,

# Parametric inference for discrete observations of diffusion processes with mixed effects


**Catherine Larédo (1)**

Joint work with Maud Delattre (2) & Valentine Genon-Catalot (3)

(1) Laboratoire MaIAGE, I.N.R.A. and LPMA, Université Denis Diderot, CNRS-UMR 7599,
(2) AgroParisTech, France ,
(3) UMR CNRS 8145, Laboratoire MAP5, Université Paris Descartes, Sorbonne Paris Cité


June 8, 2016

**DYNSTOCH 2016, June 8-10**

# General set-up

- Statistical inference for repeated measurements over time (time series data) on several experimental units.
- Among classical examples: Pharmacokinetics (PK)/ pharmacodynamics (PD) experiments:

**Two sources of stochasticity**

(1) Intrinsic randomness of individual dynamics : e.g. system noise.
(2) Population effects: Variation between individuals or experimental units (longitudinal data).

# Diffusion process on $\mathbb{R}$ with random effects

$$\begin{cases} dX(t) & = & b(X(t), \Phi)dt + \sigma(X(t), \Psi)dW(t), \\ X(0) & = & x, \\ x \in \mathbb{R}\ ,\ t \in [0, T]. \end{cases} \tag{1}$$

## $(\Phi, \Psi)$ deterministic unknown $\Rightarrow$ Classical inference

(1) **Continuous observation on [0,T]**
(e.g. Kutoyants, Lipster & Shiryaev.
$\star$ $\Psi$ identified on one sample path $\Rightarrow$ Assumption $\psi$ known.
$\star$ Inference for $\phi$: possible if $T \to \infty$ .
(2) **Discrete observations on [0,T]**: sampling $\Delta$ ($n\Delta = T$, $n$ obs.)
$\star$ $T$ fixed, $n \to \infty$: estimation of $\Psi$ (Genon-Catalot & Jacod)
$\star$ $T$ and $n \to \infty$: estimation of $(\Phi, \Psi)$ (e.g. Kessler ).

## $\Phi, \Psi$ random variables independent of $W$

Aim: estimation of unknown parameters in the distribution of $(\Phi, \Psi)$

# Repeated observations of discretized processes

**Model: $N$ *i.i.d.* processes $(X_i(t))$**

$$\begin{cases} dX_i(t) &= b(X_i(t), \Phi_i)dt + \sigma(X_i(t), \Psi_i)dW_i(t), \\ X_i(0) &= x, \\ x \in \mathbb{R}\ ,\ t \in [0, T]\ ,\ i = 1, \ldots, N, \end{cases}$$

- $(\Phi_i, \Psi_i), i = 1, \ldots, N$: *i.i.d.* random variables.
- $W_i, i = 1, \ldots, N$: : independent Wiener processes.
- $((\Phi_i, \Psi_i), i = 1, \ldots, N)$ and $W_i, i = 1, \ldots, N$ independent.

**Estimation of unknown parameters in the distribution of $(\Phi, \Psi)$**

- Fixed time interval $[0, T]$ with $T = n\Delta$, and $t_j = jT/n, j = 1, \ldots, n$.
- **Observations**: $(\ X_{i,n} := (X_i(t_j), j = 1, \ldots, n), i = 1, \ldots, N\ )$.

## General References

- Nie & Yang (2005), Nie (2006, 2007). Theoretical likelihood study. Rely on many abstract assumptions impossible to check in practice.
- Donnet, S. Samson, A. (2008). Review for mixed effects SDEs.
- Picchini, De Gaetano & Ditlevsen (2008, 2010); Picchini & Ditlevsen (2011) (approximations of the likelihood, no theoretical results)
- Delattre, Genon-Catalot & Samson (2012, 2015, 2016); Genon-Catalot & Larédo (2016); Delattre, Genon-Catalot & Larédo (2016,a,b) (parametric, likelihood methods).

## References on applications based on data sets of PK/PD dynamics

- Overgaard,Jonsson, Tornøe & Madsen (2005)
- Berglund, Sunnake, Adiels, Jirstrand & Wennberg (2011)
- Leander, J., Almquist, J., Ahlstrom, C., Gabrielsson, J. & Jirstrand, M. (2015) (concrete models and real data + many references therein)

# Likelihood for SDE with mixed effects

Parametric distribution for the $(\Phi_i, \Psi_i)$: $\nu_\vartheta(d\varphi, d\psi))$

1. Conditional likelihood of $X_{i,n}$ given $(\Phi_i = \varphi, \Psi_i = \psi)$ : $L_n(X_{i,n}, \varphi, \psi)$
   $L_n(X_{i,n}, \varphi, \psi)$= Likelihood based on $X_i^{\varphi,\psi}(t)$
   $d\, X_i^{\varphi,\psi}(t) = b(X_i^{\varphi,\psi}(t), \varphi)dt + \sigma(X_i^{\varphi,\psi}(t), \psi)\, dW_i(t),\ X_i^{\varphi,\psi}(0) = x.$
2. Integrate with respect to $\nu_\vartheta(d\varphi, d\psi))$ this conditional likelihood:

$$L_n(X_{i,n}, \vartheta) = \int L_n(X_{i,n}, \varphi, \psi)\nu_\vartheta(d\varphi, d\psi) \quad : \text{likelihood of } X_{i,n}, \quad (2)$$

3. Exact likelihood for ( $X_{i,n}, i = 1, \ldots N$ ): $L_N(\vartheta) = \prod_{i=1}^N L_n(X_{i,n}, \vartheta)$.

## Two main difficulties

⋆ Discrete observations: untractable $L_n(X_{i,n}, \varphi, \psi) \Rightarrow$ Approximations.
⋆ Integration w.r.t. $\nu_\vartheta(d\varphi, d\psi))$: no closed form in general
$\Rightarrow$ Choice of specific models $b(.,.), \sigma(.,.)$ and distributions $\nu_\vartheta(.,.)$.

