Ecological Modelling 222 (2011) 2676–2689

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Stochastic modeling of the chemostat

F. Campillo[a,*], M. Joannides[a,b,1], I. Larramendy-Valverde[b,1]

[a] MERE Project-team, INRIA/INRA, UMR MISTEA, Montpellier, France
[b] Université Montpellier 2/I3M, Montpellier, France

ARTICLE INFO

Article history:
Received 16 January 2011
Received in revised form 21 April 2011
Accepted 26 April 2011
Available online 12 June 2011

Keywords:
Stochastic differential equations
Chemostat
Pure jump process
Diffusion approximation
Tau-leap method
Monte Carlo method
Gillespie algorithm

ABSTRACT

The chemostat is classically represented, at large population scale, as a system of ordinary differential equations. Our goal is to establish a set of stochastic models that are valid at different scales: from the small population scale to the scale immediately preceding the one corresponding to the deterministic model. At a microscopic scale we present a pure jump stochastic model that gives rise, at the macroscopic scale, to the ordinary differential equation model. At an intermediate scale, an approximation diffusion allows us to propose a model in the form of a system of stochastic differential equations. We expound the mechanism to switch from one model to another, together with the associated simulation procedures. We also describe the domain of validity of the different models.

1. Introduction

The chemostat, also known as the continuous stirred tank reactor, was independently invented by Monod (1950) and Novick and Szilard (1950). A complete derivation of the simple chemostat model, with a single species and a single substrate, was proposed by Herbert et al. (1956); a complete analysis of this model could be found in Smith and Waltman (1995). The evolution of the state of a simple chemostat is usually described by a set of ordinary differential equations (ODE) derived from a mass balance principle. More precisely, if $s(t)$ denotes the concentration of nutrient (substrate) and $b(t)$ the concentration of the organism (biomass) at time t (expressed in g/L), then the couple $x(t)=(b(t), s(t))$ is the solution of the following ODE:

$$\dot{b}(t) = [\mu(s(t)) - D]b(t), \tag{1a}$$

$$\dot{s}(t) = -k\mu(s(t))b(t) + D[s^{\text{in}} - s(t)] \tag{1b}$$

where $D>0$ is the dilution rate, $s^{\text{in}}>0$ the substrate concentration in the influent, and $k>0$ the stoichiometric coefficient. The initial condition lies in the positive orthant, that is $b(0)\geq 0$ and $s(0)\geq 0$. Eq. (1) will also be denoted:

$$\dot{x}(t) = f(x(t)).$$

The specific growth rate function $\mu(s)$ is non-negative; we suppose that $\mu(0)=0$, $\mu(s)>0$ for $s>0$, $\mu(s)\leq \mu_{\max}<\infty$ and that it is continuous at 0. Commonly used models are the Monod model (uninhibited growth) $\mu(s)=\mu_{\max}s/(k_s+s)$ and the Haldane model (inhibited growth) $\mu(s)=\mu_{\max}s/(k_s+s+s^2/k_i)$ (Smith and Waltman, 1995).

The modeling process that leads to (1) relies on the fact that the stochastic effects can be neglected or averaged out, thanks to the law of large numbers. This is possible only at macroscopic scale, for large population sizes, and under homogeneity conditions. At all other scales or when the homogeneity conditions are not met, random effects cannot be neglected. This is the case at microscopic scale, in small population size, as well as all scales preceding the one where (1) is valid. This is also when the homogeneity condition is not met, e.g. in unstirred conditions. Also the accumulation of small perturbations in the context of multi-species could not be neglected. Moreover, whereas the experimental results observed in well mastered laboratory conditions match closely the ODE theoretical behavior, a noticeable difference may occur in operational conditions.

So, even if the description (1) is sufficient for a number of applications of interest, it does not account for the stochastic aspects of the problem; we aim to build a model that still relies on a mass

* Corresponding author at: MERE Project-team, INRIA/INRA, UMR MISTEA, 2 Place Viala, 34060 Montpellier Cedex 01, France.
E-mail addresses: Fabien.Campillo@inria.fr (F. Campillo), marc.joannides@univ-montp2.fr (M. Joannides), larra@math.univ-montp2.fr (I. Larramendy-Valverde).
[1] Present address: Université Montpellier 2/I3M, Case Courrier 51, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France.

0304-3800/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2011.04.027

balance principle and that encompasses the useful stochastic information.

A few papers have already addressed the stochastic modelling of the chemostat. Crump and O'Young (1979) proposed a pure jump process model for the biomass dynamics coupled with an ODE for the substrate. They described the associated stochastic simulation algorithm, also known as Gillespie algorithm, to simulate trajectories of the population size and of the substrate concentration. They derived the approximation of the two first moments conditioned upon the non-extinction of the cell population, and also the distribution for the waiting time to the population changeover.

Independently, Stephanopoulos et al. (1979) considered a two populations chemostat with random fluctuations of the dilution rate about a value $\bar{D}$ that allows coexistence. They replaced D in (1) by $\bar{D} + \dot{W}_t$ where $\dot{W}_t$ is a white Gaussian noise (i.e. formally the time-derivative of a Wiener process). Analytical and numerical solutions of the associated Fokker–Planck equation prove that extinction of one of the two populations will eventually take place.

More recently Grasman and De Gee (2005) proposed a stochastic model for the dynamics of a chemostat with three trophic levels where the stochasticity appears only in the top level trophic as a stochastic logistic model and with fluid limit dynamics for the two other levels. A diffusion model of the process is formulated and with singular perturbation methods applied to the corresponding Fokker–Planck equation an estimate of the expected extinction time of the population at top level trophic is derived. In this model, the stochastic differential equation is derived as a diffusion approximation of a birth and death process.

Finally (Imhof and Walcher, 2005) introduced a variant of the deterministic single-substrate chemostat model for which the persistence of all species is possible. To derive a stochastic model they considered a discrete-time Markov process with jumps corresponding to the deterministic system added with a centered Gaussian term, letting the time step converges to zero leads to a system of stochastic differential equations. They proved that random effects may lead to extinction in scenarios where the deterministic model predicts persistence; they also established some stochastic persistence results.

