 1$

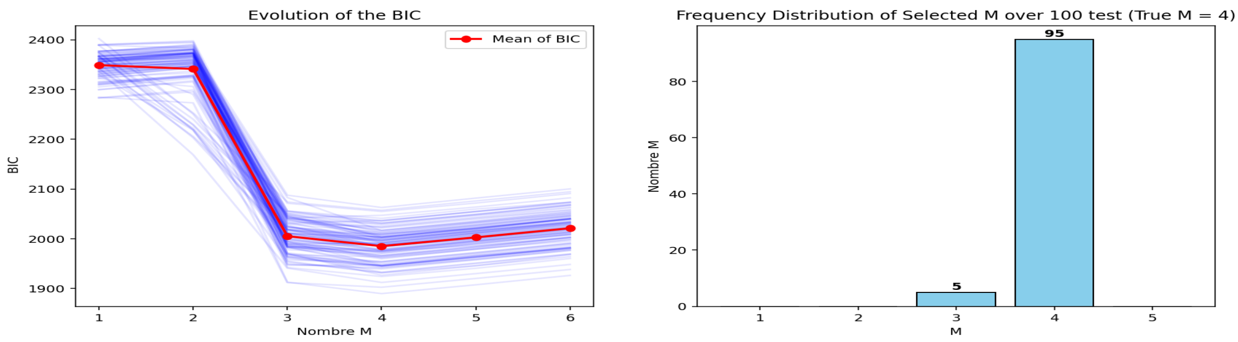


(c) $N = 500$ and $T = 10$

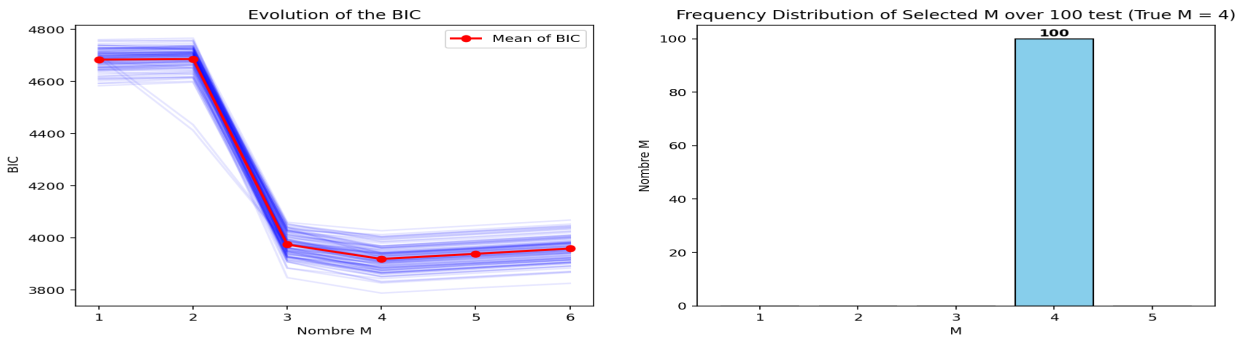


(d) $N = 1000$ and $T = 10$

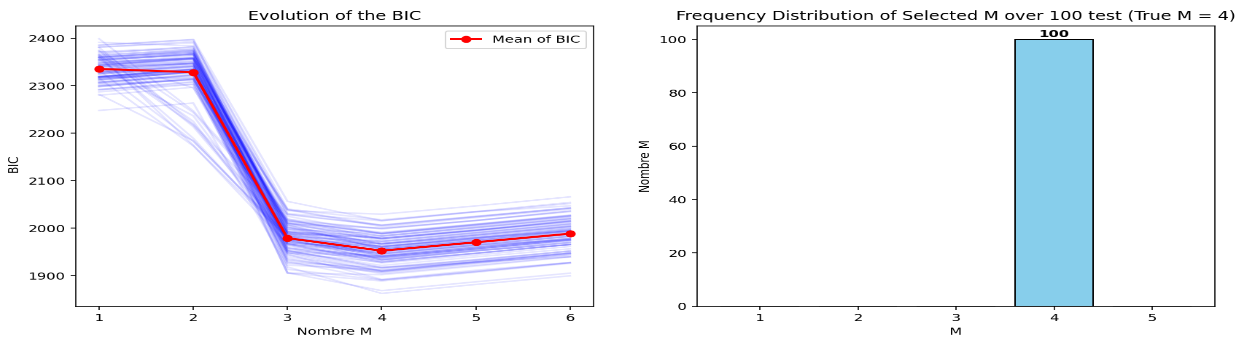
Figure 1. BIC evolution and selection frequency for $H = 0.3$ across 100 tests. (For interpretation of the references to color in the legend of this figure, the reader is referred to the electronic version of the article)



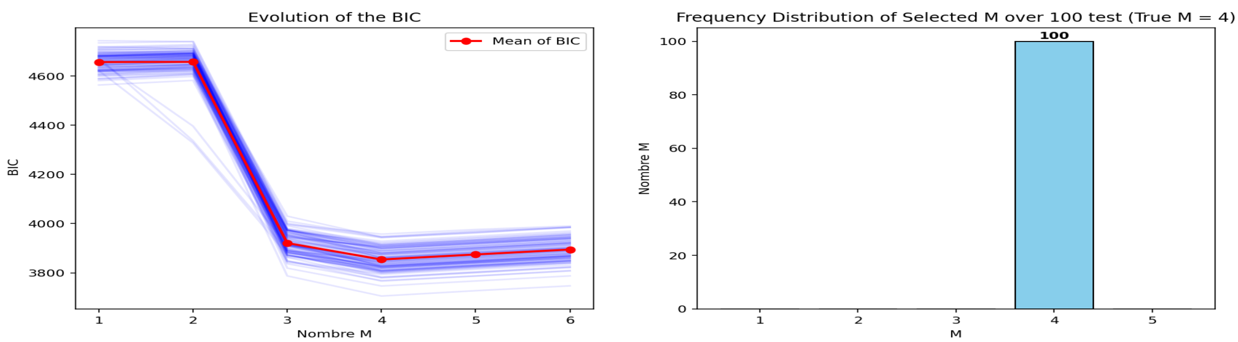
(a) $N = 500$ and $T = 1$



(b) $N = 1000$ and $T = 1$



(c) $N = 500$ and $T = 10$



(d) $N = 1000$ and $T = 10$

Figure 2. BIC evolution and selection frequency for $H = 0.7$ across 100 tests. (For interpretation of the references to color in the legend of this figure, the reader is referred to the electronic version of the article)

4. Conclusions and Discussions

We investigated parametric inference for a reflected fractional Ornstein-Uhlenbeck process with mixture Gaussian random effects. We derived an exact continuous-observation likelihood. We then obtained estimates for model parameters using the EM algorithm and demonstrated their convergence. Furthermore, we implemented numerical simulations that effectively validated the theoretical results, and we estimated the number of mixture random effects.

The proposed approach has some limitations; it does not consider the estimation of the Hurst index H and the diffusion coefficient σ. To our knowledge, and for particular fractional SDEs, with discrete observations, H and σ could be estimated separately in a preliminary step using the quadratic variations method (e.g., [3, 24]) or by the generalized method of moments (see [8, Section 6.1] and the references therein). But adapting these methods seems to be a difficult task, because here we use continuous observations for a reflected fractional **SDE**, with a mixing random effect. However, H and σ can be estimated jointly with the random effect parameters via profile or delta Laplace approximation likelihood (see [21]) or via Bayesian approaches as in [2]. However, a problem in identifying mixing parameters can arise. Therefore, the logical perspective of this paper is to address these challenges. Other improvements could be explored, notably the asymptotic normality of estimator vectors; we will address this in the near future.