# Model under study: Linear mixed effects

$N$ stochastic processes $(X_i(t); t \geq 0), i = 1, \ldots N$ on $\mathbb{R}$.

$$\begin{cases} dX_i(t) & = & \Phi_i' \, b(X_i(t)) \, dt +, \Psi_i \, \sigma(X_i(t)) \, dW_i(t), \\ X_i(0) & = & x, \\ x \in \mathbb{R} \, , \; t \in [0, T]. \end{cases} \quad (3)$$

- $(W_1, \ldots, W_N)$: $N$ independent Wiener processes.
- $((\Phi_i, \Psi_i), i = 1, \ldots, N)$ $N$ *i.i.d.* r.v. on $\mathbb{R}^d \times (0, +\infty)$, indep. of $(W_i), i = 1, \ldots N$.
- $b(.) = (b_1(.), \ldots b_d(.))'$ and $\sigma(.)$ known ; $x$ known.
- Observations: $\{X_{i,n} = (X_i(t_j), j = 1, \ldots, n), i = 1, \ldots N\}$ with $(n, N) \rightarrow \infty$.

# Random effects

Two cases not included in previous works.

## Two cases of mixed effects

1. $\Psi_i = \psi = \gamma^{-1/2}$ unknown and $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}) \Rightarrow \theta = (\gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$
2. $\Phi_i = \phi$ unknow ; $\Psi_i = \Gamma_i^{-1/2}$ with $\Gamma_i \sim G(a, \lambda) \Rightarrow \tau = (\lambda, a, \phi)$.

## Next talk (Valentine)

- $\Psi_i = \Gamma_i^{-1/2}$ with $\Gamma_i \sim G(a, \lambda)$.
- Given $\Gamma_i = \gamma$, $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega})$.

# Approximate conditional likelihood for $(X_i)$

- Derived from the Euler scheme $(Y_{i,n})$ of $(X_i)$ with $\Delta = T/n$.
- $Y_{i,n}(t_j) = Y_{i,n}(t_{j-1}) + \Delta\Phi_i' \, b(Y(t_{j-1})) + \sqrt{\Delta}\Gamma_i^{-1/2}\sigma(Y(t_{j-1})) + \epsilon_{i,j}$
- $((\epsilon_{i,j}), j = 1, \ldots, n)$ i.i.d $\mathcal{N}(0,1)$.

Conditionally on $\Phi_i = \varphi, \Psi_i = \psi$,

$$\mathcal{L}_n(X_{i,n}, \gamma, \varphi) = \gamma^{n/2} \exp\left[-\frac{\gamma}{2}(S_{i,n} + \varphi' V_{i,n}\varphi - 2\varphi' U_{i,n})\right], \quad \text{where}$$

$$S_{i,n} = S_i = \frac{1}{\Delta}\sum_{j=1}^{n} \frac{(X_i(t_j) - X_i(t_{j-1}))^2}{\sigma^2(X_i(t_{j-1}))},$$

$$V_{i,n} = V_i = \left(\sum_{j=1}^{n} \Delta \frac{b_k(X_i(t_{j-1}))b_\ell(X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))}\right)_{1\le k,\ell\le d},$$

$$U_{i,n} = U_i = \left(\sum_{j=1}^{n} \frac{b_k(X_i(t_{j-1}))(X_i(t_j) - X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))}\right)_{1\le k\le d},$$

# Approximate conditional likelihood for $(X_i)$ (2)

$$S_{i,n} = S_i = \frac{1}{\Delta} \sum_{j=1}^{n} \frac{(X_i(t_j) - X_i(t_{j-1}))^2}{\sigma^2(X_i(t_{j-1}))},$$

$$V_{i,n} = V_i = \left( \sum_{j=1}^{n} \Delta \frac{b_k(X_i(t_{j-1}))b_\ell(X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))} \right)_{1\le k,\ell \le d},$$

$$U_{i,n} = U_i = \left( \sum_{j=1}^{n} \frac{b_k(X_i(t_{j-1}))(X_i(t_j) - X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))} \right)_{1\le k \le d},$$

As $n \to \infty$, $S_{i,n}/n \to \Gamma_i^{-1}$ in probability;

$$V_{i,n} \to \left( \int_0^T \frac{b_k(X_i(s))b_\ell(X_i(s))}{\sigma^2(X_i(s))}\, ds \right)_{1\le k,\ell \le d} = V_i(T) \text{ a.s. };$$

$$U_{i,n} \to \left( \int_0^T \frac{b_k(X_i(s))}{\sigma^2(X_i(s))}\, dX_i(s) \right)_{1\le k \le d} = U_i(T) \text{ in probability.}$$

