Preprints of the 2012 20th Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, July 3-6, 2012

Particle filtering for the chemostat

Boumédiène Benyahia Fabien Campillo Brahim Cherki Jérôme Harmand

***Abstract*— We develop a particle filter approximation of the optimal nonlinear filter in the context of the chemostat. We propose a stochastic model of the chemostat together with an observation model. One of the characteristics of applications in bioprocesses is that the time between two observations is relatively large. We account for this point in the development of the particle filter by refining the prediction step of the particle filter. We present numerical tests on simulated measurements.**

I. Introduction

The control of bioprocesses are usually based on a state-space model and prior to the control phase itself, it is necessary to develop a procedure to identify the parameters of the model and the state variables [4]. First, the very structure of the state-space model is typically a drastic simplification of reality. Then, these variables are either not directly observed or observed but subject to measurement noise; observations are usually heterogeneous, asynchronous, acquired at low frequencies, but also partial (only some components are observed) and subject to high intensity noise. Finally, the initial conditions of the state process are also subject to errors. Hence the identification procedure must be robust with respect to all these aspects.

One can distinguish between deterministic and stochastic approaches. The former mainly includes observers [5]. Classical observers usually require a precise knowledge of the underlying system. Nevertheless ad hoc methods exist in specific cases: for example the asymptotic observer [14] for the chemostat remarkably does not require the knowledge of the growth function. Still observer approaches lack robustness in the presence of modeling approximations and uncertainties. These deterministic approaches focus on the reconstruction of unobserved components of the state vector and do not account for any estimation error, they apprehend the potential randomness of the model or of the observation only in terms robustness.

Stochastic methods propose an estimate of both the state vector and the error covariance, they are generally less demanding in terms of system properties; for example they do not require observability properties. In contrast with observers they easily handle measurements in discrete time, asynchronous, noisy and at low frequency. The most classical algorithm for nonlinear systems, is the extended Kalman filter (EKF), this algorithm does not generally bring about a substantial improvement compared to observers: it suffers the same lack of robustness, it gives relevant results only when the system is close to the linear/Gaussian case, possibly with high noise intensities, or when the state and observation noise intensities are low. Just like observers, it can also give good results when we can separate the "well observed" state variables, possibly "nonlinear", from the "weakly observed" but "highly linear" variables. The robustness issues have been significantly improved by unscented Kalman filters (UKF) and ensemble Kalman filters (EnKF): these approaches are clearly more robust to nonlinearities and high noise intensities [25], [24]. But as the UKF/EnKF output is still Gaussian, that is a mean estimate of the state variable coupled with an error covariance, it may have difficulties in addressing the full nonlinear/non-Gaussian case.

More recently particle filtering techniques (also called SMC for sequential Monte Carlo), which have been developed and successfully applied in target tracking and robotics [16], were also applied in the field of bioprocess [18], [27]. These techniques, and more generally numerical Bayesian methods, are almost the only ones that can fully account for the nonlinear/non-Gaussian issue. Another possibility is to numerically solve the partial differential equation of the nonlinear filtering, this approach is reliable only in small state dimension [13].

Suppose we have measurements:

$$Y_n = h_n(\theta, z_n, v_n) \tag{1}$$

at times t_n ($0 = t_0 < t_1 < t_2 \cdots$), where h_n is a known function, v_n is the measurement noise i.e. a sequence of independent and identically distributed (i.i.d.) random variables (non necessarily Gaussian), θ is an unknown parameter; z_n is the state of the system at the instant of observation t_n. It is assumed that the state process z_n is a Markov chain with transition kernel $Q^\theta_{n-1}(z, z')$ that could possibly depend on n and θ, i.e. $z' \mapsto Q^\theta_{n-1}(z_{n-1}, z')$ is the probability density function (p.d.f) of z_n, in other words z_n is drawn accordingly to $Q^\theta_{n-1}(z_{n-1}, \cdot)$ denoted as:

$$z_n \sim Q^\theta_{n-1}(z_{n-1}, \cdot)\,. \tag{2}$$

We also consider the initial state p.d.f.:

$$z_0 \sim p^\theta_0(\cdot)\,. \tag{3}$$

In many cases $z_n = Z_{t_n}$ where Z_t is solution of a stochastic

Boumédiène Benyahia and Brahim Cherki is with the LAT, University of Tlemcen, Algeria and the MODEMIC INRIA/INRA project-team
benyahia@supagro.inra.fr

This work was supported by the TREASURE research network within the framework of the 3+3 INRIA Euromed program

Fabien Campillo is with the MODEMIC INRIA/INRA project-team, Montpellier, France
Fabien.Campillo@inria.fr
Jérôme Harmand is with the LBE/INRA, Narbonne, France and the MODEMIC INRIA/INRA project-team
Jerome.Harmand@supagro.inra.fr

978-1-4673-2529-5/12/$31.00 ©2012 IEEE

differential equation (SDE):

$$\mathrm{d}Z_t = f(\theta, Z_t)\,\mathrm{d}t + \sigma(\theta, Z_t)\,\mathrm{d}w_t \tag{4}$$

where w_t is a standard Brownian motion (w_t, v_n and Z_0 are independent).

The problem of filtering and identification is to estimate the state z_n given the past observations $\{Y_\ell; \ell \leq n\}$; certain components of the parameter θ have to be estimated because they have special meaning in the problem under consideration, usually all components are estimated. The filtering problem is essentially Bayesian: given prior information on the state process in the form of a model (2) (3) which gives the law of the state process for each given value of the parameter θ, the observation model (1) allows us to determine, at least theoretically, the law of the state process given the observations for each given value of the parameter θ, see Section III.

The parameter θ can be treated in a Bayesian way, by prescribing a prior distribution, or in frequentist way by a maximum likelihood approach. In the present study *we assume that the parameter θ is known*, which is obviously not realistic, and therefor we will focus on the filtering issue. We will revisit this question and propose ways to address this issue in Section VI.

