Quasi-likelihood inference for SDE with mixed-effects * †

Maud Delattre[1] Hiroki Masuda[2]

[1]Université Paris-Saclay, INRAE, MaIAGE

[2]The University of Tokyo and JST CREST

Workshop on statistics of stochastic processes
in honor of Yury Kutoyants' 80th birthday

September 17, 2026



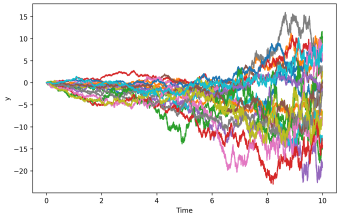
*Partially supported by JST CREST (JPMJCR2115) and JSPS KAKENHI (23K22410).

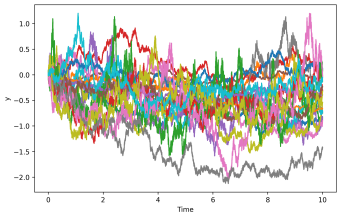
†This version: September 17, 2026

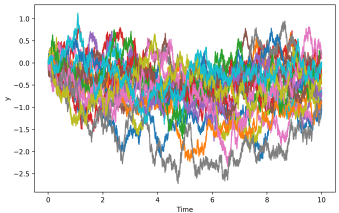
- **SDE with mixed effects [Picchini et al., 2010]**: in general,

$$Y_i(t) = Y_i(0) + \int_0^t \mu(Y_i(s), \varphi_i)ds + \int_0^t \sigma(Y_i(s), \tau_i)dw_i(s), \qquad i = 1, \ldots, N.$$

  - Pharmacokinetics: references in [Donnet and Samson, 2013]), ...
  - Neurobiology: [Ditlevsen and Samson, 2013], ...
  - Epidemic dynamics prediction: [Narci et al., 2022], ...







Contents



Delattre, M. and Masuda, H. (2025),
Quasi-likelihood inference for SDE with mixed-effects observed at high frequency.
arxiv:2508.17910

1. Proposed model
2. Quasi-likelihood estimation
3. Applications to synthetic and real data
4. Concluding remarks

1 Proposed model

2 Quasi-likelihood estimation

3 Applications to synthetic and real data

4 Concluding remarks

Underlying dynamics

$$Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(Y_i(s))ds \right) + \sqrt{\tau_i} \int_0^t c(Y_i(s); \eta) dw_i(s), \qquad i = 1, \dots, N, \quad t \in [0, T].$$

- $w_1(\cdot), w_2(\cdot), \dots$ are i.i.d. standard 1-dim. Wiener processes independent of
- i.i.d. $Y_1(0), Y_2(0), \dots$ such that $E\left[|Y_1(0)|^K\right] < \infty$ for every $K > 0$.
- $a(y) := (a_f(y), a_r(y)) : \mathbb{R} \to \mathbb{R}^{p_\varphi}$ and $c(y; \eta) : \mathbb{R} \times \Theta_\eta \to \mathbb{R}$ are m'ble func's, with η unknown.

- $\eta \in \Theta_\eta \subset \mathbb{R}^{p_\eta}$ and $\varphi_f \in \Theta_{\varphi_f} \subset \mathbb{R}^{p_{\varphi_f}}$ are **fixed-effect parameters**.
- $(\tau_1, \varphi_{r,1}), (\tau_2, \varphi_{r,2}), \dots \in (0, \infty) \times \mathbb{R}^{p_{\varphi_r}}$ are i.i.d. **random-effect parameters** s.t.
  - τ_i and $\varphi_{r,i}$ are independent.
  - $\tau_i \sim f(\tau; \theta_\tau) d\tau$ on $(0, \infty)$; $\theta_\tau \in \Theta_\tau \subset \mathbb{R}^{p_\tau}$.
  - $\varphi_{r,i} \sim N_{p_{\varphi_r}}(0, \Sigma_r)$ for a positive definite $\Sigma_r \in \mathbb{R}^{p_{\varphi_r}} \otimes \mathbb{R}^{p_{\varphi_r}}$.