# Case (1): Random effects in the drift coefficient

$\Psi_i = \psi = \gamma^{-1/2}$ unknown and $\Phi_i \sim \mathcal{N}_d(\boldsymbol{\mu}, \gamma^{-1}\boldsymbol{\Omega}) \Rightarrow \theta = (\gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$

## Proposition: Approximate likelihood for $(X_{i,n})$

If $b(.), \sigma(.)$ bounded, $C^2$ and $V_i(T)$ positive definite a.s., then

$$
\begin{aligned}
\mathcal{L}_n(X_{i,n}, \vartheta) &= \gamma^{n/2}(det(\boldsymbol{I}_d + V_i\,\boldsymbol{\Omega}))^{-1/2} \exp -\frac{\gamma}{2}(S_i + T_i(\boldsymbol{\mu}, \boldsymbol{\Omega})) \text{ with} \\
T_i(\boldsymbol{\mu}, \boldsymbol{\Omega}) &= (\boldsymbol{\mu} - V_i^{-1}U_i)'\, R_i^{-1}\, (\boldsymbol{\mu} - V_i^{-1}U_i) - U_i'\, V_i^{-1}\, U_i), \\
R_{i,n} &= R_i = V_{i,n}^{-1} + \boldsymbol{\Omega}.
\end{aligned}
$$

## Approximate Loglikelihood for $(X_i\ i = 1, \ldots, N)$

$$
\ell_{N,n}(\vartheta) = \frac{Nn}{2}\log\gamma - \frac{1}{2}\sum_{i=1}^{N}\log det(\boldsymbol{I}_d + V_i\,\boldsymbol{\Omega}) - \frac{\gamma}{2}\sum_{i=1}^{N}(S_i + T_i(\boldsymbol{\mu}, \boldsymbol{\Omega})).
$$

⋆ Formula $\mathcal{L}_n(X_{i,n}, \vartheta)$ holds if $\boldsymbol{\Omega}$. is singular.

⋆ Possibility to have both fixed and random effects in the drift coefficient.

# Estimating equations for $\vartheta = (\gamma, \boldsymbol{\mu}, \boldsymbol{\Omega})$

Derived from the "pseudo score function":

$$\mathcal{G}_{N,n}(\vartheta) = \left( \frac{\partial}{\partial \gamma} \ell_{N,n}(\vartheta), \frac{\partial}{\partial \boldsymbol{\mu}} \ell_{N,n}(\vartheta), \frac{\partial}{\partial \boldsymbol{\Omega}} \ell_{N,n}(\vartheta) \right)'$$

We study the estimators defined by the estimating equation:

$$\mathcal{G}_{N,n}(\tilde{\vartheta}_{N,n}) = 0. \quad (4)$$

$$\begin{aligned} \frac{\partial}{\partial \gamma} \ell_{N,n}(\vartheta) &= \frac{Nn}{2\gamma} - \frac{1}{2} \sum_{i=1}^{N} (S_i + T_i(\boldsymbol{\mu}, \boldsymbol{\Omega})), \\ \nabla_{\boldsymbol{\mu}} \ell_{N,n}(\vartheta) &= \gamma \sum_{i=1}^{N} A_{i,n} \text{ with } A_{i,n} = R_i^{-1}(V_i^{-1} U_i - \boldsymbol{\mu}), \\ \nabla_{\boldsymbol{\Omega}} \ell_{N,n}(\vartheta) &= -\frac{1}{2} \sum_{i=1}^{N} R_i^{-1} + \frac{\gamma}{2} \sum_{i=1}^{N} A_{i,n} A'_{i,n}. \end{aligned}$$

# Study of estimators

Different rates of convergence for $\gamma$ and $\boldsymbol{\mu}, \boldsymbol{\Omega}$. Set $q = 1 + 2d$

$$D_{N,n}\begin{pmatrix} \frac{1}{\sqrt{Nn}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{N}}\boldsymbol{I}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{N}}\boldsymbol{I}_d \end{pmatrix}, \quad \mathcal{I}(\vartheta) = \left(\begin{array}{c|c} \frac{1}{2\gamma^2} & \mathbf{0} \\ \hline \mathbf{0} & I(\vartheta) \end{array}\right), \qquad (5)$$

**(H1)**: Assume that $b(.), \sigma(.)$ $C^2$ bounded; $\sigma(.) \geq \sigma_0 > 0$.
**(H2)** $V_i(T)$ positive definite a.s..

## Theorem 1

Assume (H1)-(H2)and $I(\vartheta)$ invertible. Then if $N, n \to \infty$ and $N/n \to 0$, there exists a solution $\tilde{\vartheta}_{N,n}$ with probability tending to 1 which is consistent and such that,

$$D_{N,n}^{-1}(\tilde{\vartheta}_{N,n} - \vartheta) \to_{\mathcal{D}} \mathcal{N}_q(0, \mathcal{I}^{-1}(\vartheta)) \text{ under } \mathbb{P}_{\vartheta_0}.$$

$I(\vartheta)$: explicit covariance matrix (detailed next slide).