These three last works propose to superimpose a stochastic term on Eq. (1) in order to model the uncertainty on the phenomenon, principally due to imprecise experimental conditions. We propose instead to consider the stochastic aspect at the very beginning of the modeling process like in Crump and O'Young (1979). This approach is not individual-based per se, as it starts from the macroscopic model (1). However, the first stochastic model proposed will be described at the individual level. This method will allow for a justification of the specific structure of the stochastic perturbation that affects the mean behavior. More generally, we will outline a modeling strategy based on many available tools, either stochastic or deterministic, depending on the regularity of the phenomenon to be modeled. In this paper we focus on the modeling and simulation process rather than on the mathematical developments; moreover we make use of known mathematical results. Our goal is to establish a set of stochastic models that are valid at the scales where the deterministic model (1) is not.

The paper is organized as follows: in Section 2, we recall the origin of model (1) and the assumptions ensuring its validity. We show that since different timescales naturally appear in the problem, these assumptions need to be checked at each scale. Section 3 is devoted to the different models: the pure jump description that will be considered as the reference model is introduced in Section 3.1; the discrete time approximation, Poisson and normal, are presented in Section 3.2; the discrete-time normal approximation appeared to be a time discretization of a diffusion process given by a stochastic differential equation presented in Section 3.3. In Section 3.4 we describe the asymptotic results that bridge these different models. Section 4 is devoted to the associated simulation algorithm, Section 5 to numerical tests.

2. Scale and geometry issues in ODE model

An individual-based model should keep track of the position in space of each cell, together with their current biological states, it should also account for discrete events such as the division of a cell. Such a description of the system at the finest level could be of interest but unnecessary in view of our goals, namely to set a macroscopic model that gives account for stochastic phenomena. At this scale, the system is reduced to a $\mathbb{R}^2$-vector and its dynamics.

Model (1) is obtained according to the classical approach, by choosing a small time interval Δt on which a mass balance principle is applied. However, Δt should be large enough as we do not describe the dynamic at the timescale of jumps of one unit of substrate or bacteria but rather at the timescale of jumps of packet units. Such an interval could be called *macroscopically infinitesimal* (Gillespie, 2000).

2.1. Mass balance

Let $(\mathcal{B}_t, \mathcal{S}_t)$ denote the true concentrations at time t, assumed to be constant throughout the medium. The mass balance on interval $[t, t+\Delta t)$ reads

$$\mathcal{B}_{t+\Delta t} - \mathcal{B}_t = \Delta\mathcal{B}_t^{\text{bio}} + \Delta\mathcal{B}_t^{\text{out}}, \tag{2a}$$

$$\mathcal{S}_{t+\Delta t} - \mathcal{S}_t = \Delta\mathcal{S}_t^{\text{bio}} + \Delta\mathcal{S}_t^{\text{in}} + \Delta\mathcal{S}_t^{\text{out}} \tag{2b}$$

where $\Delta\mathcal{B}_t^{\text{bio}}$ and $\Delta\mathcal{B}_t^{\text{out}}$ are the increments of biomass due to natural growth and to the outflow respectively, within $[t, t+\Delta t)$, $\Delta\mathcal{S}_t^{\text{bio}}$, $\Delta\mathcal{S}_t^{\text{in}}$ and $\Delta\mathcal{S}_t^{\text{out}}$ are the increments of substrate due to the consumption by the biomass, the inflow and the outflow respectively, within $[t, t+\Delta t)$. Since we want to obtain an ODE, we now assume that the stochastic fluctuations are negligible relative to the increments. Again this requires Δt to be large enough, so that sufficiently many discrete events have occurred. Moreover, Δt should be taken even larger in case of inhomogeneity of the dynamics.

We denote by $(\bar{b}(t), \bar{s}(t))$ for $t = 0, \Delta t, 2\Delta t, \ldots$ the deterministic sequence constructed by using the mean increments of $(\mathcal{B}_t, \mathcal{S}_t)$:

$$\bar{b}(t+\Delta t) - \bar{b}(t) = \mathbb{E}[\Delta\mathcal{B}_t^{\text{bio}} + \Delta\mathcal{B}_t^{\text{out}}],$$
$$\bar{s}(t+\Delta t) - \bar{s}(t) = \mathbb{E}[\Delta\mathcal{S}_t^{\text{bio}} + \Delta\mathcal{S}_t^{\text{in}} + \Delta\mathcal{S}_t^{\text{out}}].$$

Next, using the mass action law for the biomass we have

$$\mathbb{E}[\Delta\mathcal{B}_t^{\text{bio}}] \simeq \mu(\bar{s}(t))\bar{b}(t)\Delta t, \quad \mathbb{E}[\Delta\mathcal{S}_t^{\text{bio}}] \simeq -k\mu(\bar{s}(t))\bar{b}(t)\Delta t \tag{3}$$

where $\mu(s)$ is the specific growth rate and $k > 0$ the stoichiometric coefficient. Note that we again require Δt to be large enough, since $\mu(s)$ and k make sense only for a sufficiently large population of bacteria. Now, since we have assumed perfect homogeneity of the medium, we get:

$$\mathbb{E}[\Delta\mathcal{B}_t^{\text{out}}] \simeq -D\bar{b}(t)\Delta t, \quad \mathbb{E}[\Delta\mathcal{S}_t^{\text{in}}] \simeq Ds^{\text{in}}\Delta t, \quad \mathbb{E}[\Delta\mathcal{S}_t^{\text{out}}] \simeq -D\bar{s}(t)\Delta t. \tag{4}$$

Note that (3) and (4) are approximations because we have used a constant value for $\bar{b}(t)$ and $\bar{s}(t)$ within $[t, t+\Delta t)$. For this approximations to be correct, none of the quantities involved should vary significantly within $[t, t+\Delta t)$. We finally obtain the construction of the sequence $(\bar{b}(t), \bar{s}(t))$ by

$$\bar{b}(t+\Delta t) - \bar{b}(t) = [\mu(\bar{s}(t)) - D]\bar{b}(t)\Delta t, \tag{5a}$$

$$\bar{s}(t+\Delta t) - \bar{s}(t) = (-k\mu(\bar{s}(t))\bar{b}(t) + D[s^{\text{in}} - \bar{s}(t)])\Delta t, \tag{5b}$$

Model (1) is obtained by letting $\Delta t \to 0$ in (5). However, since Δt is bounded from below, some care should be taken when this

limit is achieved. System (5) can be understood as the discretization of (1) using an explicit Euler scheme. Whenever there exists Δt sufficiently small, the deterministic sequence $(\bar{b}(t), \bar{s}(t))$ will be close to model (1), sampled at instants $k\Delta t$, $k \in \mathbb{N}$.