In numerous studies, the initial model is deterministic:

$$\dot{Z} = f(Z)\,, \qquad Y_n = h_n(Z_{t_n}) \tag{5}$$

and authors confine themselves to artificially adding noises to the previous equation by considering:

$$\mathrm{d}Z_t = f(Z_t)\,\mathrm{d}t + \mathrm{d}w_t\,, \tag{6a}$$

$$Y_n = h_n(Z_{t_n}) + v_n\,. \tag{6b}$$

This idea is good when the intensity of the state noise w_t and observation noise v_n are small, but in this case most methods (observers, EKF, UKF, particle filtering) give good results.

The state noise intensity is not small when the real system has a significant random component, this is the case for population dynamics in relatively small population size where the randomness is not eliminated by the law of large numbers. This noise intensity is also large when the model is a relatively crude approximation of reality, the noise can include this modeling error. In both cases it is necessary to develop a proper covariance structure, more elaborate than $\sigma(\,\cdot\,) \equiv 1$. Also, suppose the state equation $\dot{Z} = f(Z)$ is a population dynamics that leaves all the components of the state vector Z positive, then $\mathrm{d}Z_t = f(Z_t)\,\mathrm{d}t + \mathrm{d}w_t$ no longer respects this property, but models like (4) may respect that property. Therefore, the form of the function $\sigma(\,\cdot\,)$ strongly influences the behavior of the process and the stochastic part of dynamics (4) must be developed with care.

Furthermore, in bioprocess applications the intensity of the noise in the observation equation (6b) is large and again the covariance structure should be carefully investigated. This issue is based on the development of models for sensors and equation (6b) is a rough approximation, but sometimes sufficient enough, of reality in many cases. However, we cannot assume that the intensity of the observation noise is small.

The aim of this paper is to describe the modeling approach and the derivation of the particle filter with a special focus on the issue of low frequency observations. For the sake of simplicity we will limit ourselves to the case of the simple chemostat.

In Section II we present the state-space model relying on a SDE model developed in [8] where we propose a more appropriate covariance structure. In Section III we derive the equations of the optimal filter. In Section IV we develop the particle filter; numerical tests are presented in Section V.

II. The setup

A. State equation

We consider the stochastic chemostat model introduced in [8] in the form of a stochastic process $Z_t = (X_t, S_t)$ solution of the following two dimensional stochastic differential equation (SDE):

$$\mathrm{d}X_t = (\mu(S_t) - D)\,X_t\,\mathrm{d}t + c_1\,\sqrt{X_t}\,\mathrm{d}W_t^1\,, \tag{7a}$$

$$\mathrm{d}S_t = -k\,\mu(S_t)\,X_t\,\mathrm{d}t + D\,(S_{\text{in}} - S_t)\,\mathrm{d}t + c_2\,\sqrt{S_t}\,\mathrm{d}W_t^2 \tag{7b}$$

for $t \in [0, T]$; X_t and S_t are the concentrations (g/l) of biomass and substrate at time t; W_t^1 and W_t^2 are independent scalar standard Brownian motions, also independent from Z_0.

We suppose that the specific growth function is of Monod type:

$$\mu(s) = \frac{\mu_{\text{max}}\,s}{k_s + s}\,.$$

The inputs are the dilution rate D (1/h) and the input substrate concentration S_{in} (mg/l); the model parameters are the noise intensities c_1 and c_2; the kinetics parameters are the yield coefficient k, the maximum growth rate μ_{max} (1/h) and half-saturation k_s (mg/l).

The initial distribution law of the initial condition $Z_0 = (X_0, S_0)$ is denoted $p_0(z) = p_0(x, s) = p_{Z_0}(z) = p_{X_0,S_0}(x, s)$. We suppose $X_0 \geq 0$, $S_0 \geq 0$, so that $X_t \geq 0$, $S_t \geq 0$ for all $t \geq 0$.

We will use the notations:

$$f_1(x, s) = (\mu(s) - D)\,x\,,$$
$$f_2(x, s) = -k\,\mu(s)\,x + D\,(S_{\text{in}} - s)\,.$$

If the deterministic part (drift terms) of Equation (7), i.e. the classical chemostat deterministic model, is well known [29], some comments will be made later concerning the stochastic part (diffusion terms).

The model (7) is derived in [8] as a diffusion approximation of a pure jump model described at microscopic scale. This model accounts for the demographic randomness and, as discussed in Section I, it induces a sharper prior on the state process compared to the same model with constant covariance terms. The solution of [8] is positive and features interesting washout properties. A similar approach is adopted

in [19], other covariance structures are possible [21], see discussion in [8].

B. Observation equation

We suppose that we only observe the substrate concentration S_{t_n} at equally spaced discrete measurement times $t_n = n\,\Delta$; we also suppose that the standard deviation of the noise is proportional to the substrate concentration. The measurement equation is:

$$Y_n = S_{t_n} + \sigma\, S_{t_n}\, v_n \tag{8}$$

where $v_n \overset{\text{iid}}{\sim} N(0,1)$ (i.i.d. independent and identically distributed). The state noises W_t^i, the observation noise v_n and the initial condition (X_0, B_0) are supposed mutually independent. The observation parameter is the noise intensity σ.

The observation model (8) may seem surprising since the noise intensity decreases, and vanishes, with the substrate concentration. But this model is more relevant than $Y_n = S_{t_n} + \sigma\, v_n$ as, in practice, the error in the measurement of the substrate concentration becomes negligible (resp. increases) when this concentration tends toward 0 (resp. increases). The growth law $(\sigma\, S_{t_n})^2$ of the variance is of course debatable, but we want especially to emphasize the importance of developing measurement models with covariance structures that are not limited to the classical form $Y_n = S_{t_n} + \sigma\, v_n$ resulting from conventional linear/Gaussian approaches.