# Study of estimators (2)

Set $R_i(T;\mathbf{\Omega}) = V_i(T)^{-1} + \mathbf{\Omega}$;
$B_i(T;\mathbf{\Omega}) = R_i^{-1}(T;\mathbf{\Omega})$
$A_i(T;\boldsymbol{\mu},\mathbf{\Omega}) = B_i(T;\mathbf{\Omega})(V_i(T)^{-1}U_i(T) - \boldsymbol{\mu})$ .
$I(\theta)$ : explicit expression depending on $A_i$, $R_i$.

$$I(\theta) = \begin{pmatrix} \gamma\mathbb{E}_\theta B_1(T;\mathbf{\Omega}) & \gamma\mathbb{E}_\theta A_1(T;\boldsymbol{\mu},\mathbf{\Omega})'\, B_1(T;\mathbf{\Omega}) \\ \gamma\mathbb{E}_\theta B_1(T;\mathbf{\Omega})A_1(T;\boldsymbol{\mu},\mathbf{\Omega}) & \mathbb{E}_\theta\left(\gamma A_1(T;\boldsymbol{\mu},\mathbf{\Omega})A_1(T;\boldsymbol{\mu},\mathbf{\Omega})' - \frac{1}{2}B_1(T;\mathbf{\Omega})\right) \end{pmatrix}.$$

Delattre et al. (2013) : $I(\vartheta)$ covariance matrix of

$$\begin{pmatrix} \gamma A_1(T;\boldsymbol{\mu},\mathbf{\Omega}) \\ \frac{1}{2}(\gamma A_1(T;\boldsymbol{\mu},\mathbf{\Omega})A_1(T;\boldsymbol{\mu},\mathbf{\Omega})' - B_1(T;\mathbf{\Omega}) \end{pmatrix}$$

# Comments

1. Theorem holds if $\Omega$ singular:
   Possible to include mixed effects in the drift coefficient.
2. Fixed and random effects in the drift: Same rates of convergence.
3. Possible to estimate $\gamma$ from one trajectory. (by $n/S_{i,n}$) (large bias).
4. No loss of information from the discrete observations:
   Continuous observations $(X_i(t), i = 1, \ldots N)$ ($\gamma$ known, $d = 1$):
   Delattre et al.(2013): M.L.E $\hat{\theta}_{N,n}$ strongly consistent and same asymptotic variance for $\hat{\theta}_{N,n}$ and $\tilde{\theta}_{N,n}$.
5. Loss of efficiency w.r.t. direct observations of $\Phi_i \sim \mathcal{N}_2(\mu, \gamma^{-1}\omega^2)$:
   Fisher information $J_0(\mu, \omega^2)$.

## Example of Brownian motion with drift

$dX_i(t) = \Phi_i dt + \gamma^{-1/2} dW_i(t); \quad X_i(0) = 0.$

$$I(\theta) = \begin{pmatrix} \frac{\gamma}{\frac{1}{T} + \omega^2} & 0 \\ 0 & \frac{1}{2(\frac{1}{T} + \omega^2)^2} \end{pmatrix} \text{ to compare with } J_0(\mu, \omega^2) = \begin{pmatrix} \frac{\gamma}{\omega^2} & 0 \\ 0 & \frac{1}{2\omega^4} \end{pmatrix}.$$

# Assessment of the method on various examples

**Ex. 1.** Ornstein-.Uhlenbeck diffusion with one random effect.
$dX_i(t) = \phi_i X_i(t)dt + \frac{1}{\sqrt{\gamma}} dW_i(t), X_i(0) = 0$ , $\phi_i \underset{i.i.d}{\sim} \mathcal{N}(\mu, \frac{\omega^2}{\gamma})$
**Ex. 2.** Diffusion with bounded $b(.), \sigma(.)$.
$dX_i(t) = \phi_i X_i(t)^2/(1+X_i(t)^2)dt + \frac{1}{\sqrt{\gamma}} dW_i(t), X_i(0) = 0$ , $\phi_i \underset{i.i.d}{\sim} \mathcal{N}(\mu, \frac{\omega^2}{\gamma})$
**Ex. 3.** O.U diffusion with one fixed and one random effects.
$dX_i(t) = (\rho X_i(t) + \phi_i)dt + \frac{1}{\sqrt{\gamma}} dW_i(t), X_i(0) = 0$ , $\phi_i \underset{i.i.d}{\sim} \mathcal{N}(\mu, \frac{\omega^2}{\gamma})$
**Ex. 4.** O-U diffusion with two independent random effects.
$dX_i(t) = (\phi_{i1} X_i(t) + \phi_{i2})dt + \frac{1}{\sqrt{\gamma}} dW_i(t), X_i(0) = 0$ ;
$\phi_{i1} \underset{i.i.d}{\sim} \mathcal{N}(\mu_1, \frac{\omega_1^2}{\gamma})$ , $\phi_{i2} \underset{i.i.d}{\sim} \mathcal{N}(\mu_2, \frac{\omega_2^2}{\gamma})$

# Details on the simulations

1. Choice of SDEME model : Diffusion; random effects ; sampling $\Delta$ ; nb of obs. $n$ ($T = n\Delta$); nb of paths $N$
2. Each scenario: generation of 100 data sets $\Rightarrow$ Empirical mean and standard deviations in the tables.
3. Each data set: (1) draw the random effects; (2) Diffusion process path: either exact or obtained with a Euler scheme with $\delta = 0.001$
4. Comparison with direct observations of the random effects .