2.2. Geometry and scales

The mass balance established in (2) features five terms that can be gathered according to the three sources of variations. This gives rise to a geometric structure that can be emphasized by writing (1) under the form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} b(t) \\ s(t) \end{pmatrix} = \underbrace{\mu(s(t))b(t)\begin{pmatrix} 1 \\ -k \end{pmatrix}}_{\text{biology}} + \underbrace{D\begin{pmatrix} 0 \\ s^{\text{in}} \end{pmatrix}}_{\text{inflow}} - \underbrace{D\begin{pmatrix} b(t) \\ s(t) \end{pmatrix}}_{\text{outflow}} \tag{6}$$

However, whereas the geometry is well captured, the scale of the five original terms is not readable in (1) nor (6). Indeed, the fact that the approximations in (3) and (4) may be of different quality for each individual term is not exploited at all.

3. Models at different scales

In the previous section, we mentioned that the lower bound for Δt is related to the size of the population and to the regularity of the phenomenon. Often, the experimental conditions are such that the regularity of the real system is sufficient to ensure that the approximations (3) and (4) are valid, even for small Δt. In that case, system (5) is correctly approximated by (1) sampled with period Δt. If a smaller period is to be considered, then the conditions under which (5) has been obtained are not fulfilled. Particularly, the stochastic fluctuations should be accounted for.

One could think about looking for the most microscopic description, where each single bacteria is explicitly represented with its proper characteristics, namely mass, age, and location. Notice that the situation notably differs from the setting encountered in chemistry, where individuals are molecules and each reaction is instantaneous, with a strict stoichiometry. As a result, we cannot approximate the most microscopic behavior by a jump Markov process on the integer lattice.

We now introduce a stochastic process built on the same premise, that is a mean mass balance principle at a given Δt. This model will have (1) as a fluid limit as Δt goes to 0. This latter model suitably features the geometry of the chemostat but, as a limit model, cannot feature all its natural scales. The proposed stochastic models will respect both the geometry and the natural scales of the chemostat. We first establish a pure jump process representation of the chemostat at a microscopic scale, then we derive a diffusion process representation which will be valid at mesoscopic and macroscopic scales.

3.1. Pure jump model $X_t = (B_t, S_t)$

Even if we do not aim at deriving an individual-based model, we try to preserve the discrete feature in the dynamics. We achieve this by considering only *aggregated jumps* obtained by adding up small and frequent jumps resulting from individual events. The resulting stochastic process will be a pure jump process $X_t = (B_t, S_t)$, fully determined by its jumps and the corresponding jump rates; the state variable will be denoted $x = (b, s)$.

In view of (2), we are led to consider five jumps:

① biology term: biomass increase of size $\nu_1(x)$ at rate $\lambda_1(x)$;

② biology term: substrate decrease of size $\nu_2(x)$ at rate $\lambda_2(x)$;

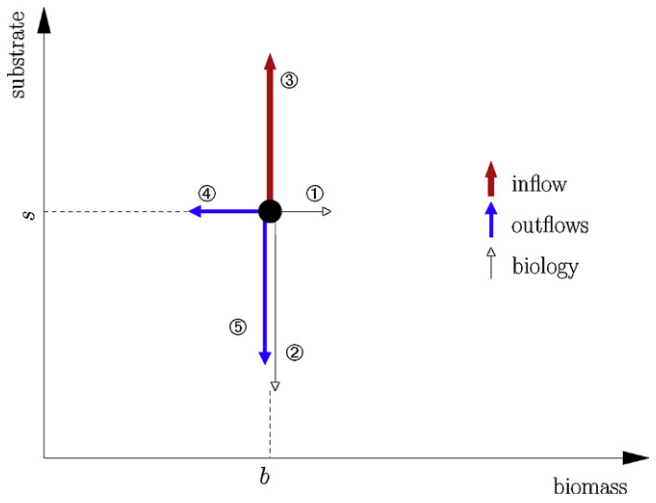

Fig. 1. In this model, from a position $x = (b, s)$ the process could jump according to 5 mechanisms (2 due to the biology, 1 inflow, and 2 outflows), the basic jump ⓘ has a length $(1/K_i)$ for $i = 1, \ldots, 5$.

③ inflow term: substrate inflow of size $\nu_3(x)$ at rate $\lambda_3(x)$;

④ outflow term: biomass outflow of size $\nu_4(x)$ at rate $\lambda_4(x)$;

⑤ outflow term: substrate outflow of size $\nu_5(x)$ at rate $\lambda_5(x)$;

See Figure 1. It remains to set the jump size rates so as to comply with the mass balance principle and the stochastic mass action law.

For a macroscopically infinitesimal Δt, denote by $\Delta X_t^{\text{b,bio}}$, $\Delta X_t^{\text{s,bio}}$, $\Delta X_t^{\text{s,in}}$, $\Delta X_t^{\text{b,out}}$, $\Delta X_t^{\text{s,out}}$ the cumulated jump of type ①, ②, ③, ④, ⑤ respectively, on state process X_t within the time interval $[t, t + \Delta t)$.

We first focus on the first two expressions. The stochastic mass action law (Wilkinson, 2006) requires

$$\mathbb{E}[\Delta X_t^{\text{b,bio}} | X_t = x] \simeq \begin{pmatrix} \mu(s)b\Delta t \\ 0 \end{pmatrix},$$

$$\mathbb{E}[\Delta X_t^{\text{s,bio}} | X_t = x] \simeq \begin{pmatrix} 0 \\ -k\mu(s)b\Delta t \end{pmatrix}.$$

Now notice that, for small Δt, the number of jumps of type ① (resp. ②) within $[t, t + \Delta t)$ is approximately $\mathcal{P}(\lambda_1(x)\Delta t)$ (resp. $\mathcal{P}(\lambda_2(x)\Delta t)$), so that

$$\mathbb{E}[\Delta X_t^{\text{b,bio}} | X_t = x] \simeq \lambda_1(x)\Delta t \nu_1(x),$$

$$\mathbb{E}[\Delta X_t^{\text{s,bio}} | X_t = x] \simeq \lambda_2(x)\Delta t \nu_2(x).$$

So we are looking for $(\lambda_i(x), \nu_i(x))$ satisfying:

$$\lambda_1(x)\nu_1(x) = \begin{pmatrix} \mu(s)b \\ 0 \end{pmatrix} \quad \text{and} \quad \lambda_2(x)\nu_2(x) = \begin{pmatrix} 0 \\ -k\mu(s)b \end{pmatrix}. \tag{7}$$

We therefore introduce the *scale* parameters K_1 and K_2 and we choose:

$$\lambda_1(x) \stackrel{\text{def}}{=} K_1\mu(s)b, \quad \nu_1 \stackrel{\text{def}}{=} \begin{pmatrix} \frac{1}{K_1} \\ 0 \end{pmatrix},$$

$$\lambda_2(x) \stackrel{\text{def}}{=} K_2 k\mu(s)b \quad \nu_2 \stackrel{\text{def}}{=} -\begin{pmatrix} 0 \\ \frac{1}{K_2} \end{pmatrix}.$$

This choice is not unique and will be explain later in Section 3.5.

