



## Maximum likelihood Estimation for Stochastic Differential Equations with two Random Effects in the Diffusion Coefficient

Alsukaini Mohammed Sari<sup>1,2</sup>, AlkreemawiWalaa khazaal<sup>1,2</sup>, Wang xiang-jun<sup>1\*</sup>

<sup>1</sup>School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei, 430074 P.R.China.

<sup>2</sup>Department of Mathematics, College of Science, Basra University, Basra, Iraq  
mohammedsari@yahoo.com<sup>1,2</sup>

### Abstract

We study  $n$  independent stochastic processes  $(X_i(t), t \in [0, T_i], i = 1, \dots, n)$ , defined by a stochastic differential equation with diffusion coefficients depending nonlinearly on a random variables  $\phi_i$  and  $\mu_i$  (the random effects). The distributions of the random effects  $\phi_i$  and  $\mu_i$  depends on unknown parameters which are to be estimated from the continuous observations of the processes  $X_i(t)$ . When the distributions of the random effects  $\phi_i, \mu_i$  are Gaussian and exponential respectively, we obtained an explicit formula for the likelihood function and the asymptotic properties (consistency and asymptotic normality) of the maximum likelihood estimator (MLE) are derived when  $n$  tend to infinity.

**Keywords:** stochastic differential equations, Maximum likelihood estimator; nonlinear random effects; strong consistency; asymptotic normality.



## Council for Innovative Research

Peer Review Research Publishing System

**Journal:** JOURNAL OF ADVANCES IN MATHEMATICS

Vol. 11, No. 10

[www.cirworld.com](http://www.cirworld.com), [editor@cirworld.com](mailto:editor@cirworld.com)



## 1 INTRODUCTION

Stochastic differential equations play an important role in many areas of science fields as physics, engineering, chemistry, neuroscience, biology, finance (Gugushvili and P. Spreij (2012)[11]). Statistical estimation of parameters in the diffusion processes has been studied for a long time; Feigin [10] provided a useful historical overview of the early studies and introduced a general asymptotic theory of maximum likelihood estimation for continuous diffusion processes. In the recent years, the stochastic differential equations with random effects have been the subject of diverse applications such as pharmacokinetic/pharmacodynamics, neuronal modeling (Delattre and Lavelle, 2013[5], Donnet and Samson, 2013[9], Picchini et al. 2010[15]). Maximum likelihood estimator of the parameters of the random effect, is generally not possible, because of the likelihood function is not available in most cases, except in (Ditlevsen and De Gaetano (2005) [5] and [8] which are found a special case of SDE and derived an explicit MLE. Many references proposed approximations for the unknown likelihood function, for general mixed SDEs an approximations of the likelihood have been proposed (Picchini and Ditlevsen, 2011[14]), linearization (Beal and Sheiner (1982)[4]), Laplace's approximation (Wolfinger, (1993)[17]) or approximating the conditional transition density of the diffusion process given the random effects by a Hermit expansion, (Ait-Sahalia (2002)[1]).

Delattre et al. (2012) [7] and alkreemawi et al. (2015) [2] are studied the maximum likelihood estimator for random effects in more generally for fixed  $T$  and  $n$  tending to infinity (for non i.i.d. sample paths, see Maitra et al. (2014) [13]) and they found an explicit expression for likelihood function and exact likelihood estimator by investigate the linear random effect in the drift (multiple and additive case respectively) together with a specific distribution for the random effect. Almost researcher studied the random effect in the drift not in diffusion except Delattre et al. (2014) [6] and alsukaini et al. (2015) [3] who used one random effect in the diffusion coefficient (linearly and nonlinearly respectively) with a specific distributions and focus on (discretely and continuously) observed SDEs respectively.

In the present work we focus on stochastic differential equation with two random effects in diffusion term and without random effect in drift term. We consider  $n$  real valued stochastic processes  $(X_i(t), t \in [0, T_i], i = 1, \dots, n)$ , with dynamics ruled by the following SDEs:

$$dX_i(t) = b(X_i(t))dt + \sigma(X_i(t), \Phi_i, \mu_i)dW_i(t), \quad X_i(0) = x^i, \quad i = 1, \dots, n \quad (1)$$

where  $W_1, \dots, W_n$  are  $n$  independent Wiener processes,  $\Phi_1, \dots, \Phi_n$  and  $\mu_1, \dots, \mu_n$  are *ni. i. d.* random variables taking values in  $(\mathbb{R} \text{ and } \mathbb{R}^+)$  respectively,  $\Phi_1, \dots, \Phi_n, \mu_1, \dots, \mu_n$  and  $W_1, \dots, W_n$  are independent and  $x^i, i = 1, \dots, n$  are known real values. The functions  $b(x)$  and  $\sigma(x)$  are known real valued. Each process  $X_i(t)$  represents an individual, the variables  $\Phi_i$  and  $\mu_i$  represents the random effects of individual  $i$ , the random variables  $\Phi_1, \dots, \Phi_n$  have a common distribution  $g(\varphi, \theta)dv(\varphi)$  on  $\mathbb{R}$  and the random variables  $\mu_1, \dots, \mu_n$  have a common distribution  $h(\mu, \beta)du(\mu)$  on  $\mathbb{R}^+$  where  $\theta$  and  $\beta$  are an unknown parameters belonging to a set  $\Theta \subset \mathbb{R}^p$  where  $v$  and  $u$  are a dominating measures.