# Ornstein-Uhlenbeck diffusion with one random effect

| | | $N=50$ | | $N=100$ | |
|---|---|---|---|---|---|
| | | $n=500$ | $n=1000$ | $n=500$ | $n=1000$ |
| | | | $(\mu_0=0, \omega_0^2=0.1, \gamma_0=4)$ | | |
| $X$ | $\tilde{\mu}$ | 0.00 (0.07) | 0.00 (0.04) | 0.00 (0.04) | 0.00 (0.03) |
| | $\tilde{\omega}^2$ | 0.09 (0.08) | 0.10 (0.04) | 0.09 (0.05) | 0.10 (0.03) |
| | $\tilde{\gamma}$ | 4.00 (0.03) | 4.00 (0.03) | 4.00 (0.02) | 4.00 (0.02) |
| $\phi$ | $\hat{\mu}$ | 0.00 (0.02) | 0.00 (0.02) | 0.00 (0.01) | 0.00 (0.02) |
| | $\hat{\omega}^2$ | 0.10 (0.02) | 0.10 (0.02) | 0.10 (0.01) | 0.10 (0.01) |

Table: Example 1. Empirical mean and standard deviation (in brackets) of the parameter estimates from 100 datasets for different values of $N$ and $n$. Estimates based on the $\phi_i$'s ($\phi$) and estimates based on the SDE ($X$) are given.

$$dX_i(t) = \phi_i X_i(t)dt + \frac{1}{\sqrt{\gamma}} dW_i(t), X_i(0) = 0 \ , \ \phi_i \underset{i.i.d}{\sim} \mathcal{N}(\mu, \frac{\omega^2}{\gamma})$$

$$dX_i(t) = (\rho X_i(t) + \phi_i)dt + \frac{1}{\sqrt{\gamma}} dW_i(t), \phi_i \underset{i.i.d}{\sim} \mathcal{N}(\mu, \frac{\omega^2}{\gamma})$$

| | | $N = 50$ | | $N = 100$ | |
|---|---|---|---|---|---|
| | | $n = 500$ | $n = 1000$ | $n = 500$ | $n = 1000$ |
| | | $(\mu_0 = 1, \omega_0^2 = 1, \gamma_0 = 10, \rho_0 = -0.1)$ | | | |
| $X$ | $\tilde{\mu}$ | 1.00 (0.05) | 1.01 (0.05) | 1.00 (0.03) | 1.00 (0.04) |
| | $\tilde{\omega}^2$ | 0.99 (0.26) | 1.01 (0.26) | 1.00 (0.21) | 0.97 (0.15) |
| | $\tilde{\gamma}$ | 10 (0.10) | 10.03 (0.06) | 10.00 (0.05) | 10.01 (0.05) |
| | $\tilde{\rho}$ | -0.10 (0.02) | -0.10 (0.01) | -0.10 (0.01) | -0.10 (0.01) |
| $\phi$ | $\hat{\mu}$ | 1.00 (0.04) | 1.00 (0.05) | 1.00 (0.03) | 1.00 (0.03) |
| | $\hat{\omega}^2$ | 1.00 (0.21) | 1.00 (0.22) | 1.00 (0.15) | 0.99 (0.14) |
| | | $(\mu_0 = 1, \omega_0^2 = 0.4, \gamma_0 = 4, \rho_0 = -0.1)$ | | | |
| $X$ | $\tilde{\mu}$ | 1.01 (0.08) | 1.02 (0.06) | 1.01 (0.06) | 1.00 (0.05) |
| | $\tilde{\omega}^2$ | 0.40 (0.14) | 0.40 (0.12) | 0.41 (0.10) | 0.40 (0.09) |
| | $\tilde{\gamma}$ | 4.01 (0.04) | 4.01 (0.02) | 4.01 (0.03) | 4.01 (0.02) |
| | $\tilde{\rho}$ | -0.11 (0.02) | -0.10 (0.01) | -0.10 (0.02) | -0.10 (0.01) |
| $\phi$ | $\hat{\mu}$ | 1.00 (0.05) | 1.00 (0.04) | 1.00 (0.03) | 1.00 (0.03) |
| | $\hat{\omega}^2$ | 0.39 (0.06) | 0.41 (0.09) | 0.41 (0.07) | 0.40 (0.06) |