We define $\Psi(y,z)$ as the *likelihood function* associated to measurement Y_n defined in (8):

$$\Psi(y,z) \overset{\text{def}}{=} \exp\Big(-\frac{1}{2\,\sigma^2\,s^2}\,|y-s|^2\Big) \tag{9}$$

that is the conditional p.d.f. of $Y_n = y$ given that $Z_t = z = (x,s)$.

III. The optimal filter

At time t_n, we have measurements $y_0, \ldots, y_n$ (denoted $y_{0:n}$) that are realizations of $Y_0, \ldots, Y_n$ (denoted $Y_{0:n}$). Based on these measurements, the nonlinear filtering problem is to give an estimation of the current unobserved component X_{t_n} of the state process together with an estimation of the associated error variance; and also to improve the estimation available of the component S_{t_n}. This could be done explicitly with the help of a nonlinear filter.

In this preliminary work *we suppose that all the model inputs, the model parameters, the kinetics parameters and the observation parameter are known.* This hypothesis is naturally unrealistic and will be commented upon later.

A. Definition

The problem is to compute the conditional p.d.f. $\pi_n(z) = \pi_n(x,s)$ of Z_{t_n} given that $Y_{0:n} = y_{0:n}$:

$$\pi_n(z) \overset{\text{def}}{=} p_{Z_{t_n}|Y_{0:n}}(z|y_{0:n})\,. \tag{10}$$

The conditional p.d.f. π_n is also called the nonlinear filter. In the following we will also need the predicted filter defined as:

$$\pi_{n^-}(z) \overset{\text{def}}{=} p_{Z_{t_n}|Y_{0:n-1}}(z|y_{0:n-1})\,.$$

The filter π_n and the predicted filter π_{n^-} gather all information available on Z_{t_n} based on the observations $Y_{0:n-1} = y_{0:n}$ and $Y_{0:n} = y_{0:n-1}$ respectively. Indeed, for any bounded function $f : \mathbb{R}_+^2 \mapsto \mathbb{R}$, the estimation $\widehat{f(Z_{t_n})}$ of $f(Z_{t_n})$ based on the observations $y_{0:n}$ is defined by:

$$\widehat{f(Z_{t_n})} \overset{\text{def}}{=} \mathbb{E}[f(Z_{t_n})|Y_{0:n} = y_{0:n}] = \int_{\mathbb{R}_+^2} f(z)\,\pi_n(z)\,\mathrm{d}z$$

it is optimal according to the mean square error criterion, that is:

$$\mathbb{E}\big[\big|\widehat{f(Z_{t_n})} - f(Z_{t_n})\big|^2\big] \le \mathbb{E}\big[\big|\phi(y_{0:n}) - f(Z_{t_n})\big|^2\big]$$

for any function $\phi(y_{0:n})$ of $y_{0:n}$.

For example, the optimal estimation of Z_{t_n} based on the observations $y_{0:n}$ is the mean:

$$\hat{Z}_{t_n} = \int_{\mathbb{R}_+^2} z\,\pi_n(z)\,\mathrm{d}z$$

of the p.d.f. $\pi_n(z)$, and the associated conditional variance is

$$\int_{\mathbb{R}_+^2} |z - \hat{Z}_{t_n}|^2\,\pi_n(z)\,\mathrm{d}z\,.$$

B. The optimal filter

The computation of the optimal filter can be achieved sequentially, the iteration $\pi_{n-1} \to \pi_n$ is done in two classic steps:

- *Prediction step:* We compute π_{n^-}:

$$\pi_{n^-}(z') = \int_{\mathbb{R}_+^2} Q_\Delta(z,z')\,\pi_{n-1}(z)\,\mathrm{d}z \tag{11}$$

  where $Q_\Delta(z,z')$ is the *transition kernel* of the state equation (7) defined by:

$$Q_\Delta(z,z') \overset{\text{def}}{=} p_{Z_{t+\Delta}|Z_t}(z'|z)\,. \tag{12}$$

  that is the conditional p.d.f. of $Z_{t+\Delta}$ given that $Z_t = z$.
- *Correction step:* Using the new observation $Y_n = y_n$, we compute π_n:

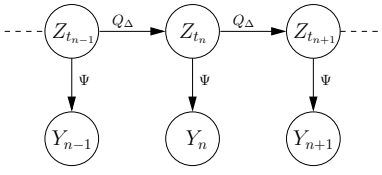
$$\pi_n(z) = \frac{\Psi(y_n,z)\,\pi_{n^-}(z)}{\int_{\mathbb{R}_+^2}\Psi(y_n,z')\,\pi_{n^-}(z')\,\mathrm{d}z'} \tag{13}$$

  where $\Psi(y,z)$ is the likelihood function (9).

The initialization at time $t = 0$ of the iterations is

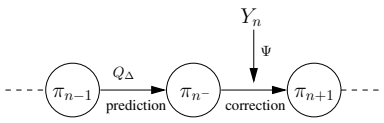
$$\pi_{0^-}(z) = \pi^0(z)\,, \quad \pi_0(z) = \frac{\Psi(y_0,z)\,\pi_{0^-}(z)}{\int_{\mathbb{R}_+^2}\Psi(y_0,z')\,\pi_{0^-}(z')\,\mathrm{d}z'}\,.$$

In conclusion, the dynamics of the nonlinear filter relies on the Markovian structure of the state equation (7) and on the so-called "memoryless channel hypothesis" of the observation equation (8); which can be represented by the following diagram:



meaning that the evolution of the state process at step n depends only on its value at step $n-1$ and that the observation Y_n depends only on the state process at step n.

The dynamics of the nonlinear filter can be represented by the following diagram:



(14)

IV. The particle filter

Except in the linear/Gaussian case and some very particular cases, this optimal filter (11) and (13) cannot be solved explicitly [3]. This is why specific approximation techniques have been developed. A first method is the Extended Kalman Filter, another one is the particle filer. We present the latter now.

