SEQUENTIAL MONTE CARLO METHOD FOR PARAMETER ESTIMATION IN DIFFUSION MODELS OF AFFINITY-BASED BIOSENSORS

Manohar Shamaiah, Xiaohu Shen, and Haris Vikalo

Department of Electrical and Computer Engineering
University of Texas at Austin, Austin, TX, USA

ABSTRACT
Estimation of the amounts of target molecules in real-time affinity-based biosensors is studied. The problem is mapped to inferring the parameters of a temporally sampled diffusion process. To solve it, we rely on a sequential Monte Carlo algorithm which generates particles using transition density of the diffusion process. The transition density is not available in a closed form and is thus approximated using Hermite polynomial expansion. Simulations and experimental results demonstrate effectiveness of the proposed scheme, and show that it outperforms competing techniques.

Index Terms—parameter estimation, stochastic differential equation, real-time biosensors, sequential Monte Carlo

1. INTRODUCTION
Recently, massively parallel affinity-based biosensors have received a lot of attention [1]-[3]. Molecular binding, which enables sensing of target molecules in affinity-based biosensors, can be modeled by a continuous-time Markov process [4]. The states of the Markov process are discrete, but under some mild conditions its dynamics can be approximated by a stochastic differential equation (SDE)

\[ dn_c = \mu(n_c, n_t)dt + \sigma(n_c, n_t)dW, \]

where \( W \) denotes the Wiener process, \( n_t \) denotes the total amount of target molecules, \( n_c(t) \) denotes the amount of those that are bound to their corresponding probes at time \( t \), the drift and diffusion are given by

\[ \mu(n_c; n_t) = k_1 \frac{n_p - n_c}{n_p} (n_t - n_c) - k_{-1} n_c, \]

\[ \sigma(n_c; n_t) = \sqrt{k_1 \frac{n_p - n_c}{n_p} (n_t - n_c) + k_{-1} n_c}. \]

where \( n_p \) is the amount of probe molecules, \( k_1 \) denotes the forward reaction rate of the capturing process, and \( k_{-1} \) denotes the backward reaction rate.

In real-time affinity-based biosensors, data is acquired at discrete points in time \( t_i = i\Delta, i = 1, 2, \ldots, N \). The ultimate goal of an experiment is to estimate the amount of targets \( n_i \) in the biological sample. However, other parameters in (1), \( n_c, k_1, \) and \( k_{-1} \), are typically also unknown and need to be inferred. If the measurements are noise-free, it would suffice to form the log-likelihood function

\[ L = \sum_{i=1}^{N} \log p(n_c(t_i)|n_c(t_{i-1})), \]

and optimize it over the values of the unknown parameters. Clearly, this requires transition densities \( p(n_c(t_i)|n_c(t_{i-1})) \). In general, for the SDE (1) with drift and diffusion (2) and (3), the closed form expression for the transition density is not known and thus need to be approximated. The simplest such approximation is based on the so-called Euler-Maruyama discretization scheme which approximates the transition density by a Gaussian distribution. However, if \( \Delta \) is not sufficiently small, the Euler approximation is inadequate. In this case, one can introduce missing values between the observations, where the transition density between consecutive augmented data is Gaussian. The Euler approximation and data augmentations for parameter estimation in real-time biosensors were discussed in [5]. The performance using Euler approximation is poor when sampling interval is not small, while the data augmentation scheme is computationally very intensive.

Hermite polynomial expansion is an efficient alternative to the Euler-Maruyama approximation scheme. The main idea of this approximation is to expand an unknown transition density using Hermite polynomials as the bases. The approach consists of transforming the diffusion process into a related one with the transition density close to a Gaussian distribution, followed by the expansion of this transition density around the Gaussian distribution [6]. It has been shown that the Hermite poly-
nominal expansion provides good approximations of the true transition density (based on studies of models where the exact transition density is known [6]).

On another note, practical biosensor measurements are noisy and hence the likelihood ratio (4) cannot be computed explicitly. To this end, we rely on a sequential Monte Carlo (SMC) algorithm to evaluate the joint posterior density of the parameters and the hidden process $n_c(t)$, and use it to estimate the unknown parameters.