Our aim is to estimate  $\psi = (\theta, \beta)$  from the continuous observations  $(X_i(t), t \in [0, T_i], i = 1, \dots, n)$ , we focus on a special case of nonlinear random effect in the diffusion coefficient in the model (1), i.e.  $\sigma(x, \Phi_i, \mu_i) = \frac{1}{\Phi_i + \mu_i} \sigma(x)$ , where  $\sigma$  is a known real function and  $\Phi_i$  is a Gaussian and  $\mu_i$  is an exponential, we find an explicit likelihood formula and the maximum likelihood estimator of  $\psi$ . We use the following sufficient statistics  $U_i$  and  $V_i$  as in [7]:

$$U_i = \int_0^{T_i} \frac{b(X_i(s))}{\sigma^2(X_i(s))} dX_i(s) \quad , \quad V_i = \int_0^{T_i} \frac{b^2(X_i(s))}{\sigma^2(X_i(s))} ds \quad , \quad i = 1, \dots, n.$$

The consistency and asymptotic normality of the maximum likelihood estimator of  $\psi$  are proved without invoking the result in the literature.

The rest of this paper is organized as follows. Section 2 contains the notation and assumptions that we will need throughout the paper. The explicit likelihood function, specific distributions for the random effects and the maximum likelihood estimators of the parameters of random effects are introduced in section 3. In section 4 we study the asymptotic properties of the maximum likelihood estimator when the random effects are Gaussian and exponential distribution respectively.

## 2 Notations and Assumptions

Consider  $n$  real valued stochastic processes  $(X_i(t), t \geq 0), i = 1, \dots, n$  with dynamics ruled by (1). The processes  $W_1, \dots, W_n$  and the random variables  $\Phi_1, \dots, \Phi_n$  and  $\mu_1, \dots, \mu_n$  are defined on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Consider the filtration  $(\mathcal{F}_t, t \geq 0)$  defined by  $\mathcal{F}_t = \sigma(\Phi_i, \mu_i, W_i(s), s \leq t, i = 1, \dots, n)$ . As  $\mathcal{F}_t = \sigma(\Phi_i, \mu_i, W_i(s), s \leq t) \vee \mathcal{F}_t^i$  with  $\mathcal{F}_t^i = \sigma(\Phi_j, \mu_j, W_j(s), s \leq t, j \neq i)$  independent of  $W_i$ , each process  $W_i$  is a  $(\mathcal{F}_t, t \geq 0)$ -Brownian motion. Moreover, the random variables  $\Phi_i, \mu_i$  are  $\mathcal{F}_0$ -measurable. We used  $\square$  referring to the end of the proofs.

We assume that:

**H1** the function  $\sigma$  belongs to  $c^2(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+)$  and for all  $x \in \mathbb{R}, 0 < \sigma_0^2 \leq \sigma^2(x, \varphi, \mu) \leq \sigma_1^2$

From **H1**, the process  $(X_i(t))$  is well define and  $(\Phi_i, \mu_i, X_i(t))$  adapted to filtration  $(\mathcal{F}_t, t \geq 0)$ . The  $n$  processes  $(\Phi_i, \mu_i, X_i(t)), i = 1, \dots, n$  are independent. For all  $\varphi, \mu$  and all  $x^i \in \mathbb{R}$ , the stochastic differential equation:



$$dX_i^{\varphi,\mu}(t) = b(X_i^{\varphi,\mu})dt + \sigma(X_i^{\varphi,\mu}, \varphi, \mu)dW_i(t), \quad X_i^{\varphi,\mu}(0) = x^i \quad (2).$$

Admits a unique strong solution process  $(X_i^{\varphi,\mu}(t), t \geq 0)$  adapted to filtration  $(\mathcal{F}_t, t \geq 0)$ . We infer that the conditional distribution of  $X_i$  given  $\varphi_i = \varphi$  and  $\mu_i = \mu$  identical to the distribution of  $X_i^{\varphi,\mu}$ .

### 3 Likelihood, A specific distributions for the random effects and Maximum likelihood estimators

#### 3.1 Likelihood

We introduce the distribution  $Q_{\varphi,\mu}^{x^i, T_i}$  of  $(X_i^{\varphi,\mu}(t), t \in [0, T_i])$ .

Let  $P_{\psi}^i = g(\varphi, \theta)dv(\varphi) \otimes h(\mu, \beta)du(\mu) \otimes Q_{\varphi}^{x^i, T_i}$  denote the joint distribution of  $(\varphi_i, \mu_i, X_i(t))$  and let  $Q_{\psi}^i$  denote the marginal distribution of  $(X_i(t), t \in [0, T_i])$ . Let us consider the following assumption:

**H2** For  $i = 1, \dots, n$  and for all  $\varphi, \mu, \varphi', \mu'$ ,  $Q_{\varphi,\mu}^{x^i, T_i} \left( \int_0^{T_i} \frac{b^2(X_i^{\varphi,\mu}(t))}{\sigma^2(X_i^{\varphi,\mu}(t), \varphi, \mu)} dt < +\infty \right) = 1$ .

**Proposition 3.1** under **H1-H2** and let  $\varphi \in \mathbb{R}, \mu \in \mathbb{R}^+$ , we have, the distribution  $Q_{\varphi,\mu}^{x^i, T_i}$  are absolutely continuous w.r.t.  $Q^i = Q_{\varphi_0, \mu_0}^{x^i, T_i}$  with density:

$$\frac{dQ_{\varphi,\mu}^{x^i, T_i}}{dQ^i}(X_i) = L_{T_i}(X_i, \varphi, \mu) = \exp \left( \int_0^{T_i} \frac{b(X_i(s))}{\sigma^2(X_i(s), \varphi, \mu)} dX_i(s) - \frac{1}{2} \int_0^{T_i} \frac{b^2(X_i(s))}{\sigma^2(X_i(s), \varphi, \mu)} ds \right) \quad (3).$$

(See Liptser and Shiryaev [12]), which is admits a continuous version  $Q^i$  a.s.