Here by "scale" we mean that jumps due to ⓘ will be of magnitude $1/K_i$ and the corresponding rates will be of magnitude K_i. Large K_i yields frequent and small jumps. Using the Poisson argument mentioned above, we see that these scale parameters K_i do not act on the mean values of the increments but on their variances (large K_i will correspond to small variances). The K_i's can thus be regarded as tuning parameters quantifying the uncertainty or regularity of the corresponding source of variation.

Reproducing this discussion with the three other types of jumps, and considering only admissible jumps (in the positive orthant), we obtain a pure jump Markov process with rate coefficients $\lambda_i(x)$ and associated jumps $\nu_i(x)$ defined in Table 1.

3.1.1. About scales parameters

In most cases $K_i \ll K_j$ for $i = 1, 4$ and $j = 2, 3, 5$, indeed the first set of scale parameters refers to the bacterial dynamics and the second set of scale parameters refers to the substrate dynamics. It is however possible to adjust the coefficients K's to the specific application considered. For example K_3 will be large in laboratory experimental conditions, but for a real implementation the substrate inflow concentration could have a large variance. Also for the outflow, in regular conditions K_4 and K_5 could be large, but in bad mixing conditions they could be smaller. Finally K_1 could be smaller than K_2, as the biomass concentration increase presents more variance than the substrate decrease (which is more regular as it is related to the diffusion of substrate across cell membranes).

As explained later, the jumps $\nu_i(x)$ are essentially constant and equal to:

$$\bar{\nu}_1 \overset{\text{def}}{=} \begin{pmatrix} \frac{1}{K_1} \\ 0 \end{pmatrix}, \quad \bar{\nu}_2 \overset{\text{def}}{=} -\begin{pmatrix} 0 \\ \frac{1}{K_2} \end{pmatrix}, \quad \bar{\nu}_3 \overset{\text{def}}{=} \begin{pmatrix} 0 \\ \frac{1}{K_3} \end{pmatrix}$$

$$\bar{\nu}_4 \overset{\text{def}}{=} -\begin{pmatrix} \frac{1}{K_4} \\ 0 \end{pmatrix}, \quad \bar{\nu}_5 \overset{\text{def}}{=} -\begin{pmatrix} 0 \\ \frac{1}{K_5} \end{pmatrix}.$$

3.1.2. Representation of X_t

The constructive description of the process X_t that has been just presented would be used for simulation purposes, see Section 4.1. Nevertheless, it should be completed by a more comprehensible and synthetic representation.

The process X_t is in fact a Markov process that can be represented as the following (jump) SDE:

$$X_t = X_0 + \sum_{i=1}^{5} \int_{(0,t]\times[0,\infty)} \nu_i(X_{u^-}) 1_{\{\nu \le \lambda_i(X_{u^-})\}} N^i(\mathrm{d}u \times \mathrm{d}\nu) \tag{8}$$

where N^i are independent random Poisson measures with intensity measure $\mathrm{d}u \times \mathrm{d}\nu$ (Lebesgue measure). The principle of this representation is simple: a random Poisson measure is associated with each of the five types of events, the indicator function $1_{\{\nu \le \lambda_i(x)\}}$ allows the corresponding clock to be of rate $\lambda_i(x)$, then at each new "i"-type event, the current state value of the chemostat x jumps to $x + \nu_i(x)$. The resulting process (8) is an exact mathematical transcription of the simulation procedure described in Section 3.1.

The law of this process is characterized by its infinitesimal generator defined by $\mathcal{A}\phi(x) = \lim_{t\to 0} \frac{1}{t}[\mathbb{E}(\phi(X_t)|X_0 = x) - \phi(x)]$. A simple computation leads to:

$$\begin{aligned} \mathcal{A}\phi(x) &= \sum_{i=1}^{5} \lambda_i(x)[\phi(x + \nu_i(x)) - \phi(x)] \\ &= \lambda(x) \int_{\mathbb{R}_+^2} [\phi(y) - \phi(x)]\rho(x, \mathrm{d}y) \end{aligned} \tag{9}$$

for all $\phi : \mathbb{R}_+^2 \mapsto \mathbb{R}$ continuous with compact support (Ethier and Kurtz, 1986, Th. 8–3.1), where:

$$\lambda(x) \overset{\text{def}}{=} \sum_{i=1}^{5} \lambda_i(x), \quad \rho(x, \mathrm{d}y) \overset{\text{def}}{=} \sum_{i=1}^{5} \bar{\lambda}_i(x) \delta_{x+\nu_i(x)}(\mathrm{d}y), \quad \bar{\lambda}_i(x) \overset{\text{def}}{=} \frac{\lambda_i(x)}{\sum_{i'=1}^{5} \lambda_{i'}(x)}$$

In Campillo et al. (2010), we prove that the Markov process X_t with infinitesimal generator $\mathcal{A}$ defined by (9) is non-explosive and admits moments of all order for all $t \ge 0$.

To describe how the process X_t can be simulated, we introduce the jump times $\tau_0 = 0$ and $\tau_n \overset{\text{def}}{=} \inf\{t > \tau_{n-1}; X_t \ne X_{\tau_{n-1}}\}$ and the embedded jump chain $Y_n \overset{\text{def}}{=} X_{\tau_n}$. It is well know that (i) Y_n is a Markov chain on $\mathbb{R}_+^2$ with transition probability $\rho(x, \mathrm{d}y)$; (ii) for all $n \ge 1$, conditionally on $Y_0, \ldots, Y_{n-1}$, the holding times $\tau_1 - \tau_0, \ldots, \tau_n - \tau_{n-1}$ are independent and exponentially distributed of intensity parameters $\lambda(Y_0), \ldots, \lambda(Y_{n-1})$, see Norris (1998). These properties are at the basis of the Gillespie simulation algorithm (see Algorithm 4.1).