REFERENCES

1. FA Attia. On a reflected ornstein-uhlenbeck process with an application. *Bulletin of the Australian Mathematical Society*, 43(3):519–528, 1991.
2. Gabriel Barragan-Ramirez, Saba Infante, Inti Becerra, and Aracelis Hernandez. A comparative review of monte carlo algorithms with numerical sde approaches for mixed effect dynamical models. *Statistics, Optimization amp; Information Computing*, 16(4):3700–3723, Jul. 2026.
3. Corinne Berzin and José R. León. Estimation in models driven by fractional Brownian motion. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 44(2):191 – 213, 2008.
4. Lijun Bo, Dan Tang, Yongjin Wang, and Xuewei Yang. On the conditional default probability in a regulated market: a structural approach. *Quantitative Finance*, 11(12):1695–1702, 2011.
5. Lijun Bo, Yongjin Wang, and Xuewei Yang. Some integral functionals of reflected sdes and their applications in finance. *Quantitative Finance*, 11(3):343–348, 2011.
6. Russell A. Boyles. On the convergence of the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(1):47–50, 1983.
7. Nesrine Chebli, Hamdi Fathallah, and Yousri Slaoui. Hybrid estimation for a mixed fractional black-scholes model with random effects from discrete time observations. *arXiv preprint arXiv:2508.07936*, 2025.
8. Nesrine Chebli, Hamdi Fathallah, and Yousri Slaoui. Random effects estimation in a fractional diffusion model based on continuous observations. *Statistical Inference for Stochastic Processes*, 28(3):13, 2025.
9. Hong Chen, David D Yao, et al. *Fundamentals of queueing networks: Performance, asymptotics, and optimization*, volume 4. Springer, 2001.
10. Min Dai, Jinqiao Duan, Junjun Liao, and Xiangjun Wang. Maximum likelihood estimation of stochastic differential equations with random effects driven by fractional brownian motion. *Applied Mathematics and Computation*, 397:125927, 2021.
11. Maud Delattre, Valentine Genon-Catalot, and Catherine Larédo. Parametric inference for discrete observations of diffusion processes with mixed effects. *Stochastic Processes and their Applications*, 128(6):1929–1957, 2018.
12. Maud Delattre, VALENTINE GENON-CATALOT, and Adeline Samson. Maximum likelihood estimation for stochastic differential equations with random effects. *Scandinavian Journal of Statistics*, 40(2):322–343, 2013.
13. Maud Delattre, Valentine Genon-Catalot, and Adeline Samson. Mixtures of stochastic differential equations with random effects: application to data clustering. *Journal of Statistical Planning and Inference*, 173:109–124, 2016.
14. Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.
15. Paul Dupuis and Hitoshi Ishii. On lipschitz continuity of the solution mapping to the skorokhod problem, with applications. *Stochastics: An International Journal of Probability and Stochastic Processes*, 35(1):31–62, 1991.
16. Hamid El Maroufy, Souad Ichi, Mohamed El Omari, and Yousri Slaoui. Nonparametric estimation for random effects models driven by fractional brownian motion using hermite polynomials. *Statistical Inference for Stochastic Processes*, 27(2):305–333, 2024.
17. Mohamed El Omari, Hamid El Maroufy, and Christiane Fuchs. Non-parametric estimation for fractional diffusion processes with random effects. *Statistics*, 2019.
18. Mohamed El Omari, Hamid El Maroufy, and Christiane Fuchs. Statistical inference for discretely observed fractional diffusion processes with random effects. *Communications in Mathematics and Statistics*, Apr 2025.