**Proof:** (see the proof of proposition 2) in [7].

By independent of individuals,  $P_{\psi} = \otimes_{i=1}^n P_{\psi}^i$  is the distribution of  $(\varphi_i, \mu_i, X_i(\cdot)), i = 1, \dots, n$  and  $Q_{\psi} = \otimes_{i=1}^n Q_{\psi}^i$  is the distribution of the sample  $(X_i(t), t \in [0, T_i], i = 1, \dots, n)$ .

We can compute the density of  $Q_{\psi}$  w.r.t.  $Q = \otimes_{i=1}^n Q^i$  as follow:

$$\frac{dQ_{\psi}}{dQ}(X_i) = \int_{\mathbb{R}^+} \int_{\mathbb{R}} L_{T_i}(X_i, \varphi, \mu) g(\varphi, \theta) h(\mu, \beta) dv(\varphi) du(\mu) = \gamma_i(X_i, \psi).$$

The distribution  $Q_{\psi}$  admits a density given by:

$$\frac{dQ_{\psi}}{dQ}(X_1, \dots, X_n) = \prod_{i=1}^n \gamma_i(X_i, \psi),$$

And the exact likelihood of whole sample  $(X_i(t), t \in [0, T_i], i = 1, \dots, n)$  is

$$\xi_n(\psi) = \prod_{i=1}^n \gamma_i(X_i, \psi).$$

#### 3.2 A Specific distributions for the random effects

Consider model (1) with nonlinear random effects in the diffusion coefficient  $\sigma(x, \varphi, \mu) = \frac{1}{\varphi + \mu} \sigma(x)$  where  $\varphi \in \mathbb{R}, \mu \in \mathbb{R}^+$  and  $b(\cdot), \sigma(\cdot)$  are known functions. We assume that:

$$\int_0^{T_i} \frac{b^2(X_i(s))}{\sigma^2(X_i(s))} ds < \infty, Q_{\varphi,\mu}^{x^i, T_i} - a. s,$$

for all  $\varphi, \mu$  and for  $i = 1, \dots, n; T_i = T, x^i = x$ , so that  $(X_i(t), t \in [0, T], i = 1, \dots, n)$  are *i. i. d.* We will use the well define statistics as follow:

$$U_i = \int_0^T \frac{b(X_i(s))}{\sigma^2(X_i(s))} dX_i(s), \quad V_i = \int_0^T \frac{b^2(X_i(s))}{\sigma^2(X_i(s))} ds \quad (4).$$

So that the density  $\gamma_i(X_i, \psi)$  is given by:

$$\gamma_i(X_i, \psi) = \int_{\mathbb{R}^+} \int_{\mathbb{R}} \exp \left( (\varphi + \mu)^2 \left( U_i - \frac{1}{2} V_i \right) \right) g(\varphi, \theta) h(\mu, \beta) dv(\varphi) du(\mu) \quad (5).$$



For a general distributions,  $g(\varphi, \theta)dv(\varphi)$  of the random effect  $\varphi_i$  and  $h(\mu, \beta)dv(\varphi)$  of the random effect  $\mu$ , it is not possible find an explicit expression for  $\gamma_i(X_i, \psi)$  above, therefor we propose a specific distributions (Gaussian  $(\lambda, \omega^2)$  for the random effects  $\varphi$  and exponential  $(\beta)$  for the random effect  $\mu$ ) which will give an explicit likelihood and then find the maximum likelihood estimators of the unknown parameters. In the next proposition an evident expression for  $\gamma_i(X_i, \psi)$  is obtained when the above distributions of the random effects is with unknown parameter  $\psi = (\lambda, \omega^2, \beta) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+$ . The true value is denoted by  $\psi_0 = (\lambda_0, \omega_0^2, \beta_0)$ .

**Proposition 3.2** suppose that  $g(\varphi, \theta)dv(\varphi) = \mathcal{N}(\lambda, \omega^2)$ , and  $h(\mu, \beta)dv(\mu) = \exp(\beta)$  then:

$$\gamma_i(X_i, \psi) = \frac{\sqrt{\pi}\beta}{\sqrt{M_i}} \exp\left(-\frac{1}{4} \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{M_i(1 - 2\omega^2 M_i)} + \frac{(\lambda^2 - \lambda(1 - 2M_i\omega^2))}{2\omega^2(1 - 2M_i\omega^2)}\right),$$

where  $M_i = U_i - \frac{1}{2}V_i$ .

**Proof:** from (5) we compute the joint density of  $(\varphi_i, \mu_i, X_i)$ :

$$\exp\left((\varphi + \mu)^2 \left(U_i - \frac{1}{2}V_i\right)\right) \times \frac{1}{\sqrt{2\pi\omega^2}} \exp\left(-\frac{1}{2\omega^2}(\varphi - \lambda)^2\right) \times \beta \exp(-\beta\mu).$$

Let  $M_i = U_i - \frac{1}{2}V_i$ , then the exponent become:

$$D_i = \varphi^2 M_i - \frac{1}{2\omega^2}(\varphi - \lambda)^2 + 2\varphi\mu M_i + \mu^2 M_i - \beta\mu. \tag{6}$$