From (9) we can deduce a more interesting representation for the process X_i. Indeed, introduce the centered random Poisson measures $\tilde{N}^i(\mathrm{d}u \times \mathrm{d}\nu) \overset{\text{def}}{=} N^i(\mathrm{d}u \times \mathrm{d}\nu) - \mathrm{d}u \times \mathrm{d}\nu$, then (8) reads:

$$\begin{aligned} X_t = X_0 &+ \sum_{i=1}^{5} \int_0^t \int_0^\infty \nu_i(X_{u^-}) 1_{\{\nu \le \lambda_i(X_{u^-})\}} \mathrm{d}u\mathrm{d}\nu \\ &+ \sum_{i=1}^{5} \int_{(0,t]\times[0,\infty)} \nu_i(X_{u^-}) 1_{\{\nu \le \lambda_i(X_{u^-})\}} \tilde{N}^i(\mathrm{d}u \times \mathrm{d}\nu) \end{aligned}$$

by letting

$$f_K(x) \overset{\text{def}}{=} \sum_{i=1}^{5} \int_0^\infty \nu_i(x) 1_{\{\nu \le \lambda_i(x)\}} \mathrm{d}\nu = \sum_{i=1}^{5} \nu_i(x)\lambda_i(x)$$

and

$$M_t^i \overset{\text{def}}{=} \sum_{i=1}^{5} \int_{(0,t]\times[0,\infty)} \nu_i(X_{u^-}) 1_{\{\nu \le \lambda_i(X_{u^-})\}} \tilde{N}^i(\mathrm{d}u \times \mathrm{d}\nu)$$

we get

$$X_t = X_0 + \int_0^t f_K(X_u)\mathrm{d}u + \sum_{i=1}^{5} M_t^i \tag{10}$$

where M_t^i are five independent square-integrable martingales.

In (10), $f_K(x)$ is "essentially" the r.h.s. function $f(x)$ of the O.D.E. (1). More precisely, let

$$\mathcal{R}_K \overset{\text{def}}{=} \left\{ x = \begin{pmatrix} b \\ s \end{pmatrix} \in \mathbb{R}_+^2; b \le \frac{1}{K_4} \text{ or } s \le \frac{1}{K_2} \text{ or } s \le \frac{1}{K_5} \right\}$$

then, for $x \notin \mathcal{R}_K$, $f_K(x) = f(x)$; and $f_K \to f$ when $K_i \to \infty$ for all $i = 1, \ldots, 5$ uniformly on any compact set of $(0, \infty)^2$ (see details in Campillo et al. (2010)). Replacing f_K by f in the decomposition (10), lead to the following system where the dependence on the K_i is enhanced:

$$\mathrm{d}B_t = (\mu(S_t)B_t - DB_t)\mathrm{d}t + \frac{\mathrm{d}\bar{m}_t^1}{\sqrt{K_1}} + \frac{\mathrm{d}\bar{m}_t^4}{\sqrt{K_4}}, \tag{11a}$$

$$\mathrm{d}S_t = (-k\mu(S_t)B_t + D(s^{\text{in}} - S_t))\mathrm{d}t + \frac{\mathrm{d}\bar{m}_t^2}{\sqrt{K_2}} + \frac{\mathrm{d}\bar{m}_t^3}{\sqrt{K_3}} + \frac{\mathrm{d}\bar{m}_t^5}{\sqrt{K_5}} \tag{11b}$$

Table 1
Rates and jumps of the five basic mechanisms of the pure jump process. Note that the jumps $\nu_i(x)$ essentially do not depend on x except for the negative jumps near the border $\{x=(b,s)\in\mathbb{R}_+^2; b=0 \text{ or } s=0\}$.

	① biomass increase (biology)	② substrate decrease (biology)	③ substrate inflow	④ biomass outflow	⑤ substrate outflow
Rate $\lambda_i(x)$	$K_1\mu(s)b$	$K_2k\mu(s)b$	K_3Ds^{in}	K_4Db	K_5Ds
Jump $\nu_i(x)$	$\begin{pmatrix}\frac{1}{K_1}\\0\end{pmatrix}$	$-\begin{pmatrix}0\\\frac{1\wedge K_2s}{K_2}\end{pmatrix}$	$\begin{pmatrix}0\\\frac{1}{K_3}\end{pmatrix}$	$-\begin{pmatrix}\frac{1\wedge K_4b}{K_4}\\0\end{pmatrix}$	$-\begin{pmatrix}0\\\frac{1\wedge K_5s}{K_5}\end{pmatrix}$

where $\tilde{m}_t^i$ are independent centered square integrable martingales given by:

$$\tilde{m}_t^1 \overset{\text{def}}{=} \frac{1}{\sqrt{K_1}}\int_0^t\int_0^\infty 1_{\{\nu\le\lambda_1(X_{u^-})\}}\tilde{N}^1(\mathrm{d}u\times\mathrm{d}\nu),$$

$$\tilde{m}_t^2 \overset{\text{def}}{=} -\frac{1}{\sqrt{K_2}}\int_0^t\int_0^\infty (1\wedge K_2S_{u^-})1_{\{\nu\le\lambda_2(X_{u^-})\}}\tilde{N}^2(\mathrm{d}u\times\mathrm{d}\nu),$$

$$\tilde{m}_t^3 \overset{\text{def}}{=} \frac{1}{\sqrt{K_3}}\int_0^t\int_0^\infty 1_{\{\nu\le\lambda_3(X_{u^-})\}}\tilde{N}^3(\mathrm{d}u\times\mathrm{d}\nu),$$

$$\tilde{m}_t^4 \overset{\text{def}}{=} -\frac{1}{\sqrt{K_4}}\int_0^t\int_0^\infty (1\wedge K_4B_{u^-})1_{\{\nu\le\lambda_4(X_{u^-})\}}\tilde{N}^4(\mathrm{d}u\times\mathrm{d}\nu),$$

$$\tilde{m}_t^5 \overset{\text{def}}{=} -\frac{1}{\sqrt{K_5}}\int_0^t\int_0^\infty (1\wedge K_5S_{u^-})1_{\{\nu\le\lambda_5(X_{u^-})\}}\tilde{N}^5(\mathrm{d}u\times\mathrm{d}\nu)$$

with the following predictable quadratic variations:

$$\langle\tilde{m}^1\rangle_t = \int_0^t \mu(S_u)B_u\mathrm{d}u, \tag{12a}$$

$$\langle\tilde{m}^2\rangle_t = \int_0^t (1\wedge K_2S_u)^2k\mu(S_u)B_u\mathrm{d}u, \tag{12b}$$

$$\langle\tilde{m}^3\rangle_t = Ds^{\text{in}}t, \tag{12c}$$

$$\langle\tilde{m}^4\rangle_t = \int_0^t (1\wedge K_4B_u)^2DB_u\mathrm{d}u, \tag{12d}$$

$$\langle\tilde{m}^5\rangle_t = \int_0^t (1\wedge K_5S_u)^2DS_u\mathrm{d}u. \tag{12e}$$

According to (11), the deterministic part of the dynamics, the drift coefficient, appears to be (essentially) the classical ODE (1); and the stochastic part of the dynamics, the martingale terms, are of order $1/\sqrt{K_i}$. These mathematical developments are detailed in Campillo et al. (2010).

