Bayesian Inference for Stochastic Differential Equation Mixed Effects Models
Imke Botha 1 Robert Kohn 2 Christopher Drovandi 1
1University of Queensland, 2University of New South Wales

Background
Stochastic differential equations (SDEs) are used to model the dynamics of complex systems over time or space. Adding random effects leads to a SDE mixed effects model (SDEMEM). Statistical inference for these models is difficult as the transition density (and therefore the likelihood) is often intractable. Adding random effects adds further complexity.

Previous methods aim to sample jointly over the space of the parameters and latent states (Whitaker et al., 2017; Picchini and Forman, 2019). However, this is very inefficient if there is high correlation between these two components. Particle MCMC (PMCMC, Andrieu et al. 2010) can be used to help overcome this, but only the simplest method in this class has been considered for SDEMEMs (Picchini and Forman, 2019).

Aim
Our aim is to develop advanced particle MCMC methods for SDEMEMs, which significantly improve the computational efficiency of exact Bayesian inference for these models.

Methods

Individual-Augmentation Pseudo-Marginal (IAPM)
In IAPM, we use an importance sampling estimate of the likelihood within a PMMH algorithm.

\[ \hat{P}(y_{1:M} \mid \theta) = \prod_{m=1}^{M} \frac{P(y_m \mid \theta)}{P(y_m \mid \hat{g}(\eta_m)} \cdot \hat{g}(\eta_m), \quad \eta_m \sim \hat{g}(\eta_m \mid \theta). \]

A naive choice of importance distribution is the prior. \( \hat{g}(\eta_m \mid \theta) = P(\eta_m \mid \theta) \).

We improve over the naive method by orders of magnitude. On this example CWPM updates the posterior samples in the following blocks: \( \eta_{1:M} \) and \( \phi_Y \).

Component-Wise Pseudo-Marginal (CWPM)
CWPM updates the parameters in the following blocks: \( \eta_{1:M}, \phi_Y, \phi_X \).CWPM is used to update \( \eta_{1:M} \) and \( \phi_Y, \phi_X \).

Log-likelihood results
We firstly consider the efficiency of the likelihood estimation. Results are shown in Table 1. For all methods (except the naive), we correlate the log-likelihood.

Table 1: Log-likelihood results for the IAPM and CWPM methods. The highlighted rows show the best combinations.

<table>
<thead>
<tr>
<th>Data</th>
<th>CWPM</th>
<th>IAPM</th>
<th>IAPM Lap-MDB</th>
<th>PMCMC Lap-MDB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (s)</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Real</td>
<td>50</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>Simulated</td>
<td>45</td>
<td>90</td>
<td>135</td>
<td>180</td>
</tr>
</tbody>
</table>

SDE Mixed Effects Model
The general form for a SDEMEM, with \( m = 1, \ldots, M \) individuals, is

\[ dX_{m,t} = \mu(X_{m,t-1}, \phi_X, \gamma_m)dt + \gamma_e e^{-\gamma_e} dBM_{m,t}, \]

where \( \gamma_m \sim P(\gamma_m) \). If this process is hidden, we obtain a state-space model with observation density \( P(y_{1:M} \mid x_{1:M}, \phi_Y) \).

The solution of (1) gives the transition density of the latent states. If an analytical solution is unavailable, numerical methods can be used, e.g. the Euler-Maruyama discretisation (EMD). The modified diffusion bridge (MDB) can also be used to simulate \( x_{m,t} \) conditional on \( y_{m,t} \).

Particle MCMC
PMCMC enables exact inference for intractable likelihood models. Two methods are:
- particle marginal Metropolis-Hastings (PMMH), which replaces the intractable likelihood with a particle filter (PF) estimate, and
- particle Gibbs (PG), which updates \( x_{1:T} \) using a conditional particle filter.

Mixed Particle Method (MPM)
MPM uses a combination of PMMH and PG to update the parameters and \( x_{1:M} \).

\[ P(\eta_{1:M}, \phi_Y, x_{1:M} \mid y) \propto \hat{P}(y_{1:M} \mid \eta_{1:M}, \phi_Y, x_{1:M}) P(\phi_Y) P(x_{1:M} \mid \eta_{1:M}, \phi_Y). \]

Variance Reduction
Mixing of the Markov chain can be very poor if the variance of the likelihood estimates is high. Generally, we want the variance of the log-ratio of likelihood estimates.

\[ \sigma^2 = \text{var} \left( \log \hat{P}(y_{1:T} \mid \theta) - \log \hat{P}(y_{1:T} \mid \theta) \right), \]

to be around 1. This may be achieved by increasing the number of particles (N) and random effects draws (L), using a more efficient proposal function and/or importance density, or correlating the log-likelihood estimates in (2) (Tran et al., 2016);

Naive method: We define the naive method as the uncorrelated IAPM algorithm with the prior as importance density and EMD as proposal function in the PF.

Example

Tumor Volume Data
We apply our methods to real data from a tumour xenography study on mice (Picchini and Forman, 2019). Figure 1 plots this data.

To fit the data, we consider an adaptation of a SDEMEM that was used by Picchini and Forman (2019) for unperturbed growth.

\[ \begin{align*}
Y_{m,t} &= X_{m,t} + \epsilon_{m,t}, \quad \epsilon_{m,t} \sim N(0, \sigma^2), \\
\frac{dX_{m,t}}{dt} &= \left( \beta + \frac{1}{2\tau} \left( 1 - e^{-\gamma e} \right) \right) dt + \gamma e^{-\gamma e} dBM_{m,t}, \\
X_{m,0} &\sim N\left( \mu_0, \sigma^2_{x0} \right), \\
(3,1,1,1,1,0.5,1)^T &\sim N\left( 3.1, -1, 1, 1, 1, 0.5, 1 \right)^T. 
\end{align*} \]

Log-likelihood results
We use the Laplace-MDB importance density (IAPM) and the MDB proposal function for both datasets.

Mixed Particle Method (MPM)
We improve over the naive method by orders of magnitude. On this example CWPM gave the best results, however the best method to use in any particular situation greatly depends on the model and data.

MCMC Results

![Figure 2: Univariate posterior density plots.](image-url)