We will compute the first part  $(\varphi^2 M_i - \frac{1}{2\omega^2}(\varphi - \lambda)^2 + 2\varphi\mu M_i)$  of the exponent as follow:

$$\begin{aligned} \varphi^2 M_i - \frac{1}{2\omega^2}(\varphi - \lambda)^2 + 2\varphi\mu M_i &= \left(M_i - \frac{1}{2\omega^2}\right)\varphi^2 + \left(\frac{\lambda}{\omega^2} + 2\mu M_i\right)\varphi - \frac{\lambda^2}{2\omega^2} \\ &= \frac{-1}{2}\left(\frac{1}{\omega^2} - 2M_i\right)\left(\varphi^2 - 2\frac{\lambda + 2\omega^2 M_i \mu}{1 - 2\omega^2 M_i}\varphi\right) - \frac{\lambda^2}{2\omega^2} \\ &= \frac{-1}{2}\left(\frac{1}{\omega^2} - 2M_i\right)\left[\left(\varphi - \frac{\lambda + 2\omega^2 M_i \mu}{1 - 2\omega^2 M_i}\right)^2 - \left(\frac{\lambda + 2\omega^2 M_i \mu}{1 - 2\omega^2 M_i}\right)^2\right] - \frac{\lambda^2}{2\omega^2} \\ &= \frac{-1}{2}\left(\frac{1 - 2\omega^2 M_i}{\omega^2}\right)\left(\varphi - \frac{\lambda + 2\omega^2 M_i \mu}{1 - 2\omega^2 M_i}\right)^2 + \frac{(\lambda + 2\omega^2 M_i \mu)^2}{2\omega^2(1 - 2\omega^2 M_i)} - \frac{\lambda^2}{2\omega^2}. \end{aligned}$$

Now, by splitting the result into two parts that are independent and dependent on the random effect  $\varphi$  respectively, the integral of the dependent part is the integral of a Gaussian density.

Then the first integral in (5) with respect to  $\varphi$  yields the following result :

$$\frac{1}{\sqrt{1 - 2\omega^2 M_i}} \exp\left(\frac{(\lambda + 2\omega^2 M_i \mu)^2}{2\omega^2(1 - 2\omega^2 M_i)} - \frac{\lambda^2}{2\omega^2}\right).$$

By substituting in (5), the second part of the exponent is:

$$\begin{aligned} E_i &= \frac{(\lambda + 2\omega^2 M_i \mu)^2}{2\omega^2(1 - 2\omega^2 M_i)} + \mu^2 M_i - \beta\mu - \frac{\lambda^2}{2\omega^2} \\ &= \left(\frac{2\omega^2 M_i^2}{1 - 2\omega^2 M_i} + M_i\right)\mu^2 - \left(\beta - \frac{2\lambda M_i}{1 - 2\omega^2 M_i}\right)\mu + \frac{\lambda^2 - \lambda(1 - 2\omega^2 M_i)}{2\omega^2(1 - 2\omega^2 M_i)} \\ &= \frac{M_i}{1 - 2\omega^2 M_i}\left(\mu^2 - \left(\frac{\beta(1 - 2\omega^2 M_i) - 2\lambda M_i}{M_i}\right)\mu\right) + \frac{\lambda^2 - \lambda(1 - 2\omega^2 M_i)}{2\omega^2(1 - 2\omega^2 M_i)} \\ &= \frac{M_i}{1 - 2\omega^2 M_i}\left[\left(\mu - \frac{1}{2}\frac{\beta(1 - 2\omega^2 M_i) - 2\lambda M_i}{M_i}\right)^2 - \left(\frac{1}{2}\frac{\beta(1 - 2\omega^2 M_i) - 2\lambda M_i}{M_i}\right)^2\right] + \frac{\lambda^2 - \lambda(1 - 2\omega^2 M_i)}{2\omega^2(1 - 2\omega^2 M_i)} \\ &= -\frac{1}{2}\frac{M_i}{\omega^2 M_i - \frac{1}{2}}\left(\mu - \frac{1}{2}\frac{\beta(1 - 2\omega^2 M_i) - 2\lambda M_i}{M_i}\right)^2 - \frac{1}{4}\frac{(\beta(1 - 2\omega^2 M_i) - 2\lambda M_i)^2}{M_i(1 - 2\omega^2 M_i)} + \frac{\lambda^2 - \lambda(1 - 2\omega^2 M_i)}{2\omega^2(1 - 2\omega^2 M_i)}. \end{aligned}$$

Now, by rearrange the second integral we see that the first part is normal depend on the random effect  $\mu$  with mean is

$$\delta_i = \frac{1}{2}\frac{\beta(1 - 2\omega^2 M_i) - 2\lambda M_i}{M_i},$$



And variance,

$$\rho_i^2 = \frac{\omega^2 M_i - \frac{1}{2}}{M_i},$$

Then, the conditional distribution of  $(\phi_i, \mu_i)$  given  $X_i$  is  $\mathcal{N}(\delta_i, \rho_i^2)$ .

And hence,

$$\gamma_i(X_i, \psi) = \frac{\sqrt{\pi}\beta}{\sqrt{M_i}} \exp\left(-\frac{1}{4} \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{M_i(1 - 2\omega^2 M_i)} + \frac{\lambda^2 - \lambda(1 - 2M_i\omega^2)}{2\omega^2(1 - 2M_i\omega^2)}\right). \square$$

### 3.3 Maximum likelihood estimator for $\psi = (\lambda, \omega^2, \beta)$ :

Since we find the density  $\gamma_i(X_i, \psi)$ , a natural approach to estimate  $\psi$  is the maximum likelihood estimation, so in order to find MLE  $\hat{\psi}$ , the likelihood function is written as:

$$\xi_n(\psi) = \prod_{i=1}^n \frac{\sqrt{\pi}\beta}{\sqrt{M_i}} \exp\left(-\frac{1}{4} \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{M_i(1 - 2\omega^2 M_i)} + \frac{(\lambda^2 - \lambda(1 - 2M_i\omega^2))}{2\omega^2(1 - 2M_i\omega^2)}\right).$$