3.2. Discrete time approximations

3.2.1. Poisson approximation $\tilde{X}_{t_n}=(\tilde{B}_{t_n},\tilde{S}_{t_n})$

For any small $\Delta t>0$ given, let $t_n=n\Delta t$. We propose a discrete time Poisson approximation $(\tilde{X}_{t_n})_{n\ge0}$ of $(X_t)_{t\ge0}$: on the interval $[t_n, t_{n+1})$ we froze the rate functions $\lambda_i(X_t)$ to $\lambda_i(X_{t_n})$ so that we get a Poisson distribution. The jumps $\nu_i(X_t)$ are also frozen to $\nu_i(X_{t_n})$. Let $\tilde{X}_0=X_0$, the approximation is defined by:

$$\tilde{X}_{t_{n+1}} = \tilde{X}_{t_n} + \sum_{i=1}^5 \nu_i(\tilde{X}_{t_n})\mathcal{P}_n^i(\Delta t\lambda_i(\tilde{X}_{t_n})) \tag{13}$$

where $(\mathcal{P}_n^i(\rho))_{n\in\mathbb{N},i=1\cdots5}$ are independent Poisson variables with intensities ρ. We have:

$$\mathbb{E}[\tilde{X}_{t_{n+1}}|\tilde{X}_{t_n}=x] = x + f_K(x). \tag{14}$$

In other words, the infinitesimal increments of the conditional mean follow the O.D.E. (1). Also:

$$\begin{aligned}\mathrm{cov}[\tilde{X}_{t_{n+1}}|\tilde{X}_{t_n}=x] &= \sum_{i=1}^5 \mathrm{cov}[\nu_i(x)\mathcal{P}_n^i(\Delta t\lambda_i(\tilde{X}_{t_n}))|\tilde{X}_{t_n}=x]\\ &= \begin{pmatrix}\tilde{\Sigma}_1^2 & 0\\ 0 & \tilde{\Sigma}_2^2\end{pmatrix}\end{aligned} \tag{15a}$$

with

$$\tilde{\Sigma}_1^2 = \Delta t\left\{\frac{1}{K_1}\mu(s)b + \frac{1}{K_4}(1\wedge K_4b)^2Db\right\}, \tag{15b}$$

$$\tilde{\Sigma}_2^2 = \Delta t\left\{\frac{1}{K_2}(1\wedge K_2s)^2k\mu(s)b + \frac{1}{K_3}Ds^{\text{in}} + \frac{1}{K_5}(1\wedge K_5s)^2Ds\right\}. \tag{15c}$$

3.2.2. Diffusion approximation $\tilde{\xi}_{t_n}=(\tilde{\beta}_{t_n},\tilde{\sigma}_{t_n})$

In (13), the variable $\mathcal{P}_n^i(\Delta t\lambda_i(x))$ is Poisson distributed with parameter $\Delta t\lambda_i(x)$. When this parameter is large (greater than 10 or 20) then this last distribution is very close to the normal distribution of mean $\Delta t\lambda_i(x)$ and variance $\Delta t\lambda_i(x)$. Hence, we get a (discrete time) normal approximation $(\tilde{\xi}_{t_n})_{n\ge0}$ of $(X_t)_{t\ge0}$ by letting $\tilde{\xi}_0=X_0$ and, conditionally on $\tilde{\xi}_{t_n}=x$:

$$\tilde{\xi}_{t_{n+1}} = x + \sum_{i=1}^5 \nu_i(x)\mathcal{N}_n^i$$

where $\mathcal{N}_n^i$ are 5 independent Gaussian random variables: $\mathcal{N}_n^i\sim\mathcal{N}(\lambda_i(x)\Delta t, \lambda_i(x)\Delta t)$. So conditionally on $\tilde{\xi}_{t_n}=x$, $\tilde{\xi}_{t_{n+1}}$ is normal with mean (14) and covariance matrix (15).

Let $\tilde{\xi}_{t_n}=(\tilde{\beta}_{t_n},\tilde{\sigma}_{t_n})$, given $\tilde{\beta}_{t_n}=b$ and $\tilde{\sigma}_{t_n}=s$:

$$\begin{aligned}\tilde{\beta}_{t_{n+1}} &= b + [\mu(s) - (1\wedge K_4b)D]b\Delta t\\ &\quad + \sqrt{\Delta t\frac{\mu(s)b}{K_1}}w_n^1 + \sqrt{\Delta t\frac{(1\wedge K_4b)^2Db}{K_4}}w_n^4\end{aligned} \tag{16a}$$

$$\begin{aligned}\tilde{\sigma}_{t_{n+1}} &= s + [-(1\wedge K_2s)k\mu(s)b + Ds^{\text{in}} - (1\wedge K_5s)Ds]\Delta t\\ &\quad + \sqrt{\Delta t\frac{(1\wedge K_2s)^2k\mu(s)b}{K_2}}w_n^2 + \sqrt{\Delta t\frac{Ds^{\text{in}}}{K_3}}w_n^3\\ &\quad + \sqrt{\Delta t\frac{(1\wedge K_5s)^2Ds}{K_5}}w_n^5\end{aligned} \tag{16b}$$

where w_n^i are i.i.d. $\mathcal{N}(0,1)$ random variables.

In both approximations (13) and (16), no mechanism prevents the processes $\tilde{X}_{t_n}$ or $\tilde{\xi}_{t_n}$ from staying within the positive orthant $\mathbb{R}_+^2$. An ad hoc solution is to set the concentration to 0 whenever it becomes negative, see Section 4.