Assumptions in brief

$$Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(Y_i(s))ds \right)$$
$$+ \sqrt{\tau_i} \int_0^t c(Y_i(s); \eta) dw_i(s), \qquad i = 1, \ldots, N, \quad t \in [0, T],$$
$$\tau_i \sim f(\tau; \theta_\tau)d\tau, \qquad \varphi_{r,i} \sim N_{p_{\varphi_r}}(0, \Sigma_r).$$

- We observe $\{Y_i(t_{ij}) : j = 0, 1, \ldots, n;\ i = 1, \ldots, N\}$, where $t_{ij} = jh$ with $h := T/n$ (fixed T).
- About (a, c): smoothness, non-degeneracy of c, moment conditions, and identifiability.
- Diffusion random-effect distribution $\tau_i \sim f(\tau; \theta_\tau)d\tau$:
  - $\exists C_\tau > 0 \quad \max_{k=0,1,2,3} \max_{l=0,1} \sup_{\theta_\tau} \left| \partial_{\theta_\tau}^k \partial_\tau^l \log f(\tau; \theta_\tau) \right| \lesssim 1 + \tau^{C_\tau} + \tau^{-C_\tau}$;
  - $f(\tau; \theta_\tau) = f(\tau; \theta_{\tau,0})$ ($\mathcal{L}(\tau_1)$-a.e.) if and only if $\theta_\tau = \theta_{\tau,0}$;
  - $\forall \theta_\tau \in \overline{\Theta_\tau}\ \forall K_\tau \in \mathbb{R}, \quad \int_{0+}^{\infty} y^{K_\tau} f(y; \theta_\tau)dy < \infty.$

1 Proposed model

2 Quasi-likelihood estimation

3 Applications to synthetic and real data

4 Concluding remarks

$$Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(Y_i(s))ds \right) + \sqrt{\tau_i} \int_0^t c(Y_i(s); \eta)dw_i(s)$$

1. Profile quasi-likelihood for the diffusion parameters
2. Marginal quasi-likelihood for the drift parameters

To put it simply: Multi-stage estimation procedure

$$
\begin{aligned}
Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(Y_i(s))ds \right) \\
+ \sqrt{\tau_i} \int_0^t c(Y_i(s); \eta) dw_i(s), \qquad i = 1, \ldots, N, \quad t \in [0, T], \\
\tau_i \sim f(\tau; \theta_\tau)d\tau, \qquad \varphi_{r,i} \sim N_{p_{\varphi_r}}(0, \Sigma_r).
\end{aligned}
$$

Asymptotics for $n, N \to \infty$ with $n^{\mathsf{a}'} \lesssim N \lesssim n^{\mathsf{a}''}$ $(0 < \mathsf{a}' \leq \mathsf{a}'' < 1)$

1. Ignore the drift and focus on the **"diffusion" parameter** $\vartheta_1 := (\eta, \theta_\tau) \in \Theta_\eta \times \Theta_\tau$

$$
\left(\sqrt{nN}(\widehat{\eta} - \eta_0),\ \sqrt{N}(\widehat{\theta}_\tau - \theta_{\tau,0}) \right) \xrightarrow{\mathcal{L}} N_{p_\eta + p_\tau} \left(0,\ \mathrm{diag}\left(Q_{11,0}^{-1},\ \mathcal{I}_{12}^{-1} \right) \right)
$$

2. With $\widehat{\vartheta}_1$, estimate the **"drift" parameter** $\vartheta_2 := (\varphi_f, \Sigma_r) \in \Theta_{\varphi_f} \times \Theta_{\Sigma_r}$

$$
\sqrt{N}(\widehat{\vartheta}_2 - \vartheta_{2,0}) \xrightarrow{\mathcal{L}} N_{q_\varphi} \left(0,\ \mathcal{J}_{2,0}^{-1} \right)
$$

Remarks: Why stepwise?