$$dX_i(t) = (\phi_{i1}X_i(t) + \phi_{i2})dt + \frac{1}{\sqrt{\gamma}}dW_i(t)$$

| | | $N=50$ | | $N=100$ | |
|---|---|---|---|---|---|
| | | $n=500$ | $n=1000$ | $n=500$ | $n=1000$ |
| | | $(\mu_{1,0}=-0.1, \mu_{2,0}=1, \omega^2_{1,0}=0.1, \omega^2_{2,0}=1, \gamma_0=10, \rho_0=-0.1)$ | | | |
| $X$ | $\tilde{\mu}_1$ | -0.10 (0.03) | -0.10 (0.02) | -0.10 (0.02) | -0.10 (0.01) |
| | $\tilde{\omega}^2_1$ | 0.10 (0.04) | 0.10 (0.02) | 0.10 (0.03) | 0.10 (0.02) |
| | $\tilde{\mu}_2$ | 1.00 (0.07) | 1.00 (0.05) | 1.00 (0.04) | 1.00 (0.04) |
| | $\tilde{\omega}^2_2$ | 0.96 (0.28) | 0.97 (0.28) | 0.94 (0.19) | 0.99 (0.20) |
| | $\tilde{\gamma}$ | 10.03 (0.09) | 10.02 (0.07) | 10.03 (0.07) | 10.02 (0.05) |
| $\phi$ | $\hat{\mu}_1$ | -0.10 (0.02) | -0.10 (0.01) | -0.10 (0.01) | -0.10 (0.01) |
| | $\hat{\omega}^2_1$ | 0.10 (0.02) | 0.10 (0.02) | 0.10 (0.01) | 0.10 (0.02) |
| | $\hat{\mu}_2$ | 1.00 (0.05) | 1.00 (0.04) | 1.00 (0.03) | 1.00 (0.03) |
| | $\hat{\omega}^2_2$ | 1.00 (0.21) | 1.01 (0.23) | 0.98 (0.13) | 1.00 (0.14) |
| | | $(\mu_{1,0}=-0.1, \mu_{2,0}=1, \omega^2_{1,0}=0.04, \omega^2_{2,0}=0.4, \gamma_0=4)$ | | | |
| $X$ | $\tilde{\mu}_1$ | -0.10 (0.04) | -0.11 (0.02) | -0.10 (0.02) | -0.10 (0.01) |
| | $\tilde{\omega}^2_1$ | 0.04 (0.02) | 0.04 (0.01) | 0.04 (0.01) | 0.04 (0.01) |
| | $\tilde{\mu}_2$ | 0.98 (0.08) | 1.01 (0.06) | 1.00 (0.05) | 1.00 (0.05) |
| | $\tilde{\omega}^2_2$ | 0.37 (0.15) | 0.37 (0.13) | 0.39 (0.11) | 0.41 (0.10) |
| | $\tilde{\gamma}$ | 4.02 (0.03) | 4.00 (0.03) | 4.01 (0.02) | 4.00 (0.08) |

# Results for random effects in the drift coefficient

- RESULTS: satisfactory overall.
- Three designs : $(N, n) = (50, 500); (100, 500); (100, 1000)$:
- Model parameters estimated with very little bias.
- Increasing $N$ and $n$ reduces the bias and the standard deviation.
- $N = 100, n = 100$ :estimations similar to those based direct observation of the random effects.
- Good results in Ex. 3 with both random and fixed effects in the drift.
- Gives evidence of the validity of our method for singular covariance matrix $\Omega$ of the random effects.

# Case (2): Random effect in the diffusion coefficient

- $\Phi_i = \varphi$ unknown; $\Psi_i = \Gamma_i^{-1/2}$ with $\Gamma_i \sim G(a, \lambda)$
- Parameter: $\tau = (\lambda, a, \varphi)$
- Conditionally on $\Phi_i = \varphi, \Psi_i = \psi$, approximate likelihood for $(X_{i,n})$:

$$\mathcal{L}_n(X_{i,n}, \gamma, \varphi) = \gamma^{n/2} \exp\left[-\frac{\gamma}{2}(S_{i,n} + \varphi' V_{i,n}\varphi - 2\varphi' U_{i,n})\right].$$

- Integrating w.r.t the $\Gamma_i$ :

$$\widetilde{\Lambda}_n(X_i, \tau) = \frac{\lambda^a \Gamma(a + (n/2))}{\Gamma(a)(\lambda + \frac{1}{2}(S_i - 2\varphi' U_i + \varphi' V_i \varphi))^{a+(n/2)}}, \qquad (6)$$

- For the $N$ paths approximate loglikelihood,

$$\widetilde{\ell}_{N,n}(\tau) = \sum_{i=1}^{N} \log \widetilde{\Lambda}_n(X_i, \tau).$$

# Recap of the terms in the likelihood

$$S_{i,n} = S_i = \frac{1}{\Delta} \sum_{j=1}^{n} \frac{(X_i(t_j) - X_i(t_{j-1}))^2}{\sigma^2(X_i(t_{j-1}))},$$

$$V_{i,n} = V_i = \left( \sum_{j=1}^{n} \Delta \frac{b_k(X_i(t_{j-1})) b_\ell(X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))} \right)_{1 \leq k,\ell \leq d},$$

$$U_{i,n} = U_i = \left( \sum_{j=1}^{n} \frac{b_k(X_i(t_{j-1}))(X_i(t_j) - X_i(t_{j-1}))}{\sigma^2(X_i(t_{j-1}))} \right)_{1 \leq k \leq d},$$