We introduce the simplest implementation of the particle filter, namely the *sequential importance sampling with resampling* (SISR) also called bootstrap filter [16]. The idea is to obtain an empirical approximation of π_n of the form:

$$\pi_n(z) \simeq \pi_n^N(z) \stackrel{\text{def}}{=} \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_n^i}(z)$$

where $\xi_n^1, \ldots, \xi_n^N$ are the particles, $\delta_\xi(z)$ is the Dirac measure centered on ξ. This "particle" representation of the filter allows for simple computations of the integral terms of the prediction and correction steps of the optimal filter.

The particle filter is a sequential Monte Carlo method where π_n^N is computed from π_{n-1}^N and the new observation $Y_n = y_n$.

Starting from the particles $(\xi_{n-1}^i)_{i=1\cdots N}$, the new particles $(\xi_n^i)_{i=1\cdots N}$ are computed in two steps: the prediction (or mutation step) and the correction (or selection step).

Prediction (mutation)

(x, s) given
$\delta_{\text{filter}} \leftarrow \Delta / M_{\text{filter}}$
for $m = 1 : M_{\text{filter}}$ **do**
  $w_1 \sim N(0,1)$, $w_2 \sim N(0,1)$
  $x' \leftarrow \max(0, x + f_1(x,s)\,\delta_{\text{filter}} + c_1\sqrt{x}\,\sqrt{\delta_{\text{filter}}}\,w_1)$
  $s' \leftarrow \max(0, s + f_2(x,s)\,\delta_{\text{filter}} + c_2\sqrt{s}\,\sqrt{\delta_{\text{filter}}}\,w_2)$
  $x \leftarrow x'$, $s \leftarrow s'$
end for

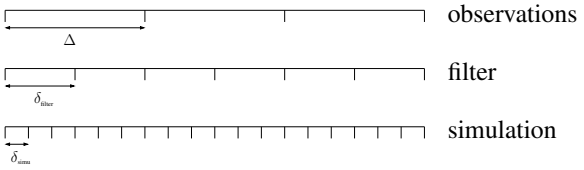
Algorithm 1: *Simulation of Equation* (7) *with an Euler-Maruyama scheme leading to the approximation* $\tilde{Q}_\Delta$ *of the transition kernel* Q_Δ *given by* (12)*.*

The kernel Q_Δ and the prediction step (11) are complex so we replace them by their empirical counterparts. First we have to simulate predicted particles $\tilde{\xi}_n^i \sim \tilde{Q}_\Delta(\xi_{n-1}^i, \cdot)$, but the transition kernel Q_Δ is not explicit; we approximate it using an Euler-Maruyama approximation of (7), that is:

$$\tilde{X}_{\tilde{t}_m} = \Big[\tilde{X}_{\tilde{t}_{m-1}} + f_1(\tilde{X}_{\tilde{t}_{m-1}}, \tilde{S}_{\tilde{t}_{m-1}})\,\delta_{\text{filter}} + c_1\sqrt{\tilde{X}_{\tilde{t}_{m-1}}}\,\sqrt{\delta_{\text{filter}}}\,w_m^1\Big]_+ \quad (15a)$$

$$\tilde{S}_{\tilde{t}_m} = \Big[\tilde{S}_{\tilde{t}_{m-1}} + f_2(\tilde{X}_{\tilde{t}_{m-1}}, \tilde{S}_{\tilde{t}_{m-1}})\,\delta_{\text{filter}} + c_2\sqrt{\tilde{S}_{\tilde{t}_{m-1}}}\,\sqrt{\delta_{\text{filter}}}\,w_m^2\Big]_+ \quad (15b)$$

where w_m^1 and w_m^2 are $N(0,1)$ independent random variables. Here we use a time step $\delta_{\text{filter}} = \Delta / M_{\text{filter}}$ and $\tilde{t}_m = m\,\delta_{\text{filter}}$; indeed, for the filter time step δ_{filter} we subdivide the observation time step Δ:



also for the simulation time step δ_{simu}, which will be used in Section V, we subdivide the filter time step δ_{filter}.

The positive part operator $[\,\cdot\,]_+$ is used to ensure the positivity of the solution.

Starting from $(\tilde{X}_0, \tilde{S}_0) = \xi_{n-1}^i$, we simulate M_{filter} time iterations of (15) to get the predicted particle:

$$\tilde{\xi}_n^i = (\tilde{X}_{\tilde{t}_{M_{\text{filter}}}}, \tilde{S}_{\tilde{t}_{M_{\text{filter}}}}) = (\tilde{X}_\Delta, \tilde{S}_\Delta)\,.$$

We denote this step by:

$$\tilde{\xi}_n^i \sim \tilde{Q}_\Delta(\xi_{n-1}^i, \cdot\,)\,.$$

corresponding to the Algorithm 1.

Correction (selection)

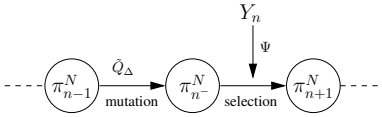
In a second step a likelihood weight ω_n^i is associated with each predicted particle $\tilde{\xi}_n^i$. This weight:

$$\omega_n^i \propto \Psi(y_n, \tilde{\xi}_n^i) \quad (16)$$

is proportional to the likelihood of the last observation and $\sum_{i=1}^N \omega_n^i = 1$. Then these particles are selected accordingly to these weights, i.e. sampled from the empirical distribution $\sum_{i=1}^N \omega^i\,\delta_{\tilde{\xi}_n^i}$ independently one from the other:

$$\xi_n^1, \ldots, \xi_n^N \stackrel{\text{iid}}{\sim} \sum_{i=1}^{N} \omega_n^i\,\delta_{\tilde{\xi}_n^i}\,. \quad (17)$$

See Algorithm 2 for a complete description of the particle filter. The dynamics of the particle filter can be represented by the following diagram:



which is equivalent to the diagram of the nonlinear filter (14)