- $\vartheta = (\vartheta_1, \vartheta_2) \in \operatorname{argmax} \sum_i \sum_j \log \iint \phi(Y_i(t_i j); (\vartheta, \tau_i, \varphi_{r,i})\text{-func.}) dP^{\tau_i} dP^{\varphi_{r,i}}$?
- Possible to obtain the explicit QLF via the same concept as [Fujinaga and Masuda, 2023]:

$$\mathbb{H}^{(1)}_{N,n}(\vartheta_1) := -\frac{nN}{2}\log(2\pi h) + \frac{nN}{2}\log\gamma - \frac{1}{2}\sum_{i,j}\log S_{i,j-1}(\eta) - \lambda N \log\delta - N\log K_\lambda(\delta\gamma)$$
$$+ \sum_i \log K_{\lambda - n/2}\left(\gamma\sqrt{\delta^2 + Q_i(\eta)}\right) + \frac{1}{2}\left(\lambda - \frac{n}{2}\right)\sum_i \log\left(\delta^2 + Q_i(\eta)\right),$$

where $Q_i(\eta) := \sum_j S^{-1}_{i,j-1}(\eta)\, y^2_{ij}$.

- Should be theoretically more efficient,
- while the numerical optimization involves the modified Bessel function $K_{\lambda - n/2}$,
- hence much less stable and suffers from the usual local-optimum problem,
- and possible one-step improvement.

Stage 1: Profile quasi-likelihood for diffusion parameters $\vartheta_1 = (\eta, \theta_\tau)$

- Euler approximation ($S_{i,j-1}(\eta) := c^2_{i,j-1}(\eta)$, $\Delta_j w_i := w_i(t_{i,j}) - w_i(t_{i,j-1})$, etc.):

$$\begin{aligned} Y_{ij} &\overset{\vartheta}{=} Y_{i,j-1} + \tau_i \varphi_i \cdot \int_j a(Y_i(s))ds + \sqrt{\tau_i} \int_j c(Y_i(s);\eta)dw_i(s) \\ &\approx Y_{i,j-1} + \tau_i \varphi_i \cdot a_{i,j-1}h + \sqrt{\tau_i}\, c_{i,j-1}(\eta)\Delta_j w_i \\ &\approx Y_{i,j-1} + \sqrt{\tau_i}\, c_{i,j-1}(\eta)\Delta_j w_i. \end{aligned}$$

- **Gaussian quasi-likelihood (GQLF)** regarding $\{\tau_i\}_i$ as a fixed effect parameter:

$$(\eta, \tau_1, \ldots, \tau_N) \mapsto \sum_{i,j} \log \phi\big(\overline{y}_{ij}; Y_{i,j-1}, h\tau_i\, S_{i,j-1}(\eta)\big).$$

- Solve the associated quasi-likelihood eq. with respect to $\tau_1, \ldots, \tau_N$: for $\overline{y}_{ij} := h^{-1/2}\Delta_j Y_i$,

$$\widehat{\tau}_i(\eta) := \frac{1}{n}\sum_j S^{-1}_{i,j-1}(\eta)\, \overline{y}^2_{ij}, \qquad i = 1, \ldots, N.$$

- QLF when both $\{Y_{ij}\}_{i,j}$ and $\{\tau_i\}_i$ were observed:

$$(\eta, \theta_\tau) \mapsto \sum_{i,j} \log \phi\left(Y_{ij}; Y_{i,j-1}, h\tau_i\, S_{i,j-1}(\eta)\right) + \sum_i \log f(\tau_i; \theta_\tau). \tag{2.1}$$

- Replace τ_i in (2.1) by $\widehat{\tau}_i(\eta)$ to obtain the first-stage QLF $\mathbb{H}_1(\vartheta_1)$:

$$(\eta, \theta_\tau) \mapsto \sum_{i,j} \log \phi\left(Y_{ij}; Y_{i,j-1}, h\,\widehat{\tau}_i(\eta)\, S_{i,j-1}(\eta)\right) + \sum_i \log f(\widehat{\tau}_i(\eta); \theta_\tau).$$