19. Mohammed Er-rachidi, Hamid El Maroufy, and Mohamed El Omari. Moments method of estimation for diffusion processes with random effects driven by fractional brownian motion with real application. *Journal of Statistical Computation and Simulation*, 0(0):1–22, 2026.
20. Souad Ichi and Hamid El Maroufy. Nonparametric estimation for fractional geometric model with random effects, with application to the nasdaq-100 stock prices. *Journal of Statistical Theory and Practice*, 19(4):62, 2025.
21. Nelson T. Jamba, Patr Filipe, Gon Jacinto, and Carlos A. Braumann. Stochastic differential equations mixed model for individual growth with the inclusion of genetic characteristics. *Statistics, Optimization amp; Information Computing*, 12(2):298–309, Dec. 2023.
22. Christine Keribin. Consistent estimation of the order of mixture models. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 49–66, 2000.
23. Takis Konstantopoulos and Si-Jian Lin. Fractional brownian approximations of queueing networks. In *Stochastic networks*, pages 257–273. Springer, 1996.
24. K. Kubilius and D. Melichov. Quadratic variations and estimation of the hurst index of the solution of sde driven by a fractional brownian motion. *Lithuanian Mathematical Journal*, 50(4):401–417, 2010.
25. Chihoon Lee and Jian Song. On drift parameter estimation for reflected fractional ornstein–uhlenbeck processes. *Stochastics*, 88(5):751–778, 2016.
26. Brian G Leroux and Martin L Puterman. Maximum-penalized-likelihood estimation for independent and markov-dependent mixture models. *Biometrics*, pages 545–558, 1992.
27. Robert S. Liptser and Albert N. Shiryaev. *Statistics of Random Processes II: Applications*, volume 2. Springer Science & Business Media, 2001.
28. Geoffrey J McLachlan and Thriyambakam Krishnan. *The EM algorithm and extensions*. John Wiley & Sons, 2008.
29. Yuliya S Mishura. *Stochastic calculus for fractional Brownian motion and related processes*. Springer, 2008.
30. El Omari Mohamed and Hamid El Maroufy. Nonparametric estimation for small fractional diffusion processes with random effects. *Stochastic Analysis and Applications*, 38(6):1084–1101, 2020.
31. Shohei Nakajima and Yasutaka Shimizu. Parameter estimation of stochastic differential equation driven by small fractional noise. *Statistics*, 56(4):919–934, 2022.
32. Ilkka Norros, Esko Valkeila, and Jorma Virtamo. An elementary approach to a girsanov formula and other analytical results on fractional brownian motions. 1999.
33. Umberto Picchini and Susanne Ditlevsen. Practical estimation of high dimensional stochastic differential mixed-effects models. *Computational Statistics & Data Analysis*, 55(3):1426–1444, 2011.
34. BLS Prakasa Rao. *Asymptotic theory of statistical inference*. John Wiley & Sons, Inc., 1987.
35. Josh Reed, Amy Ward, and Dongyuan Zhan. On the generalized drift skorokhod problem in one dimension. *Journal of Applied Probability*, 50(1):16–28, 2013.
36. Josh Reed and Amy R Ward. A diffusion approximation for a generalized jackson network with reneging. In *Proceedings of the 42nd annual Allerton conference on communication, control, and computing*, 2004.
37. Luigi M Ricciardi. Stochastic population theory: diffusion processes. In *Mathematical Ecology: An Introduction*, pages 191–238. Springer, 1986.
38. Anatolıı V Skorohod. Stochastic differential equations for a bounded region. *Theor. Probability Appl*, 6:264–274, 1961.
39. Amy R Ward and Peter W Glynn. A diffusion approximation for a markovian queue with reneging. *Queueing systems*, 43:103–128, 2003.
40. Amy R Ward and Peter W Glynn. Properties of the reflected ornstein–uhlenbeck process. *Queueing Systems*, 44:109–123, 2003.
41. Amy R Ward and Peter W Glynn. A diffusion approximation for a gi/gi/1 queue with balking or reneging. *Queueing Systems*, 50:371–400, 2005.
42. C. F. Jeff Wu. On the convergence properties of the em algorithm. *The Annals of Statistics*, 11(1):95–103, 1983.