{initialization}
$\tilde{\xi}^1, \dots, \tilde{\xi}^N \overset{\text{iid}}{\sim} \text{law}(Z_0)$
$\omega^i \leftarrow \Psi(y_0, \tilde{\xi}^i)$ for $i = 1 : N$ {likelihood weights}
$\omega^i \leftarrow \omega^i / \sum_{i'=1}^N \omega^{i'}$ for $i = 1 : N$ {normalization}
$\xi^1, \dots, \xi^N \overset{\text{iid}}{\sim} \sum_{i=1}^N \omega^i \, \delta_{\xi^i}$ {selection}
save $(0, (\xi^1, \dots, \xi^N))$

{iterations}
for $n = 1, 2, \dots$ **do**
$\tilde{\xi}^i \sim \tilde{Q}_\Delta(\xi^i, \mathrm{d}z')$ for $i = 1 : N$ {mutation}
$\omega^i \leftarrow \Psi(y_n, \tilde{\xi}^i)$ for $i = 1 : N$ {likelihood weights}
$\omega^i \leftarrow \omega^i / \sum_{i'=1}^N \omega^{i'}$ for $i = 1 : N$ {normalization}
$\xi^1, \dots, \xi^N \overset{\text{iid}}{\sim} \sum_{i=1}^N \omega^i \, \delta_{\xi^i}$ {selection}
save $(n\,\Delta, (\xi^1, \dots, \xi^N))$
end for

Algorithm 2: *Bootstrap particle filter: allows to compute an approximation of π_{n^-} of the form $\frac{1}{N}\sum_{i=1}^N \delta_{\tilde{\xi}^i}(\mathrm{d}z)$ and an approximation of $\pi_n(\mathrm{d}z)$ of the form $\frac{1}{N}\sum_{i=1}^N \delta_{\xi^i}(\mathrm{d}z)$. In the mutation step, predicted particles $\tilde{\xi}^i$ are sampled independently of one another.*

Outputs of the filter

One of the advantages of the particle approximation approach is that the estimation of the state process is straightforward. Indeed, for any function $f : \mathbb{R}_+^2 \mapsto \mathbb{R}$ the optimal estimation $\mathbb{E}[f(Z_{t_n})|Y_{0:n} = y_{0:n}]$ of $f(Z_{t_n})$ is simply approximated by:

$$\widehat{f(Z_{t_n})}^N = \int_{\mathbb{R}_+^2} f(z)\, \pi_n^N(z)\, \mathrm{d}z = \frac{1}{N}\sum_{i=1}^N f(\xi_i^N)\,.$$

For example, the estimation of Z_{t_n} is $\hat{Z}_{t_n}^N = \frac{1}{N}\sum_{i=1}^N \xi_i^N$. Also the probability for the state process Z_{t_n} to be in a given subset D of $\mathbb{R}_+^2$ is simply the proportion of particles in the domain D:

$$\mathbb{P}[Z_{t_n} \in D | Y_{0:n} = y_{0:n}] \simeq \frac{1}{N}\sum_{i=1}^N 1_D(\xi_n^i)\,.$$

About the resampling step (17)

In particle filtering, the resampling step (17) deserves special attention; indeed it can be very time-consuming and can also affect the output of the filter in terms of variance. Several algorithms are available [15], for our simulations we chose the residual resampling. The measurement frequency is usually low in bioprocess applications, and the resampling (17) is required only at the time of measurement.

Comments

Concerning the noise intensities c_1, c_2 and σ: first the stochastic model (7) and more specifically the diffusion terms $c_1\sqrt{X_t}$ and $c_2\sqrt{S_t}$ are derived from an analysis of the demographic noise [8]. Some development and confrontation to data remain to be done to clarify and confirm the structure of these terms. In the context of particle filtering, these terms are used in the prediction/mutation step: the structure of these terms determine the way the particles explore the state space. A uniform structure like "$+c_i\,\mathrm{d}W_t^i$" may be too vague and lead the particles into irrelevant areas of the state space, which can in turn lead to a loss of track (particles with zero or almost zero likelihood). The role of the c_i's is then clear: too small, the filter output variance will be small but the filter capacity to follow the evolution of the real state process is limited; too large, the latter capacity is more effective but the filter output variance will be large; far too large, again the filter will lose its ability to track the real state space vector. In conclusion, the c_i's are less model parameters to estimate than filter parameters to tune.

Concerning the noise intensity σ, it can be evaluated from the nature and the analysis of the sensors; it also plays a specific role in the particle filter algorithm. The observation equation (8) derives from the fact that the standard deviation of the noise is proportional to the substrate concentration. The likelihood function, which is used to weight the particles, is directly obtained from this observation equation, in particular its exponential form is due to the Gaussian hypothesis. The parameter σ is used to sharpen the likelihood function, and like for c_i's should be suitably tuned.

With respect to the nature of the noise: the state noises W_t^i are supposed to be independent Wiener processes (i.e. the time derivative of W_t^i is a white Gaussian noise), the observation noise v_n is supposed to be independent and identically distributed $N(0,1)$ (again a discrete time white Gaussian noise), moreover they are supposed to be mutually independent and independent from the initial condition Z_0. These hypotheses lead to the simplest model, one could propose other possibilities that will inevitably lead to the addition of parameters to the model or components to the state space vector.

V. Numerical tests

We propose two numerical tests with observations simulated from (8) and with state process simulated from the Euler-Maruyama scheme (15) with a time step $\delta_{\text{simu}} = \Delta/M_{\text{simu}}$ where M_{simu} is a multiple of M_{filter}.

A. Test 1: high frequency observations

Final time is $T = 1000$ (h) and we use the same discretization step for the observations, the simulation of the SDE and the filter: $\Delta = \delta_{\text{simu}} = \delta_{\text{filter}} = 0.5$ (i.e. $M_{\text{simu}} = M_{\text{filter}} = 1$). The parameters of the filter (inputs and model parameters) are: dilution rate $D = 0.01\,\text{h}^{-1}$; input substrate concentration $S_{\text{in}} = 100\,\text{mg/l}$; stoichiometric coefficient $k = 10$; maximum growth rate $\mu_{\text{max}} = 0.3\,\text{h}^{-1}$; half-saturation $k_s = 10\,\text{mg/l}$. The final time $T = 1000$ thus corresponds to 10 times the retention time $\frac{1}{D}$.