- **The two-stage quasi-MLE** $\widehat{\vartheta}_1 = (\widehat{\eta}, \widehat{\theta}_\tau)$:

$$\widehat{\eta} \in \operatorname*{argmax}_\eta \mathbb{H}_{11}(\eta), \quad \mathbb{H}_{11,N,n}(\eta) := -\frac{1}{2} \sum_i \left\{ \sum_j \log S_{i,j-1}(\eta) + n \log \left(\frac{1}{n} \sum_j S_{i,j-1}^{-1}(\eta)\, \overline{y}_{ij}^2 \right) \right\},$$

$$\widehat{\theta}_\tau \in \operatorname*{argmax}_{\theta_\tau} \mathbb{H}_{12}(\widehat{\eta}, \theta_\tau), \quad \mathbb{H}_{12,N,n}(\eta, \theta_\tau) := \sum_i \log f(\widehat{\tau}_i(\eta); \theta_\tau).$$

Stage 2: Marginal quasi-likelihood for drift parameters $\vartheta_2 = (\varphi_f, \Sigma_r)$

- The Euler approximation in this second stage:

$$Y_{ij} \overset{\vartheta}{\approx} Y_{i,j-1} + \tau_i \, \varphi_i \cdot a_{i,j-1} h + \sqrt{\tau_i} \, c_{i,j-1}(\eta) \Delta_j w_i.$$

- Auxiliary Gaussian quasi-likelihood of $\{\varphi_i\} \mapsto \prod_i \widehat{g}_i(\varphi_i)$, where

$$\widehat{g}_i(\varphi_i) := \exp\left(\sum_j \log \phi\left(Y_{ij}; Y_{i,j-1} + h\widehat{\tau}_i \, \varphi_i \cdot a_{i,j-1}, \, h\widehat{\tau}_i \, \widehat{S}_{i,j-1} \right) \right),$$

  with $\widehat{S}_{i,j-1} = c^2_{i,j-1}(\widehat{\eta})$, $\widehat{\tau}_i = \widehat{\tau}_i(\widehat{\eta})$.

- Then, we integrate out the variable $\varphi_1, \ldots, \varphi_N$ with respect to the Gaussian distribution $N_{p_\varphi}(\mu, \Sigma)^{\otimes N}$ **as if** $\Sigma = \mathrm{diag}(O, \Sigma_r)$ **is non-degenerate**, as in [Delattre et al., 2018].

- The resulting expression: for $i = 1, \dots, N$,

$$\vartheta_2 \mapsto \int_{\mathbb{R}^{p_\varphi}} \widehat{g}_i(\varphi_i)\phi_{p_\varphi}(\varphi_i; \mu, \Sigma)d\varphi_i$$
$$= \widehat{C}_i \det\left(\widehat{M}_i^{-1} + \Sigma\right)^{-1/2} \exp\left(-\frac{1}{2}(\widehat{M}_i^{-1}\widehat{v}_i - \mu)^\top (\widehat{M}_i^{-1} + \Sigma)^{-1}(\widehat{M}_i^{-1}\widehat{v}_i - \mu)\right), \quad (2.2)$$

  where $\widehat{C}_i$ denotes an $\mathcal{F}_{i,T} \vee \sigma(\widehat{\vartheta}_1)$-measurable r.v. not depending on (μ, Σ, φ_i), and where

$$\widehat{M}_i := \widehat{\tau}_i\, h \sum_j \widehat{S}_{i,j-1}^{-1} a_{i,j-1}^{\otimes 2}, \qquad \widehat{v}_i := \sum_j \widehat{S}_{i,j-1}^{-1} a_{i,j-1} \Delta_j Y_i.$$

- Taking the product of (2.2) for $i \leq N$ and then taking the logarithm, we arrive at

$$\widehat{\vartheta}_2 = (\widehat{\varphi}_f, \widehat{\Sigma}_r) \in \operatorname{argmax} \mathbb{H}_2, \quad \mathbb{H}_2(\vartheta_2) := \sum_i \log \phi_{p_\varphi}\left(\widehat{M}_i^{-1}\widehat{v}_i;\, \mu,\, \widehat{M}_i^{-1} + \operatorname{diag}(O, \Sigma_r)\right).$$