And hence, the logarithm of likelihood function is,

$$\begin{aligned} \mathcal{L}_n(\psi) &= \log \prod_{i=1}^n \xi_n(\psi) & (7) \\ &= \log \pi^{\frac{1}{2}} \beta^n + \frac{1}{2} \log(M_i) - \sum_{i=1}^n \left[ \frac{1}{4} \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{M_i(1 - 2\omega^2 M_i)} - \frac{(\lambda^2 - \lambda(1 - 2M_i\omega^2))}{2\omega^2(1 - 2M_i\omega^2)} \right]. \end{aligned}$$

We will study the following score function

$G_n(\psi) = \left( \frac{\partial}{\partial \mu} \mathcal{L}_n(\psi), \frac{\partial}{\partial \beta} \mathcal{L}_n(\psi), \frac{\partial}{\partial \omega^2} \mathcal{L}_n(\psi) \right)'$ , where  $x'$  denotes the transpose of  $x$ , such that:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \mathcal{L}_n(\psi) &= \sum_{i=1}^n \left[ \frac{\beta(1 - 2M_i\omega^2) - 2\lambda M_i}{1 - 2\omega^2 M_i} + \frac{2\lambda - (1 - 2M_i\omega^2)}{2\omega^2(1 - 2M_i\omega^2)} \right], \\ \frac{\partial}{\partial \omega^2} \mathcal{L}_n(\psi) &= \sum_{i=1}^n \left[ \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)\beta}{1 - 2M_i\omega^2} - \frac{1}{2} \frac{(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{(1 - 2\omega^2 M_i)^2} - \frac{\lambda^2(2 - 8M_i\omega^2)}{(2\omega^2(1 - 2M_i\omega^2))^2} + \frac{\lambda}{2(\omega^2)^2} \right], \\ \frac{\partial}{\partial \beta} \mathcal{L}_n(\psi) &= \sum_{i=1}^n \left[ \frac{1}{\beta} - \frac{\frac{1}{2}\beta(1 - 2M_i\omega^2)}{M_i} + \lambda \right]. \end{aligned}$$

When  $\omega_0^2, \beta_0$  are known, the explicit estimator  $\lambda_0$  is:

$$\hat{\lambda}_n = \frac{\sum_{i=1}^n \frac{1 - 2M_i\omega_0^2\beta_0}{2M_i}}{\sum_{i=1}^n \frac{1 - 2M_i\omega_0^2}{1 - 2M_i\omega_0^2}},$$

and when  $\omega_0^2, \lambda_0$  are known, the explicit estimator of  $\beta_0$  is:

$$\hat{\beta}_n = \begin{cases} \frac{-2n\lambda + \sqrt{4n^2\lambda_0^2 + 8nK_i}}{2K_i}, \\ \frac{-2n\lambda - \sqrt{4n^2\lambda_0^2 + 8nK_i}}{2K_i}, \end{cases}$$

where  $K_i = \sum_{i=1}^n \frac{1 - 2M_i\omega_0^2}{M_i}$ .

If all the parameters are unknown, the MLEs of  $\psi_0 = (\lambda_0, \omega_0^2, \beta_0)$  are given by the system:

$$\hat{\lambda}_N = \left( \sum_{i=1}^n \frac{1 - 2M_i\hat{\omega}_N^2\hat{\beta}_N}{2M_i} \right) \left( \sum_{i=1}^n \frac{1 - 2M_i\hat{\omega}_N^2}{1 - 2M_i\hat{\omega}_N^2} \right)^{-1}$$



$$\hat{\beta}_n = \begin{cases} \frac{-2n\lambda + \sqrt{4n^2\hat{\lambda}_n^2 + 8nK_i}}{2K_i}, \\ \frac{-2n\lambda - \sqrt{4n^2\hat{\lambda}_n^2 + 8nK_i}}{2K_i} \end{cases}$$

Such that  $\hat{\beta}_n$  is a maximum likelihood estimator defined as any solution of

$$\mathcal{L}_n(\hat{\beta}_n) = \sup_{\psi \in \Theta} \mathcal{L}_N(\psi).$$

$$\sum_{i=1}^n \frac{1}{2} \frac{(\hat{\beta}_n(1 - 2M_i\hat{\omega}_n^2) - 2\hat{\lambda}_n M_i)^2}{\hat{\lambda}_n(1 - 2\hat{\omega}_n^2 M_i)^2} - \frac{(\hat{\beta}_n(1 - 2M_i\hat{\omega}_n^2) - 2\hat{\lambda}_n M_i)\hat{\beta}_n}{\hat{\lambda}_n(1 - 2M_i\hat{\omega}_n^2)} - \frac{\hat{\lambda}_n^2(2 - 8M_i\hat{\omega}_n^2)}{\hat{\lambda}_n(2\hat{\omega}_n^2(1 - 2M_i\hat{\omega}_n^2))^2} = \sum_{i=1}^n \frac{1}{2(\hat{\omega}_n^2)^2}.$$