3.3. Diffusion model $\xi_t = (\beta_t, \sigma_t)$

3.3.1. A stochastic differential equation model

System (16) is the Euler–Maruyama time discretization of the diffusion process $\xi_t = (\beta_t, \sigma_t)$ solution of the following SDE:

$$\begin{aligned} d\beta_t &= [\mu(\sigma_t) - (1 \wedge K_4\beta_t)D]\beta_t dt \\ &\quad + \sqrt{\frac{\mu(\sigma_t)\beta_t}{K_1}} dW_t^1 + \sqrt{\frac{(1 \wedge K_4\beta_t)^2 D\beta_t}{K_4}} dW_t^4 \end{aligned}$$

$$\begin{aligned} d\sigma_t &= [-(1 \wedge K_2\sigma_t)k\mu(\sigma_t)\beta_t + Ds^{in} - (1 \wedge K_5\sigma_t)D\sigma_t]dt \\ &\quad + \sqrt{\frac{(1 \wedge K_2\sigma_t)^2 k\mu(\sigma_t)\beta_t}{K_2}} dW_t^2 + \sqrt{\frac{Ds^{in}}{K_3}} dW_t^3 \\ &\quad + \sqrt{\frac{(1 \wedge K_5\sigma_t)^2 D\sigma_t}{K_5}} dW_t^5 \end{aligned}$$

where W_t^i are independent standard Wiener processes. Note that this result can be obtained directly from the process X_t without the help of the discrete-time approximation. Indeed the infinitesimal generator of process X_t given by (9) is a difference operator, and by Taylor development, it can be approximated by a second order differential operator corresponding to a diffusion process (Ethier and Kurtz, 1986).

For large K_i's, we can replace the terms $(1 \wedge K_4\beta_t)$, $(1 \wedge K_2\sigma_t)$, $(1 \wedge K_5\sigma_t)$ by 1; we can also group the Brownian motions and finally get:

$$d\beta_t = [\mu(\sigma_t) - D]\beta_t dt + \sqrt{\frac{\mu(\sigma_t)\beta_t}{K_1} + \frac{D\beta_t}{K_4}} dW_t^b \tag{17a}$$

$$\begin{aligned} d\sigma_t &= [-k\mu(\sigma_t)\beta_t + D(s^{in} - \sigma_t)]dt \\ &\quad + \sqrt{\frac{k\mu(\sigma_t)\beta_t}{K_2} + \frac{Ds^{in}}{K_3} + \frac{D\sigma_t}{K_5}} dW_t^s \end{aligned} \tag{17b}$$

where W^b and W^s are independent standard Wiener processes.

3.3.2. Behavior of the system of SDE's near the axes

System (16) is the Euler–Maruyama time discretization of the SDE (17) (for large K_i's). Even if the diffusion approximation is only valid for large values of the biomass and the substrate, we can study the behavior of (17) near the axes.

As for the discrete-time normal approximation, we should clarify the boundary conditions. As we well see, the component β_t given by (17a) will remain positive, but the component σ_t given by (17b) could become negative. We must first require that $\mu(s) = 0$ for $s < 0$. Then, note that each equation of (17) is related to the following well-known CIR model for interest rates:

Remark 3.1 *(Cox–Ingersoll–Ross model).* Consider the one-dimensional SDE:

$$dX_t = (a + bX_t)dt + \sigma\sqrt{X_t} dW_t, \quad X_0 = x_0 \geq 0. \tag{18}$$

with $a \geq 0$, $b \in \mathbb{R}$, $\sigma > 0$. According to (Lamberton and Lapeyre, 1996, Prop. 6.2.4), for all $x_0 \geq 0$, X is a continuous process taking values in $\mathbb{R}^+$, and let $\tau = \inf\{t \geq 0, X_t = 0\}$, then:

(i) If $a \geq \sigma^2/2$, then $\tau = \infty$ $\mathbb{P}_x$–a.s.;
(ii) if $0 \leq a < \sigma^2/2$ and $b \leq 0$ then $\tau < \infty$ $\mathbb{P}_x$–a.s.;
(iii) if $0 \leq a < \sigma^2/2$ and $b > 0$ then $\mathbb{P}_x(\tau < \infty) \in (0, 1)$.

In the first case, X never reaches 0. In the second case X a.s. reaches the state 0, in the third case it may reach 0. If $a = 0$ then the state 0 is absorbing.

It is clear that $\beta = 0$ is an absorbing state for (17a), and when $\beta = 0$, (17b) reduces to

$$d\sigma_t = D(s^{in} - \sigma_t)dt + \sqrt{\frac{Ds^{in}}{K_3} + \frac{D\sigma_t}{K_5}} dW_t^s$$

and from Remark 3.1, the solution of this SDE will stay on the half-line $[-(K_5/K_3)s^{in}, \infty)$ and:

(i) if $s^{in}((1/K_3) + (1/K_5)) \geq (1/2K_5^2)$ then σ_t never reaches $-(K_5/K_3)s^{in}$;
(ii) if $s^{in}((1/K_3) + (1/K_5)) < (1/2K_5^2)$ then σ_t reaches $-(K_5/K_3)s^{in}$ in finite time and is reflected.

Indeed, it is enough to apply It formula to $\tilde{\sigma}_t = (Ds^{in}/K_3) + (D\sigma_t/K_5)$ and to use Remark 3.1. Note that, as K_5 is large, condition (i) is more realistic than condition (ii).

To extend the definition of (17) for negative value of σ, let suppose that $\mu(\sigma) = 0$ for $\sigma \leq 0$. As we seen, β_t will stay non-negative and $\beta = 0$ is an absorbing state. Also $\sigma_t \geq -(K_5/K_3)s^{in}$ and for large K_5 this state will be repulsive. Note that for small values of σ_t, as the K_i are large, the diffusion term in (17b) will be small and the drift part will be dominated by Ds^{in} so that σ_t will increase fast and its probability to be negative will be small.

The fact that the substrate concentration could be "negative" is due to the normal approximation. This approximation is valid for large values of concentration and the validity of the diffusion system (17) is questionable for small concentration. Nonetheless we can study its properties.

A possibility to get an SDE with positive solution is to consider an SDE with boundary condition (Ikeda and Watanabe, 1981, §IV-7) by adding a local time in $\{\sigma = 0\}$ to the Eq. (17b). In the present work, we only consider that the solution of the system (17) remains in the domain $\mathcal{D} = [0, \infty) \times [-(K_5/K_3)s^{in}, \infty)$.