As $n \to \infty$, $S_{i,n}/n \to \Gamma_i^{-1}$ in probability;

$$V_{i,n} \to \left( \int_0^T \frac{b_k(X_i(s)) b_\ell(X_i(s))}{\sigma^2(X_i(s))} \right)_{1 \leq k,\ell \leq d} = V_i(T) \text{ a.s. };$$

$$U_{i,n} \to \left( \int_0^T \frac{b_k(X_i(s))}{\sigma^2(X_i(s))} \, dX_i(s) \right)_{1 \leq k \leq d} = U_i(T) \text{ in probability.}$$

# Estimating $\tau$

$$\widetilde{G}_{N,n}(\tau) = \left( \frac{\partial}{\partial\lambda}\widetilde{\ell}_{N,n}(\tau) \ \ \frac{\partial}{\partial a}\widetilde{\ell}_{N,n}(\tau) \ \ \frac{\partial}{\partial\varphi}\widetilde{\ell}_{N,n}(\tau) \right)'. \qquad (7)$$

Estimators $\widetilde{\tau}_{N,n}$ such that $\widetilde{G}_{N,n}(\widetilde{\tau}_{N,n}) = 0$.
Central random variable approximating $\Gamma_i^{-1}$:

$$\zeta_i(\tau) = \zeta_i = \frac{\lambda + \frac{1}{2}\left(S_i - 2\varphi U_i + \varphi^2 V_i\right)}{a + (n/2)}. \qquad (8)$$

Three estimating equations ($\psi(z) = \Gamma'(z)/\Gamma(z)$ ):

$$\frac{\partial}{\partial\lambda}\widetilde{\ell}_{N,n}(\tau) = \sum_{i=1}^{N}\left(\frac{a}{\lambda} - \zeta_i^{-1}\right), \frac{\partial}{\partial\varphi}\widetilde{\ell}_{N,n}(\tau) = \sum_{i=1}^{N}\zeta_i^{-1}(U_i - \varphi V_i).$$

$$\frac{\partial}{\partial a}\widetilde{\ell}_{N,n}(\tau) = \sum_{i=1}^{N}(\log\lambda - \psi(a) - \log\zeta_i) + N\left(\psi(a + (n/2)) - \log\left(a + (n/2)\right)\right),$$

Need to control the moments of $\zeta_i(\tau)$ and $\zeta_i^{-1}(\tau)$

# Convergence theorem

## Theorem 2

Assume (H1)-(H2). If $a > 5$ and $N, n$ tend to infinity with $N/n \to 0$. Then, there exists a solution $\widetilde{\tau}_{N,n}$ to the estimating equations with probability tending to 1, consistent and such that $\sqrt{N}(\widetilde{\tau}_{N,n} - \tau)$ converges in distribution under $\mathbb{P}_\tau$ to $\mathcal{N}_3(0, \mathcal{V}^{-1}(\tau))$ where $I_0(\lambda, a)$ is defined and

$$\mathcal{V}(\tau) = \left(\begin{array}{c|c} I_0(\lambda, a) & \mathbf{0} \\ \hline \mathbf{0} & \mathbb{E}_\tau\left(\Gamma \int_0^T \frac{b^2(X(s))}{\sigma^2(X(s))} ds\right) \end{array}\right).$$

For the first two components of $\widetilde{\tau}_{N,n}$, the constraint $N/n^2 \to 0$ is enough.

$I_0(\lambda, a)$: Fisher information of direct observations of $\Gamma_i$

$$I_0(\lambda, a) = \begin{pmatrix} \frac{a}{\lambda^2} & -\frac{1}{\lambda} \\ -\frac{1}{\lambda} & \Psi'(a) \end{pmatrix}$$

# Comments

- Estimator of $(\lambda, a)$ based on the indirect obs. $(X_i)$asympt. equivalent to the MLE of $(\lambda, a)$ based on the $\Gamma_i$'s if $N/n^2 \to 0$.
- Contrary to Case (1), same rates of convergence for $(\lambda, a)$ and $\varphi$.
- Estimation of fixed $\varphi, \gamma$ for $N$ paths:
  $dX_i(t) = \varphi b(X_i(t))dt + \gamma^{-1/2}\sigma(X_i(t))dW_i(t), X_i(0) = x, i = 1, \dots, N.$

$$\hat{\varphi}_{N,n} = \frac{\sum_{i=1}^N U_i}{\sum_{i=1}^N V_i}, \quad \hat{\gamma}_{N,n} = \frac{nN}{\sum_{i=1}^N S_i + \hat{\varphi}_{N,n}^2 V_i - 2\hat{\varphi}_{N,n} U_i}.$$

- $\sqrt{Nn}(\hat{\gamma}_{N,n} - \gamma) \to \mathcal{N}(0, 2\gamma^2)$.
- If $N/n \to 0$, $\hat{\varphi}_{N,n}$ $\sqrt{N}(\hat{\varphi}_{N,n} - \varphi)$ converge in distribution to $\mathcal{N}(0, (\gamma \mathbb{E} V(T))^{-1})$.
- Result obtained for $\tilde{\varphi}_{N,n}$ in Theorem 2 not surprising.