The second derivatives of  $\mathcal{L}_N(\psi)$  with respect to the parameters is as follow:

$$\frac{\partial^2}{\partial \lambda^2} \mathcal{L}_n(\psi) = \sum_{i=1}^n \left[ \frac{-2M_i}{1 - 2\omega^2 M_i} + \frac{1}{\omega^2(1 - 2M_i\omega^2)} \right] \quad (8),$$

$$\begin{aligned} \frac{\partial^2}{\partial \omega^2 \partial \omega^2} \mathcal{L}_n(\psi) &= \sum_{i=1}^n \frac{-4\lambda M_i^2 \beta}{(1 - 2\omega^2 M_i)^2} + \frac{2M_i \beta (\beta(1 - 2M_i\omega^2) - 2\lambda M_i)}{(1 - 2\omega^2 M_i)^2} - \frac{\lambda}{(\omega^2)^3} \\ &- \frac{2M_i(\beta(1 - 2M_i\omega^2) - 2\lambda M_i)^2}{(1 - 2\omega^2 M_i)^3} + \frac{8M_i^2 \lambda^2}{(2\omega^2(1 - 2\omega^2 M_i))^2} - \frac{2\lambda^2(2 - 8\omega^2)^2}{(2\omega^2(1 - 2\omega^2 M_i))^3} \quad (9), \end{aligned}$$

$$\frac{\partial^2}{\partial \beta^2} \mathcal{L}_n(\psi) = \sum_{i=1}^n \left[ -\frac{(1 - 2M_i\omega^2)}{M_i} \right] \quad (10),$$

$$\begin{aligned} \frac{\partial^2}{\partial \lambda \partial \omega^2} \mathcal{L}_n(\psi) &= \sum_{i=1}^n \left[ \frac{-4\lambda M_i^2}{(1 - 2\omega^2 M_i)^2} - \frac{\lambda(1 - 4M_i\omega^2)}{(\omega^2(1 - 2M_i\omega^2))^2} + \frac{1}{2(\omega^2)^2} \right] \\ &= \frac{\partial^2}{\partial \omega^2 \partial \lambda} \mathcal{L}_N(\psi) \quad (11), \end{aligned}$$

$$\frac{\partial^2}{\partial \lambda \partial \beta} \mathcal{L}_n(\psi) = n = \frac{\partial^2}{\partial \beta \partial \lambda} \mathcal{L}_n(\psi) \quad , \quad \frac{\partial^2}{\partial \omega^2 \partial \beta} \mathcal{L}_n(\psi) = n\beta = \frac{\partial^2}{\partial \beta \partial \omega^2} \mathcal{L}_n(\psi) \quad (12).$$

And the information matrix ,

$$I(\psi) = \begin{pmatrix} E_{\psi} \left( \frac{\partial^2}{\partial \lambda^2} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \lambda \partial \omega^2} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \lambda \partial \beta} \mathcal{L}_n(\psi) \right) \\ E_{\psi} \left( \frac{\partial^2}{\partial \omega^2 \partial \lambda} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \omega^2 \partial \omega^2} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \omega^2 \partial \beta} \mathcal{L}_n(\psi) \right) \\ E_{\psi} \left( \frac{\partial^2}{\partial \beta \partial \lambda} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \beta \partial \omega^2} \mathcal{L}_n(\psi) \right) & E_{\psi} \left( \frac{\partial^2}{\partial \beta^2} \mathcal{L}_n(\psi) \right) \end{pmatrix} \quad (13),$$

is the covariance matrix of the vector

$$\begin{pmatrix} \frac{\partial}{\partial \lambda} \mathcal{L}_n(\psi) \\ \frac{\partial}{\partial \omega^2} \mathcal{L}_n(\psi) \\ \frac{\partial}{\partial \beta} \mathcal{L}_n(\psi) \end{pmatrix}.$$

The following lemma is important to investigate the consistency and asymptotic normality of the estimators.

**Lemma 3.1** If  $H_i = -2M_i$ , for all  $\psi = (\lambda, \omega^2, \beta) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+$  and all  $m \in \mathbb{R}$ ,

$$E_{\psi} \left( \exp \left( m \frac{M_i}{1 + \omega^2 H_i} \right) \right) < +\infty.$$

The proof of the above lemma is similar to proof of lemma 1 in [7] and omitted.

**Remark 3.1** from lemma (3.1) and proof of proposition (7) in [7], the second derivatives of log-likelihood function  $\mathcal{L}_n(\psi)$  has finite expectation.



We need the following additional assumptions to prove the asymptotic properties:

- H3 The parameter set  $\Theta$  is a compact subset of  $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ .
- H4 The true value  $\psi_0$  belongs to  $\Theta^o$ .
- H5 The matrix  $I(\psi_0)$  is invertible.

## 4 Asymptotic properties of maximum likelihood estimator

### 4.1 Consistency of MLE

We consider the theorem 7.49 and 7.54 of schervish (1995) [16] and verify the regularity conditions in this theorem for our purpose with suppose that  $\Omega$  is compact. The consistency here is strong consistency that is mean any solution of the likelihood is consistent.

#### Theorem 1 [16]:

Let  $\{x_n\}_{n=1}^\infty$  be conditionally *i. i. d* given  $\theta$  with density  $f_1(x|\theta)$  with respect to a measure  $\nu$  on a space  $(\mathcal{X}^1, \mathcal{B}^1)$ . Fix  $\theta_0 \in \Omega$ , and define, for each  $M \subseteq \Omega$  and  $x \in \mathcal{X}^1$ ,

$$Z(M, x) = \inf_{\alpha \in M} \log \frac{f_1(x|\theta_0)}{f_1(x|\alpha)}.$$

Assume that for each  $\theta \neq \theta_0$ , there is an open set  $N_\theta$  such that  $\theta \in N_\theta$  and that  $E_{\theta_0} Z(N_\theta, X_i) > -\infty$ .

Also assume that  $f_1(x|\cdot)$  is continuous at  $\theta$  for every  $\theta$ , a.s.  $[P_{\theta_0}]$ .

Then if  $\hat{\theta}_n$  is the MLE of  $\theta$  corresponding to  $n$  observations. It holds that  $\lim_{n \rightarrow \infty} \hat{\theta}_n = \theta_0$  a.s.  $[P_{\theta_0}]$ .