The state noise intensities are $c_1 = c_2 = 0.03$ while the observation noise intensity is $\sigma = 0.2$. The initial law is $\pi^0(\mathrm{d}x, \mathrm{d}s) = \mathcal{N}(0.2, 0.5^2) \otimes \mathcal{N}(1, 0.5^2)$.

The number of particles is $N = 1000$.

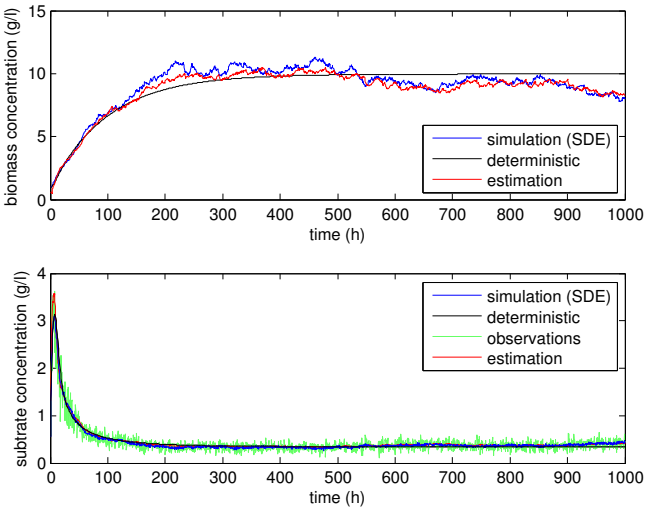


Fig. 1. Test 1. The simulated process $t \to (X_t, S_t)$ (—), the observation process $t_n \to Y_n$ (—), the estimates $t \to (\hat{X}_t, \hat{S}_t)$ (—) and the deterministic trajectory $t \to (x(t), s(t))$ (—) obtained with $c_1 = c_2 = 0$.

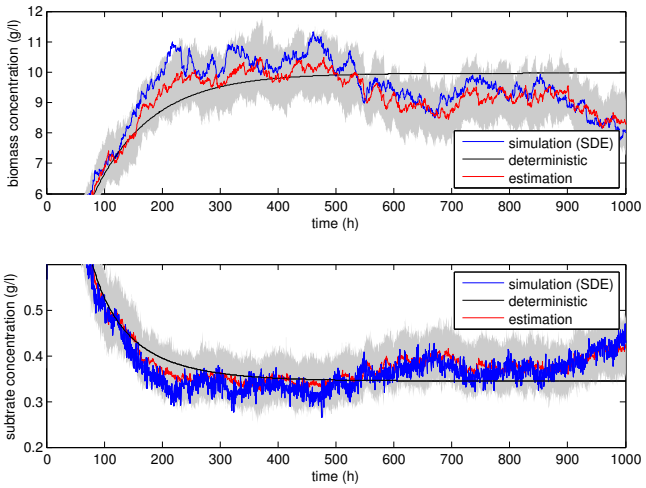


Fig. 2. Test 1. Same graphics as Fig. 1 but without the observation process and with the grey "tubes" (■) around the estimates $\hat{X}_t$ and $\hat{S}_t$ which are the minimum and maximum values taken by the x and s components of the particles.

B. Test 2: low frequency observations

This test is identical to Test 1 except that we do not use the same time step for the simulation of the SDE, for the observations and for the filter: $\Delta = 10$, $\delta_{\text{simu}} = \delta_{\text{filter}} = 0.5$ (i.e. $M_{\text{simu}} = M_{\text{filter}} = 20$). It should be noted that in this test, the particle filter is much faster than in Test 1, the respective CPU times being 17.3945 (s) for Test 1 and 0.84054 (s) for Test 2. This is due to the fact that the resampling step is time consuming, as it cannot be vectorized, and as it is used for each new measurement, it penalizes the Test 1.

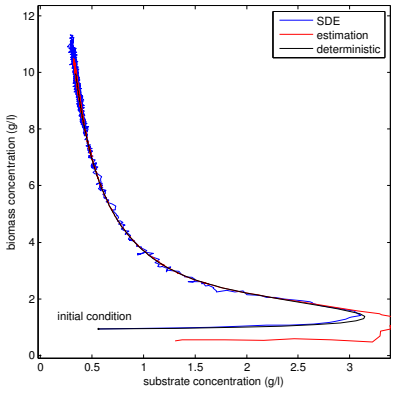


Fig. 3. Phase plot for Test 1 with the simulated process $t \to (X_t, S_t)$ (—), the estimates $t \to (\hat{X}_t, \hat{S}_t)$ (—) and the deterministic trajectory $t \to (x(t), s(t))$ (—) obtained with $c_1 = c_2 = 0$.

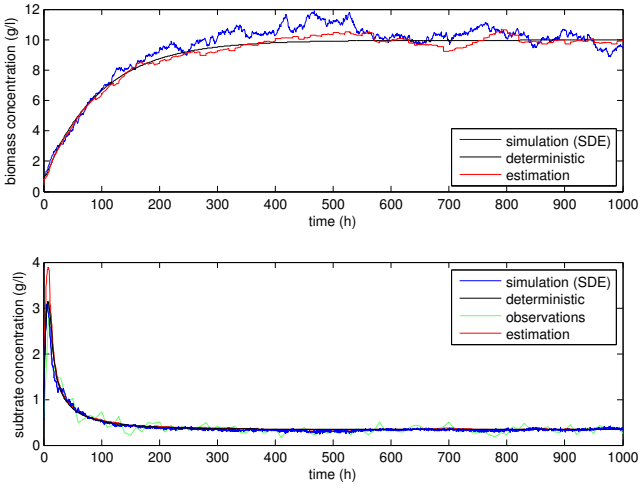


Fig. 4. Test 2. The simulated process $t \to (X_t, S_t)$ (—), the observation process $t_n \to Y_n$ (—), the estimates $t \to (\hat{X}_t, \hat{S}_t)$ (—) and the deterministic trajectory $t \to (x(t), s(t))$ (—) obtained with $c_1 = c_2 = 0$.