2. HERMITE POLYNOMIAL APPROXIMATION

Consider the general SDE model of the form

$$dX = \mu_X(X; \theta)dt + \sigma_X(X; \theta)dW.$$  

To find the Hermite polynomial expansion of the transition density $p(x|X_k; \theta)$, we first need to transform $X$ into a diffusion process with a transition density close to the Gaussian distribution. Start by introducing the diffusion

$$Y = \gamma(X) = \int_X^x \frac{du}{\sigma_X(u; \theta)};  \tag{5}$$

governed by the equation

$$dY = \mu_Y(Y; \theta)dt + dW;  \tag{6}$$

where

$$\mu_Y(Y; \theta) = \frac{\mu_X(X; \theta)}{\sigma_X(X; \theta)} - \frac{1}{2} \frac{\partial \sigma_X(X; \theta)}{\partial X};  \tag{7}$$

and $X = \gamma^{-1}(Y)$. Next, $Y$ is transformed to $Z = (Y - Y_k)/\sqrt{\Delta}$. Distribution $p(z|Y_k; \theta)$ can be approximated by keeping the first $J$ terms in its Hermite expansion,

$$p(z|Y_k; \theta) = \phi(z) \sum_{j=0}^J \eta_j(\Delta, Y_k; \theta)H_j(z),$$

where

$$H_j(z) = e^{-\frac{z^2}{2}} \frac{d^j}{dz^j}(e^{-\frac{z^2}{2}}), \quad \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}},$$

and

$$\eta_j(\Delta, Y_k; \theta) = \frac{1}{J!} E[H_j(z)|Y(t_k) = Y_k; \theta],$$

with $\eta_0(\Delta, Y_k; \theta) = 1$. Note that the Hermite polynomials are mutually orthogonal with respect to a normal distribution, i.e., $\int_{-\infty}^{\infty} \phi(z)H_m(z)H_n(z) = \sqrt{\pi}/\delta(n - m)$. Assume that $\eta_j(\Delta, Y_k; \theta)$ is $K$ times continuously differentiable with respect to $\Delta$ around the origin. Then it can be approximated by means of the MacLaurin expansion

$$\eta_j(\Delta, Y_k; \theta) = \frac{1}{J!} \sum_{i=0}^K \lim_{y \to Y_k} A_{ij}[H_j(z)] \frac{\Delta^i}{i!} + O(\Delta^{K+1}),$$

where

$$A_{ij}(\psi) = \psi^{ij} \frac{d^{2j}}{dy^{2j}} + \frac{1}{2} d^{2j+2}.$$  

Finally, the desired transition density is obtained as

$$p(x|X_k; \theta) = \frac{p(z|Y_k; \theta)}{\sqrt{\Delta} \sigma_X(X; \theta)}.$$  

On the other hand, a closed form approximation for the loglikelihood of the transition density is given by [7]

$$l_X(x|x_0; \theta) = -\frac{1}{2} \log(\sigma_X^2(x; \theta)) - \frac{1}{2} \log(2\pi \Delta) - \frac{1}{2\Delta} (\gamma(x; \theta) - \gamma(x_0; \theta))^2 + \sum_{k=0}^{\infty} C_Y^{(k)}(\gamma(x; \theta)) \gamma(x_0; \theta) \frac{\Delta^k}{k!},$$  

where

$$C_Y^{(0)}(y|y_0; \theta) = (y - y_0) \int_0^1 \mu_Y(y_0 + u(y - y_0); \theta) du,$$

$$C_Y^{(k)}(y|y_0; \theta) = k \int_0^1 C_Y^{(k-1)}(y_0 + u(y - y_0); \theta) u^{k-1} du, k \geq 1,$$

with the integrand function $G_Y^{(k)}(y|y_0; \theta)$ given by

$$G_Y^{(1)}(y|y_0; \theta) = \frac{1}{2} \left( \mu_Y(y; \theta) + \frac{\partial \mu_Y(y; \theta)}{\partial y} \right),$$  

and

$$G_Y^{(k)}(y|y_0; \theta) = \frac{1}{2} \left( \frac{\partial^2 C_Y^{(k-1)}(y|y_0; \theta)}{\partial y^2} \right) + \sum_{h=1}^{k-2} \left( \frac{(k-1)!}{h!(k-1-h)!} \frac{\partial G_Y^{(h)}(y|y_0; \theta)}{\partial y} \frac{\partial G_Y^{(k-1-h)}(y|y_0; \theta)}{\partial y} \right), k \geq 2.$$  