We note that for any  $x$ ,  $f_1(x|\psi) = \lambda_1(x, \psi) = \lambda(x, \psi)$ . Which is clearly continuous in  $\psi$ . In our case we compute for every  $\psi \neq \psi_0$ :

$$\begin{aligned} \log \frac{f_1(x|\psi_0)}{f_1(x|\psi)} &= \log \frac{\beta_0}{\beta} - \frac{1}{4} \frac{(\beta_0(1-2\omega_0^2 M_i) - 2\lambda_0 M_i)^2}{M_i(1-2\omega_0^2 M_i)} + \frac{\lambda_0^2 - \lambda_0(1-2\omega_0^2 M_i)}{2\omega_0^2(1-2\omega_0^2 M_i)} \\ &\quad + \frac{1}{4} \frac{(\beta(1-2\omega^2 M_i) - 2\lambda M_i)^2}{M_i(1-2\omega^2 M_i)} + \frac{\lambda^2 - \lambda(1-2\omega^2 M_i)}{2\omega^2(1-2\omega^2 M_i)} \\ &= \log \frac{\beta_0}{\beta} - \frac{1}{4} \left( \frac{(\beta_0^2(1-2\omega_0^2 M_i)^2 - 4\beta_0 \lambda_0 M_i(1-2\omega_0^2 M_i) + 4\lambda_0^2 M_i^2)^2}{M_i(1-2\omega_0^2 M_i)} + \frac{\lambda_0^2}{2\omega_0^2(1-2\omega_0^2 M_i)} - \frac{\lambda_0}{2\omega_0^2} \right) + \\ &\quad \frac{1}{4} \left( \frac{(\beta^2(1-2\omega^2 M_i)^2 - 4\beta \lambda M_i(1-2\omega^2 M_i) + 4\lambda^2 M_i^2)^2}{M_i(1-2\omega^2 M_i)} + \frac{\lambda^2}{2\omega^2(1-2\omega^2 M_i)} - \frac{\lambda}{2\omega^2} \right) \\ &= \log \frac{\beta_0}{\beta} - \frac{1}{4} \frac{\beta_0^2(1-2\omega_0^2 M_i)}{M_i} + \beta_0 \lambda_0 - \frac{\lambda_0^2 M_i}{1-2\omega_0^2 M_i} + \frac{\lambda_0^2}{2\omega_0^2(1-2\omega_0^2 M_i)} - \frac{\lambda_0}{2\omega_0^2} + \frac{1}{4} \frac{\beta^2(1-2\omega^2 M_i)}{M_i} - \beta \lambda + \frac{\lambda^2 M_i}{1-2\omega^2 M_i} \\ &\quad + \frac{\lambda^2}{2\omega^2(1-2\omega^2 M_i)} - \frac{\lambda}{2\omega^2} \\ &= \log \frac{\beta_0}{\beta} - \frac{1}{4} \left( \frac{\beta_0^2(1-2\omega_0^2 M_i)}{M_i} - \frac{\beta^2(1-2\omega^2 M_i)}{M_i} \right) - \left( \frac{\lambda_0^2 M_i}{1-2\omega_0^2 M_i} - \frac{\lambda^2 M_i}{1-2\omega^2 M_i} \right) \\ &\quad + \left( \frac{\lambda_0^2}{2\omega_0^2(1-2\omega_0^2 M_i)} + \frac{\lambda^2}{2\omega^2(1-2\omega^2 M_i)} \right) - (\beta \lambda - \beta_0 \lambda_0) - \left( \frac{\lambda_0}{2\omega_0^2} - \frac{\lambda}{2\omega^2} \right). \end{aligned}$$

We note that  $E_{\psi_0} \log \frac{\beta_0}{\beta}$ ,  $E_{\psi_0}(\beta \lambda - \beta_0 \lambda_0)$  and  $E_{\psi_0} \left( \frac{\lambda_0}{2\omega_0^2} - \frac{\lambda}{2\omega^2} \right)$  are finite and by using lemma(3.1),  $E_{\psi_0} \left( \frac{\beta_0^2(1-2\omega_0^2 M_i)}{M_i} - \frac{\beta^2(1-2\omega^2 M_i)}{M_i} \right)$ ,  $E_{\psi_0} \left( \frac{\lambda_0^2 M_i}{1-2\omega_0^2 M_i} - \frac{\lambda^2 M_i}{1-2\omega^2 M_i} \right)$  and  $E_{\psi_0} \left( \frac{\lambda_0^2}{2\omega_0^2(1-2\omega_0^2 M_i)} + \frac{\lambda^2}{2\omega^2(1-2\omega^2 M_i)} \right)$  are also finite, and by assume that  $N_\psi = (\underline{\lambda}, \bar{\lambda}) \times (\underline{\omega^2}, \bar{\omega^2}) \times (\underline{\beta}, \bar{\beta})$ , follows that  $E_{\psi_0} Z(N_\psi, X_i) > -\infty$ , hence  $\lim_{n \rightarrow \infty} \hat{\psi}_N = \psi_0$  a.s.  $[P_{\psi_0}]$ . □

### 4.2 Asymptotic normality of MLE

To corroborate asymptotic normality of MLE  $\hat{\psi}$ , we investigate the conditions in the next theorem provided in ([16], Theorem 7.63):