# Assessment of the method on various examples

Several models are simulated:

**Ex. 5.** Brownian motion with drift and random diffusion coefficient

$dX_i(t) = \rho dt + \Gamma_i^{-1/2} dW_i(t), X_i(0) = 0$ , $\Gamma_i \underset{i.i.d}{\sim} G(a, \lambda)$

**Ex. 6.** Ornstein-Uhlenbeck with random diffusion coef.

$dX_i(t) = \rho X_i(t)dt + \Gamma_i^{-1/2} dW_i(t), X_i(0) = x$ , $\Gamma_i \underset{i.i.d}{\sim} G(a, \lambda)$

**Ex. 7.** Diffusion with varying $\sigma(.)$

$dX_i(t) = \rho X_i(t)dt + \Gamma_i^{-1/2} \sqrt{1 + X_i(t)^2} dW_i(t), X_i(0) = x$ , $\Gamma_i \underset{i.i.d}{\sim} G(a, \lambda)$

Results given in Tables 5-7.

Estimations based on the processes and estimations based on direct observations of the random effects are compared

$$dX_i(t) = \rho dt + \Gamma_i^{-1/2} dW_i(t)$$

| | | $N = 50$ | | | $N = 100$ | | |
|---|---|---|---|---|---|---|---|
| | | $n = 500$ | $n = 1000$ | $n = 10000$ | $n = 500$ | $n = 1000$ | $n = 10000$ |
| | | | | $(a_0 = 8, \lambda_0 = 2, \rho_0)$ | | | |
| $X$ | $\tilde{a}$ | 6.03 | 7.03 | 8.32 | 6.05 | 6.88 | 8.14 |
| | | (0.73 - 1.57) | (0.91 - 1.57) | (1.80 - 1.57) | (0.44 - 1.11) | (0.73 - 1.11) | (1.11 - 1.11) |
| | $\tilde{\lambda}$ | 1.50 | 1.76 | 2.08 | 1.51 | 1.71 | 2.04 |
| | | (0.20 - 0.40) | (0.24 - 0.40) | (0.44 - 0.40) | (0.12 - 0.29) | (0.19 - 0.29) | (0.30 - 0.29) |
| | $\tilde{\rho}$ | -0.99 | -1.00 | -1.00 | -1.00 | -1.00 | -1.00 |
| | | (0.03 - 0.03) | (0.02 - 0.02) | (0.01 - 0.01) | (0.02 - 0.02) | (0.02 - 0.02) | (0.01 - 0.01) |
| $\psi$ | $\hat{a}$ | 8.23 | 8.57 | 8.42 | 8.38 | 8.32 | 8.27 |
| | | (1.77 - 1.57) | (1.76 - 1.57) | (2.15 - 1.57) | (1.16 - 1.11) | (1.35 - 1.11) | (1.31 - 1.11) |
| | $\hat{\lambda}$ | 2.06 | 2.14 | 2.10 | 2.09 | 2.07 | 2.07 |
| | | (0.46 - 0.40) | (0.45 - 0.40) | (0.53 - 0.40) | (0.30 - 0.29) | (0.34 - 0.29) | (0.34 - 0.29) |

Table 1: Example 5. Empirical mean and, in brackets, (empirical standard deviation - theoretical standard deviation) of the parameter estimates from 100 datasets for different values of $N$ and $n$. Estimates based on the $\psi_i$'s ($\psi$) and estimates based on the SDE ($X$) are given.

$$dX_i(t) = \rho X_i(t)dt + \Gamma_i^{-1/2} dW_i(t),\ \Gamma_i \underset{i.i.d}{\sim} G(a, \lambda)$$

| | | $N = 50$ | | | $N = 100$ | | |
|---|---|---|---|---|---|---|---|
| | | $n = 500$ | $n = 1000$ | $n = 10000$ | $n = 500$ | $n = 1000$ | $n = 10000$ |
| | | | | $(a_0 = 8, \lambda_0 = 2, \rho_0)$ | | | |
| $X$ | $\tilde{a}$ | 6.18 (0.66) | 6.79 (0.99) | 8.20 (1.61) | 6.02 (0.50) | 6.70 (0.63) | 7.91 (1.04) |
| | $\tilde{\lambda}$ | 1.52 (0.17) | 1.69 (0.27) | 2.03 (0.42) | 1.49 (0.14) | 1.66 (0.16) | 1.96 (0.25) |
| | $\tilde{\rho}$ | -1.00 (0.10) | -1.01 (0.06) | -0.99 (0.02) | -1.00 (0.06) | -1.00 (0.04) | -1.00 (0.01) |
| $\psi$ | $\hat{a}$ | 8.71 (1.80) | 8.12 (1.94) | 8.20 (1.61) | 8.16 (1.30) | 8.10 (1.19) | 8.10 (1.16) |
| | $\hat{\lambda}$ | 2.16 (0.45) | 2.04 (0.51) | 2.07 (0.47) | 2.04 (0.33) | 2.02 (0.30) | 2.03 (0.29) |

Table 1: Example 6. Empirical mean and standard deviation (in brackets) of the parameter estimates from 100 datasets for different values of $N$ and $n$. Estimates based on the $\psi_i$'s ($\psi$) and estimates based on the SDE ($X$) are given.