VI. Discussion and perspectives

We have developed a first implementation of a particle filter on a chemostat model. One challenge is to manage the fact that the frequency of observations is low, hence it is necessary to simulate the state equation between two observations. This prediction step can be done in an efficient way as it is highly parallelizable.

There exists many extensions of this work, we now explore some interesting leads.

Improved prediction step

In particle filter, the sum of the weights:

$$\sum_{i=1}^{N} \Psi(y_n, \tilde{\xi}_n^i)$$

used in the normalizing step (16) is a local likelihood: the higher it is, the better the filter behaves i.e. the more informative is $\pi_n^N(z)$. In most bioprocess applications, the

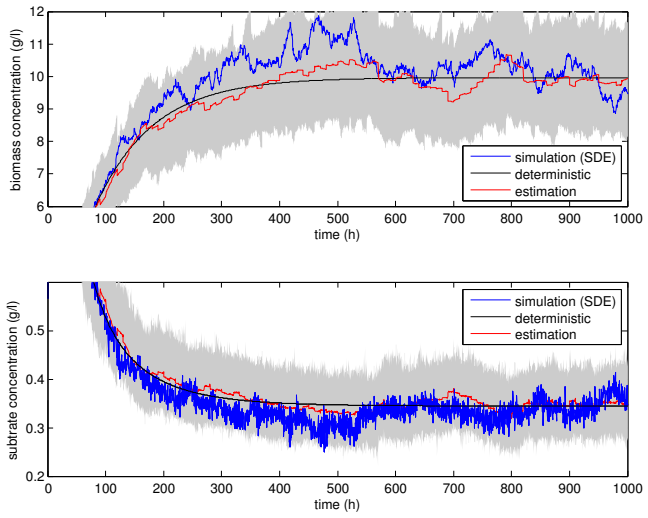


Fig. 5. Test 2. Same graphics as Fig. 4 but without the observation process and with the grey "tubes" (■) around the estimates $\hat{X}_t$ and $\hat{S}_t$ that are the minimum and maximum values taken by the x and s components of the particles.

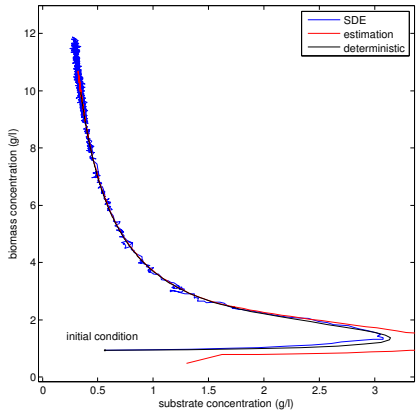


Fig. 6. Phase plot for Test 2 with the simulated process $t \to (X_t, S_t)$ (—), the estimates $t \to (\hat{X}_t, \hat{S}_t)$ (—) and the deterministic trajectory $t \to (x(t), s(t))$ (—) obtained with $c_1 = c_2 = 0$.

elapsed time between two consecutive measurements is long enough to significantly improve the prediction step. One can indeed sample as many particles as required to get N particles with large enough likelihood weights. It is also possible to call on specific techniques like the progressive correction proposed in [28].

Smoothing

If $T = n_{\max}\,\Delta$ is the final time of the experiment or of the bioreactor exploitation, it could be pertinent to compute the smoother [17], [6], that is:

$$\bar{\pi}_n(z) \stackrel{\text{def}}{=} p_{Z_{t_n}|Y_{0:n_{\max}}}(z|y_{0:n_{\max}})\,,$$

that is the distribution law of Z_{t_n} given that all the observations $Y_{0:n_{\max}} = y_{0:n_{\max}}$ available during the complete bioreactor experiment period. The smoother cannot be calculated sequentially, but it may be tractable in bioprocess applications where the elapsed time between two consecutive measurements is long enough to allow for cumbersome computations.

About the parameter estimation

In this work, the parameters are assumed to be known. This is not realistic and there are several ways to simultaneously identify the parameters and the state process.

The most classical idea to extend filtering procedure to the following augmented state:

$$\tilde{z}_n = \begin{pmatrix} z_n \\ \theta_n \end{pmatrix}\,, \tag{18a}$$

to consider the following artificial dynamics for the parameter:

$$\theta_{n+1} = \theta_n + \tilde{\sigma}\,\tilde{w}_n \tag{18b}$$

and to apply any filtering technique to the joint state dynamics (2) and (18b). This approach has been proposed with the EKF [26] and the UKF [24]. The difficulty is that with $\tilde{\sigma}$ small it fails at exploring the parameter space and with $\tilde{\sigma}$ large the covariance error for the parameter is too large. For too small $\tilde{\sigma}$, standard particle filters fail due to the particle degeneracy phenomenon, this issue can be addressed by specific sequential Monte Carlo procedures that use kernel approximation techniques [12], [11].

The advantage of the augmented state techniques is that that are sequential and can be achieved online in case of real-time constraint. It should be noticed that it is a Bayesian approach as a prior distribution on θ should be specified to initiate the dynamics (18b).

Another possibility is to treat the parameter in a frequentist way and to approximate the maximum likelihood estimate with direct maximization or with the Expectation/Maximization (EM) procedure [9]. The EM method coupled with the particle filter has been proposed in the context of bioprocess in [20].

Bayesian numerics should have a great impact in bioprocess applications where the low frequency of measurements allows for cumbersome numerical procedures, which is the case of the Bayesian numerics. There already exits applications of standard Monte Carlo Markov chain (MCMC) in the field of bioprocesses [22]. More elaborate procedures could be applied [10], [7], see [23] for an overview of MCMC for state-space models. Also particle Markov chain Monte Carlo (pMCMC) algorithms, which combine MCMC and SMC, proposed in [1] have already been applied in the field [19].