**Theorem 2 [16]:** Let  $\Omega$  be a subset of  $\mathbb{R}^d$  and let  $\{x_n\}_{n=1}^\infty$  be conditionally *i. i. d* given  $\theta$  with density  $f_1(\cdot|\theta)$ . Let  $\hat{\theta}_N$  be an MLE. Assume that  $\hat{\theta}_n \xrightarrow{P} \theta_0$  under  $P_\theta$  for all  $\theta$ . Assume that  $f_1(x|\theta)$  has continuous second partial



derivatives with respect to  $\theta$  and that differentiation can be passed under the integral sign. Assume that there exist  $H_r(x, \theta)$  such that, for each  $\theta_0 \in \text{int}(\Omega)$  and each  $k, j$ ,

$$\sup_{\|\theta - \theta_0\| \leq r} \left| \frac{\partial^2}{\partial \theta_k \partial \theta_j} \log f_{X_1|\theta}(x|\theta_0) - \frac{\partial^2}{\partial \theta_k \partial \theta_j} \log f_{X_1|\theta}(x|\theta) \right| \leq H_r(x, \theta_0), \quad (14)$$

With

$$\lim_{r \rightarrow 0} E_{\theta_0} H_r(x, \theta_0) = 0. \quad (15)$$

Assume that the fisher information matrix  $I(\theta)$  is finite and non-singular. Then under  $P_{\theta_0}$ ,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} \mathcal{N}(\mathbf{0}, I^{-1}(\theta_0)).$$

The maximum likelihood estimator  $\hat{\psi}$  is almost sure consistency (see section 4.1 above) that is mean  $\hat{\psi}_n \xrightarrow{P} \psi_0$  under  $P_\psi$ ,  $\forall \psi$ . The differentiation can be passed under the integral sign (see [7], proof of proposition 5) and from **H4** we find  $\psi_0 \in \text{int}(\Omega)$ . From (8),(9),(10),(11) and (12), we deduce that  $\frac{\partial^2}{\partial \theta_k \partial \theta_j} \log f_{X_1|\theta}(x|\theta)$  is differentiable in  $\psi = (\lambda, \omega^2, \beta)$ . From remark (3.1), the derivatives has finite expectation, Hence (14) and (15) holds, we obtain that the information matrix  $I(\theta)$  is finite and from **H5**,  $I(\theta)$  is invertible, hence  $\hat{\psi}$  is asymptotically normal.  $\square$

## References

- [1] Ait-Sahalia Y. (2002). Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach, *Econometrica* **70**, no.1, 223-262.
- [2] Alkreemawi W.K, Alsukaini M. S. and Wang X.J. (2015). On Parameters Estimation in Stochastic Differential Equations with Additive Random Effects, *journal of advances in mathematics*, **11**, no.3 5018-5028.
- [3] Alsukaini M. S., Alkreemawi W.K., and Wang, X.J. (2015). Asymptotic Properties of MLE in Stochastic Differential Equations with Random Effects in the Diffusion Coefficient. *International Journal of Contemporary Mathematical Sciences*, **10**, no. 6, 275 - 286.
- [4] Beal S. and Shiner L. (1982). Estimating population kinetics, *Critical Reviews in Biomedical Engineering*, **8** 195 - 222.
- [5] Delattre M. and Lavielle M. (2013). Coupling the SAEM algorithm and the extended Kalman filter for maximum likelihood estimation in mixed-effects diffusion models, *Statistics and Its Interface*, **6** 519 - 532. <http://dx.doi.org/10.4310/sii.2013.v6.n4.a10>.
- [6] M. Delattre, V. Genon-Catalot, and A. Samson, Estimation of population parameters in stochastic differential equations with random effects in the diffusion coefficient, *Preprint MAP*, **5** (2014), 2014 - 07.
- [7] Delattre M., Genon-Catalot V. and Samson A. (2012). Maximum likelihood estimation for stochastic differential equations with random effects, *Scandinavian Journal of Statistics*, **40** 322 - 343. <http://dx.doi.org/10.1111/j.1467-9469.2012.00813.x>.
- [8] Ditlevsen S. and De Gaetano A. (2005), Mixed effects in stochastic differential equation models, *REVSTAT Statistical Journal*, **3** 137 - 153.
- [9] Donnet S. and Samson A. (2013). A review on estimation of stochastic differential equations for pharmacokinetic-pharmacodynamics models, *Advanced Drug Delivery Reviews*, **65** 929-939. <http://dx.doi.org/10.1016/j.addr.2013.03.005>.
- [10] Feigin P.D. (1976). Maximum likelihood estimation for continuous-time stochastic processes, *Advances in Applied Probability* 712-736.
- [11] Gugushvili S. and Spreij P. (2012). Parametric inference for stochastic differential equations: a smooth and match approach, *ALEA, Lat. Am. J. Probab. Math. Stat.* **9** (2), 609–635.
- [12] Liptser R. S. and Shiryaev A. N. (2001). *Statistics of Random Processes I. General Theory*, 2nd edition. Springer-Verlag, Berlin, Heidelberg,
- [13] Maitra T. and Bhattacharya S. (2014). On asymptotic related to classical inference in stochastic differential equations with random effects, *ArXiv: 1407.3968v1*, 1 - 12.
- [14] Picchini U. and Ditlevsen S. (2011). Practical estimation of high dimensional stochastic differential mixed-effects models, *Computational Statistics & Data Analysis*, **55** 1426 - 1444. <http://dx.doi.org/10.1016/j.csda.2010.10.003>.
- [15] Picchini U., De Gaetano A. and Ditlevsen S. (2010). Stochastic differential mixed-effects models, *Scand. J. Statist.*, **37**, 67 - 90.
- [16] M. J. Schervish. (1995). *Theory of Statistics*, Springer-Verlag, New York. <http://dx.doi.org/10.1007/978-1-4612-4250-5>.
- [17] Wolfinger R. (1993) Laplace's approximation for nonlinear mixed models, *Biometrika*, **80** 791 - 795. ISSN 0006-3444. <http://dx.doi.org/10.1093/biomet/80.4.791>.