3. APPROXIMATING TRANSITION DENSITY OF MOLECULAR BINDING PROCESS

Following the steps outlined in Section 2, we here evaluate the Hermite polynomial expansion of the transition density of the process (1). For convenience, let us denote $X = n_c$ and $\theta = n_t$, so that drift (2) and diffusion (3) can be written as

$$\mu_X(X; \theta) = aX^2 - bX + c,$$

$$\sigma_X(X; \theta) = \sqrt{aX^2 - b_1X + c},$$

where $a = \frac{k_1}{n_p}$, $b = \frac{k_{1t}}{n_p} + k_1 + k_{-1}$, $c = k_{1t}$, $b_1 = \frac{k_{1t}n_t}{n_p} + k_1 - k_{-1}$. It is straightforward to show that

$$\gamma(X; \theta) = \begin{cases} 
-\frac{1}{\sqrt{a}} \cosh^{-1} \left( \sqrt{a} \frac{b_1 - X}{\sqrt{b_1^2 - X^2}} \right), & \text{if } b_1^2 > 4ac \\
-\frac{1}{\sqrt{a}} \sinh^{-1} \left( \sqrt{a} \frac{b_1 - X}{\sqrt{b_1^2 - X^2}} \right), & \text{if } b_1^2 < 4ac \\
\frac{1}{\sqrt{a}} \log \left( X - \frac{b_1}{2a} \right), & \text{if } b_1^2 = 4ac 
\end{cases},$$  

(9)
and that \( \gamma^{-1}(Y; \theta) = \)
\[
\begin{cases}
\frac{b_1}{2a} - \frac{1}{\sqrt{a}} \sqrt{\frac{b_1^2}{4a} - c} \cosh \left( \sqrt{a}Y \right), & \text{if } b_1^2 > 4ac \\
\frac{b_1}{2a} + \frac{1}{\sqrt{a}} \sqrt{c - \frac{b_1^2}{4a}} \sinh \left( \sqrt{a}Y \right), & \text{if } b_1^2 < 4ac \\
\frac{b_1}{2a} + \exp \left( \sqrt{a}Y \right), & \text{if } b_1^2 = 4ac
\end{cases}
\] (10)

Note that while deriving (9) from (5), we obtain three different integrals based on the discriminant of the quadratic form \( ax^2 - b_1x + c \). Moreover, note that the parameters should be such that the quadratic form is positive (and so is the integral (5)).

By substituting (9) and (10) in (7), we obtain expressions for \( \mu_Y(Y; \theta) \). The resulting expression is used to derive coefficients \( \mathcal{C}^{(k)}_{Y}(y; y_0; \theta) \). These expressions are somewhat cumbersome to write and are omitted for brevity of the presentation. It suffices to say that we truncate the series in (8) to \( k = 1 \) and hence only need to determine \( \mathcal{C}^{(0)}_{Y}(y; y_0; \theta) \) and \( \mathcal{C}^{(1)}_{Y}(y; y_0; \theta) \). We evaluate these expressions using Mathematica.

4. NOISY OBSERVATIONS CASE: SEQUENTIAL MONTE CARLO

Measurements in biosensors are corrupted by noise and thus we need to estimate both the parameters and the hidden states \( X \). We adopt the kernel smoothing based sequential Monte Carlo algorithm with resampling step [8]. The proposal distribution used to generate samples relies on the modified diffusion bridge construct that employs the Hermite polynomial approximation of the transition density derived in Section 3. The static parameters are modeled using an artificial evolution. This approach is sometimes referred to as kernel smoothing [9].

In a general SDE model with state transition density \( p(x_{t+1}|x_t, \theta) \) and observation dynamics described by \( p(y_{t+1}|x_{t+1}, \theta) \), the joint posteriori distribution of the parameters and the states is given by \( p(x_{t+1}, \theta|y_{1:t+1}) \propto p(y_{t+1}|x_{t+1}, \theta)p(x_{t+1}|\theta, y_{1:t})p(\theta|y_{1:t}) \). SMC method represents this joint distribution at time \( t \) by the samples \( \{x^{(i)}_t, \theta^{(i)}_t : i = 1, \ldots, N\} \) and their associated weights \( \{w^{(i)}_t : i = 1, \ldots, N\} \). In the kernel smoothing approach, \( p(\theta|y_{1:t}) \) is given by

\[
p(\theta|y_{1:t}) \approx \sum_{j=1}^{N} w^{(j)}_t \mathcal{N}(m^{(j)}_t, h^2 V_t),
\] (11)