Other perspectives

The most interesting perspective is to test the robustness of the filter in the case of model mismatch: use a filter with a given growth function $\mu(s)$ as the observations correspond to a different growth function $\tilde{\mu}(s)$. Another interesting question is to determine which growth function, among a given finite set of growth functions $\{\mu_i(s); i = 1, \ldots, I\}$, is

underlying a given set of observations. Bayesian numerics, SMC and MCMC, provide promising solutions to these problems and may also extend to issues of change detection or control [2].

REFERENCES

[1] Christophe Andrieu, Arnaud Doucet, and Roman Holenstein. Particle markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010.

[2] Christophe Andrieu, Arnaud Doucet, Sumeetpal S. Singh, and Vladislav B. Tadić. Particle methods for change detection, identification and control. *Proceedings of the IEEE*, 92(3):423–438, 2004.

[3] Alan Bain and Dan Crişan. *Fundamentals of stochastic filtering*. Springer Verlag, 2008.

[4] Georges Bastin and Denis Dochain. *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier Science, 1990.

[5] Gildas Besançon, editor. *Nonlinear Observers and Applications*. Springer, 2007.

[6] Mark Briers, Arnaud Doucet, and Simon Maskell. Smoothing algorithms for state–space models. *Annals of the Institute of Statistical Mathematics*, 62(1):61–89, 2010.

[7] Fabien Campillo, Philippe Cantet, Rivo Rakotozafy, and Vivien Rossi. Méthodes MCMC en interaction pour l'évaluation de ressources naturelles. *ARIMA*, 8:64–80, 2008.

[8] Fabien Campillo, Marc Joannides, and Irène Larramendy-Valverde. Stochastic modeling of the chemostat. *Ecological Modelling*, 222(15):2676–2689, 2011.

[9] Fabien Campillo and François Le Gland. MLE for partially observed diffusions: direct maximization vs. the EM algorithm. *Stochastic Processes and their Applications*, 33(2):245–274, 1989.

[10] Fabien Campillo, Rivo Rakotozafy, and Vivien Rossi. Parallel and interacting Markov chain Monte Carlo algorithm. *Mathematics and Computers in Simulation*, 79:3424–3433, 2009.

[11] Fabien Campillo and Vivien Rossi. Convolution particle filter for parameter estimation in general state-space models. *IEEE Transactions on Aerospace and Electronic Systems*, 45(3):1063–1071, 2009.

[12] Tao Chen, Julian Morris, and Elaine Martin. Particle filters for state and parameter estimation in batch processes. *Journal of Process Control*, 15(6):665 – 673, 2005.

[13] Dan Crisan and Boris Rozovskii. *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2011.

[14] Denis Dochain. State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control*, 13(8):801 – 818, 2003.

[15] Randal Douc, Olivier Cappé, and Eric Moulines. Comparison of resampling schemes for particle filtering. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, ISPA*, 2005.

[16] Arnaud Doucet, Nando de Freitas, and Neil J. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer–Verlag, New York, 2001.

[17] Arnaud Doucet and Adam M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In *[13]*, 2009.

[18] Guillaume Goffaux and Alain Vande Wouwer. *Bioprocess state estimation: some classical and less classical approaches*, pages 111–128. Berlin: Springer, 2005.

[19] Andrew Golightly and Darren J. Wilkinson. Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo. *Interface Focus*, September 2011.

[20] R. Bhushan Gopaluni, Thomas B. Schön, and Adrian G. Wills. Particle filter approach to nonlinear system identification under missing observations with a real application. In *Proceedings of the 15th IFAC Symposium on System Identification (SYSID)*, Saint-Malo, France, July 2009, July 2009.

[21] Lorens Imhof and Sebastian Walcher. Exclusion and persistence in deterministic and stochastic chemostat models. *Journal of Differential Equations*, 217(1):26–53, 2005.

[22] Seunghee S. Jang, Hector De la Hoz, Amos Ben-zvi, William C. McCaffrey, and R. Bhushan Gopaluni. Parameter estimation in models with hidden variables : An application to a biotech process. *The Canadian Journal of Chemical Engineering*, 90(3):690–702, 2012.

[23] Nikolas Kantas, Arnaud Doucet, Sumeetpal S. Singh, and Jan Marian Maciejowski. An overview of sequential Monte Carlo methods for parameter estimation in general state-space models. In *15th IFAC Symposium on System Identification (SYSID), Saint-Malo, France*, 2009.

[24] Giancarlo Marafioti, Sihem Tebbani, Dominique Beauvois, Giuliana Becerra-Celis, Arsène Isambert, and M. Hovd. Unscented Kalman filter state and parameter estimation in a photobioreactor for microalgae production. In *Proceedings of International Symposium on Advanced Control of Chemical Processes*, Istanbul, Turquie, 2009.

[25] Ali Mesbah, Adrie E. M. Huesman, Herman J. M. Kramer, and Paul M. J. Van den Hof. A comparison of nonlinear observers for output feedback model-based control of seeded batch crystallization processes. *Journal of Process Control*, January 2011.

[26] Jan Kloppenborg Møller, Kirsten Riber Bergmann, Lasse Engbo Christiansen, and Henrik Madsen. Development of a restricted state space stochastic differential equation model for bacterial growth in rich media. *Journal of Theoretical Biology*, 305:78 – 87, 2012.

[27] Olga L. Q. Montoya, Gustavo Scaglia, Fernando di Sciascio, and Vicente Mut. Numerical methods based strategy and particle filter state estimation for bio process control. In *IEEE International Conference on Industrial Technology (ICIT)*, pages 1–6, 2008.

[28] Christian Musso, Nadia Oudjane, and François Le Gland. Improving regularized particle filters. In Arnaud Doucet, Nando de Freitas, and Neil J. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, chapter 12, pages 247–271. Springer–Verlag, New York, 2001.

[29] Hal L. Smith and Paul E. Waltman. *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge University Press, 1995.