1 Proposed model

2 Quasi-likelihood estimation

3 Applications to synthetic and real data

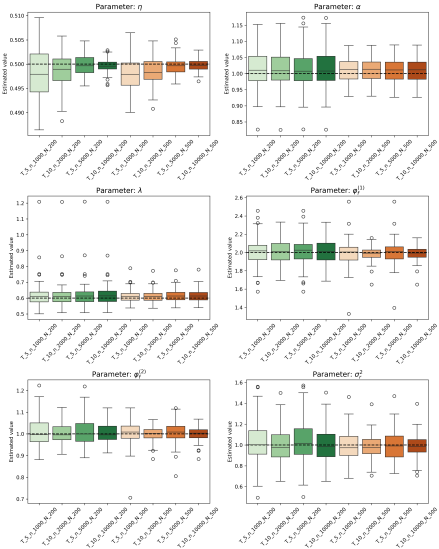
4 Concluding remarks

Simulation

$$Y_i(t) = Y_i(0) - \tau_i \Bigg(\int_0^t \frac{\varphi_f^{(1)} Y_i(s) + \varphi_f^{(2)}}{\sqrt{1 + Y_i(s)^2}} ds + \int_0^t \frac{\varphi_{r,i} Y_i(s)}{\sqrt{1 + Y_i(s)^2}} \Bigg) ds + \sqrt{\tau_i} \int_0^t \exp(\eta \operatorname{atan}(Y_i(s))) dw_i(s)$$

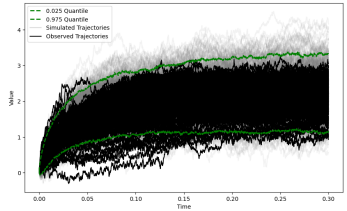
$$\tau_i \sim \text{Weibull}(\alpha, \lambda), \ \varphi_{r,i} \sim_{i.i.d.} N_1(0, \sigma_r^2),$$

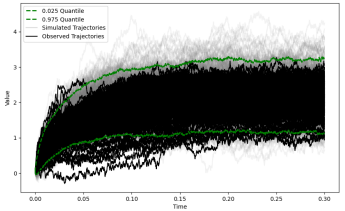
- $\eta = 0.5$, $\alpha = 1$, $\lambda = 0.6$, $\varphi_f^{(1)} = 2$, $\varphi_f^{(2)} = 1$, $\sigma_r^2 = 1$.
- The green (resp. orange) boxplots correspond to $N = 200$ (resp. $N = 500$).

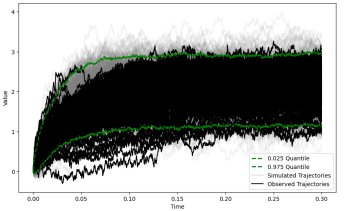


Neuronal data (membrane potentials) from the R package `mixedsde`

- 1000 simulated (gray) and observed (black) trajectories of membrane potentials. The green shaded area indicates the 95% prediction intervals for the simulated trajectories, providing a range within which future observations are expected to fall.







1 Proposed model

2 Quasi-likelihood estimation

3 Applications to synthetic and real data

4 Concluding remarks



Delattre, M. and Masuda, H. (2025),
Quasi-likelihood inference for SDE with mixed-effects observed at high frequency.
arxiv:2508.17910

Summary

- Observed data $\{Y_i(t_{ij}) : j = 0, 1, \ldots, n;\ i = 1, \ldots, N\}$, where $t_{ij} = jh$ with $h := T/n$ from

$$Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(Y_i(s))ds \right) + \sqrt{\tau_i} \int_0^t c(Y_i(s); \eta) dw_i(s),$$
$$\tau_i \sim f(\tau; \theta_\tau)d\tau, \qquad \varphi_{r,i} \sim N_{p_{\varphi_r}}(0, \Sigma_r).$$