where \( \mathcal{N}(m, V) \) denotes the multivariate normal density with mean \( m \) and variance \( V \). This is a Gaussian mixture with sample weights \( w^{(j)}_t \) which adds random jitter to the particles. The mean and variance of the Gaussian kernel are set to \( m^{(j)}_t = \lambda h^{(j)} + (1 - \lambda) \bar{\theta}_t \) and \( h^2 V_t \), respectively, where \( \bar{\theta}_t \) denotes the previous (prior) mean estimate and \( V_t \) is its corresponding variance. The tuning parameters \( \lambda \) and \( h \) are given by \( \lambda = \frac{0.05}{2\sigma} \), \( h = \sqrt{1 - \lambda^2} \), where \( \delta \) is a discount factor typically chosen as 0.95 – 0.99. With this choice of parameters, the Gaussian mixture in (11) retains mean \( \bar{\theta}_t \) and has the correct variance \( V_t \) [9].

4.1. Proposal density

**Euler proposal density:** As a reference, first recall the proposal density based on the Euler approximation,

\[
\pi_{t|t-1}(X_t|X_{t-1}, \theta, Y_t) = p(X_t|X_{t-1}) \sim \mathcal{N}(\mu_t, \sigma_t^2)
\]

where the mean and variance are given by \( \mu_t = X_{t-1} + \mu_X(X_{t-1}; \theta) \Delta \) and \( \sigma_t^2 = \sigma_X(X_{t-1}; \theta) \Delta \).

**Modified diffusion construct:** Diffusion bridge construct incorporates noisy observations \( Y_t \) in the proposal density. In particular,

\[
\pi_{t|t-1}(X_t|X_{t-1}, \theta, Y_t) \sim \mathcal{N}(\mu_t, \sigma_t^2)\]

(12)

where

\[
\mu_t = \text{X}_{t-1} + \left( \mu_t + \frac{\sigma_t^2}{\sigma_t^2 + \sigma_Y^2} (Y_t - \text{X}_{t-1} - \mu_k \Delta \mu_k) \right) \Delta,
\]

\[
\sigma_t^2 = \sigma_Y^2 - \frac{\sigma_t^2}{\sigma_t^2 + \sigma_Y^2} \sigma_t^2 \Delta,
\]

\( \sigma_Y^2 \) denotes the observation noise variance, \( \mu_t = X_{t-1} + \mu_X(X_{t-1}; \theta) \Delta \) and \( \sigma_t^2 = \sigma_X(X_{t-1}; \theta) \Delta \).

On the other hand, the weights \( w_t^{(j)} \) are given by

\[
w_t^{(j)} = p(y_t|x_t^{(j)}) \frac{p(x_t|x_{t-1}, \theta_{t-1})}{\pi_{t|t-1}(x_t|y_{1:t}, \theta_{t-1})}.
\]

Here \( p(x_t|x_{t-1}, \theta_{t-1}) \) is obtained via the Hermite polynomial approximation described in Section 3. The above weights are normalized and resampling is performed based on the normalized weights. The SMC algorithm employing the above steps is formalized below.

1. **Initialization** Set \( t = 1 \). Draw \( \{x_1^{(i)}, \theta_1^{(i)} : i = 1, \ldots, N\} \) using observation \( Y_1 \) as the prior for \( \{x_1^{(i)}\} \) and a prior for \( \{\theta^{(i)}\} \).

2. **Iterations** For \( t \geq 2 \):

(i) For \( i = 1, \ldots, N_p \): Draw samples from the proposal density \( \pi_{t|t-1}(X_t|X_{t-1}, \theta, Y_t) \).

(ii) Compute \( p(x_t|x_{t-1}, \theta_{t-1}) \) using an approximate transition density (12).

(iii) Compute weights \( w_t^{(i)} \),

\[
w_t^{(i)} = \frac{p(Y_t|x_t^{(i)})}{\pi_{t|t-1}(x_t^{(i)}|y_{t-1}, \theta_{t-1})}.
\]
In this paper, we presented a sequential Monte Carlo technique for estimating parameters of a diffusion process modeling real-time affinity-based biosensors. We employed Hermite polynomial expansion to approximate generally unavailable transition density of the molecular binding process. Simulation studies indicate that the Hermite approximation based technique provides significantly better accuracy compared to Euler-Maruyama method, especially for large sampling times Δ. The effectiveness of the algorithm is also demonstrated on experimental data.

7. REFERENCES