- Profile and marginal quasi-likelihoods (**essential** for these simple forms):

$$\widehat{\eta} \in \operatorname*{argmax}_{\eta} \mathbb{H}_{11}(\eta), \quad \mathbb{H}_{11,N,n}(\eta) := -\frac{1}{2} \sum_i \left\{ \sum_j \log S_{i,j-1}(\eta) + n \log \left(\frac{1}{n} \sum_j S_{i,j-1}^{-1}(\eta)\, \overline{y}_{ij}^2 \right) \right\},$$
$$\widehat{\theta}_\tau \in \operatorname*{argmax}_{\theta_\tau} \mathbb{H}_{12}(\widehat{\eta}, \theta_\tau), \quad \mathbb{H}_{12,N,n}(\eta, \theta_\tau) := \sum_i \log f(\widehat{\tau}_i(\eta); \theta_\tau),$$
$$\widehat{\vartheta}_2 = (\widehat{\varphi}_f, \widehat{\Sigma}_r) \in \operatorname{argmax} \mathbb{H}_2, \quad \mathbb{H}_2(\vartheta_2) := \sum_i \log \phi_{p_\varphi} \left(\widehat{M}_i^{-1} \widehat{v}_i;\ \mu,\ \widehat{M}_i^{-1} + \operatorname{diag}(O, \Sigma_r) \right).$$

- $\begin{pmatrix} \sqrt{nN}(\widehat{\eta} - \eta_0) \\ \sqrt{N}(\widehat{\theta}_\tau - \theta_{\tau,0}) \end{pmatrix}$ and $\sqrt{N}(\widehat{\vartheta}_2 - \vartheta_{2,0})$ are asymp. normal under $n^{\mathsf{a}'} \lesssim N \lesssim n^{\mathsf{a}''}$ $(0 < \mathsf{a}' \leq \mathsf{a}'' < 1)$.

Agenda: Model selection issue

$$Y_i(t) = Y_i(0) + \tau_i \left(\varphi_f \cdot \int_0^t a_f(X_i(s), Y_i(s))ds + \varphi_{r,i} \cdot \int_0^t a_r(X_i(s), Y_i(s))ds \right)$$
$$+ \sqrt{\tau_i} \int_0^t c(X_i(s), Y_i(s); \eta)dw_i(s), \qquad i = 1, \ldots, N, \quad t \in [0, T].$$

- **Model-complexity and explanatory-variable/process selection**
- Sparse-type high-dimensional modeling through the **Spike-and-Slab type device**
- Ongoing with the French group at INRAE (Maud Delattre, Guillaume Kon Kam King)

//

References

Delattre, M. and Masuda, H. (2025),
Quasi-likelihood inference for SDE with mixed-effects observed at high frequency.
arxiv:2508.17910

Delattre, M., Genon-Catalot, V., and Larédo, C. (2018).
Approximate maximum likelihood estimation for stochastic differential equations with random effects in the drift and the diffusion.
Metrika, 81(8):953–983.

Ditlevsen, S. and Samson, A. (2013).
Introduction to stochastic models in biology.
Stochastic biomathematical models: with applications to neuronal modeling, pages 3–35.

Donnet, S. and Samson, A. (2013).
A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models.
Advanced drug delivery reviews, 65(7):929–939.

Fujinaga, Y. and Masuda, H. (2023).
Mixed-effects location-scale model based on generalized hyperbolic distribution.
Jpn. J. Stat. Data Sci., 6(2):669–704.

Narci, R., Delattre, M., Larédo, C., and Vergu, E. (2022).
Inference in gaussian state-space models with mixed effects for multiple epidemic dynamics.
Journal of Mathematical Biology, 85(4):40.

Picchini, U., Gaetano, A. d., and Ditlevsen, S. (2010).
Stochastic differential mixed-effects models.
Scandinavian Journal of Statistics, 37(1):67–90.