3.4. Asymptotic analysis

The convergence of the pure jump model (8) or of the diffusion approximation (17) to the deterministic model (1) as all the $K_i \to \infty$ can be rigorously established.

Let X_t^K be the pure jump model defined at the beginning of Section 3.1, or as the solution of the Eq. (8) for a given $K \stackrel{\text{def}}{=} (K_1, K_2, K_3, K_4, K_5)$. Let ξ_t^K be the solution of the SDE (17). Let $x(t)$ be the EDO model solution of Eq. (1). Then X_t^K converges toward $x(t)$ in the following way: for all $T > 0$ and all $\delta > 0$,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \left\| X_t^K - x(t) \right\| \geq \delta\right) \to 0$$

as $K_i \to \infty$ for all $i = 1 \ldots 5$. This result is not surprising if we consider the representation (11) of $(X_t)_{t \geq 0}$; it was obtained in a context of martingale convergence theorems in Kurtz (1970, 1971) or in a more general context of convergence of sequences of infinitesimal generators in Ethier and Kurtz (1986).

We can also prove the same type of convergence for the process ξ_t^K. Indeed, in Eq. (17) the scale coefficients appear as $1/\sqrt{K_i}$ in the diffusion part of the SDE, and the convergence clearly holds as all the K_i tends to infinity.

3.5. Other models

As already noticed, the choice of $(\lambda_i(x), \nu_i(x))$ satisfying (7) is not unique. We choose not to make the jump sizes depend on the state value x (except for the boundary conditions), only the jump

rates depend on x. Another possibility is to choose jump sizes that depend on the state value x. For example instead of the choice of Table 1, we can choose:

$$\lambda_1(x) = K_1\mu(s), \quad \lambda_2(x) = K_2 k\mu(s), \quad \lambda_3(x) = K_3 D, \quad \lambda_4(x) = K_4 D, \quad \lambda_5(x) = K_5 D$$

$$\nu_1(x) = \begin{pmatrix} \frac{b}{K_1} \\ 0 \end{pmatrix}, \quad \nu_2(x) = -\begin{pmatrix} 0 \\ \frac{b}{K_2} \end{pmatrix}, \quad \nu_3(x) = \begin{pmatrix} 0 \\ \frac{s^{\text{in}}}{K_3} \end{pmatrix},$$

$$\nu_4(x) = -\begin{pmatrix} \frac{b}{K_4} \\ 0 \end{pmatrix}, \quad \nu_5(x) = -\begin{pmatrix} 0 \\ \frac{s}{K_5} \end{pmatrix}$$

(if we neglect the boundary condition). Then in place of (17) we have the following set of equations:

$$\mathrm{d}\beta_t = [\mu(\sigma_t)\beta_t - D\beta_t]\mathrm{d}t + \sqrt{\frac{\mu(\sigma_t)}{K_1}}\beta_t \mathrm{d}W_t^1 + \sqrt{\frac{D}{K_4}}\beta_t \mathrm{d}W_t^4 \tag{19a}$$

$$\mathrm{d}\sigma_t = [-k\mu(\sigma_t)\beta_t + Ds^{\text{in}} - D\sigma_t]\mathrm{d}t + \sqrt{\frac{k\mu(\sigma_t)}{K_2}}\beta_t \mathrm{d}W_t^2 + \sqrt{\frac{D}{K_3}}s^{\text{in}}\mathrm{d}W_t^3 + \sqrt{\frac{D}{K_5}}\sigma_t \mathrm{d}W_t^5 \tag{19b}$$

where W^i are independent standard Wiener processes.

3.5.1. A three components model

Instead of the five components ①–⑤, we can consider a case with three independent sources of jump variation. This example strictly preserves the geometry (6) by considering three independent sources of jump variation: ①′ biology term: biomass increase and substrate decrease at scale K_1'; ②′ inflow term: substrate inflow at scale K_2'; ③′ outflow term: biomass and substrate outflow at scale K_3'. Again the jump sizes and rates should be chosen so as to satisfy the mass balance principle and the stochastic mass action, with no canonical choice. An ad hoc choice is

$$\lambda_1'(x) = K_1'\mu(s)b, \quad \lambda_2'(x) = K_2'D, \quad \lambda_3'(x) = K_3'D$$

$$\nu_1'(x) = \frac{1}{K_1'}\begin{pmatrix} 1 \\ -k \end{pmatrix}, \quad \nu_2'(x) = \frac{1}{K_2}\begin{pmatrix} 0 \\ s^{\text{in}} \end{pmatrix}, \quad \nu_3'(x) = \frac{1}{K_3'}\begin{pmatrix} -b \\ -s \end{pmatrix}$$

(if we neglect the boundary condition). This setting forces the jumps to be directed along the corresponding vector field, which is a strong constraint. In particular, the stoichiometry is strictly respected: the production of 1 unit of biomass requires *exactly* k units of substrate. Moreover the outflow jump is always radial, so that the increments of biomass and substrate are again strongly linked. Notice that for this particular choice of λ_3' and ν_3', the jump rate is constant but the jump size is not. In other words, the jump carries information both in the direction and the intensity of the variation. This will affect the qualitative behavior of the process and of its diffusion approximation, regarding extinction for example.

As for our canonical model, for large K_i's, we obtain a SDE for the diffusion approximation of the jump process:

$$\mathrm{d}\begin{pmatrix} \beta_t \\ \sigma_t \end{pmatrix} = \underbrace{\left[\mu(\sigma_t)\beta_t \begin{pmatrix} 1 \\ -k \end{pmatrix} + D\begin{pmatrix} 0 \\ s^{\text{in}} \end{pmatrix} - D\begin{pmatrix} \beta_t \\ \sigma_t \end{pmatrix}\right]}_{\text{biology} \quad \text{inflow} \quad \text{outflow}} \mathrm{d}t + \underbrace{\sqrt{\frac{\mu(\sigma_t)\beta_t}{K_1'}}\begin{pmatrix} 1 \\ -k \end{pmatrix}\mathrm{d}W_t^1}_{\text{biology}} + \underbrace{\sqrt{\frac{D}{K_2'}}\begin{pmatrix} 0 \\ s^{\text{in}} \end{pmatrix}\mathrm{d}W_t^2}_{\text{inflow}} + \underbrace{\sqrt{\frac{D}{K_3'}}\begin{pmatrix} \beta_t \\ \sigma_t \end{pmatrix}\mathrm{d}W_t^3}_{\text{outflow}}$$

that could be compared to (6). Notice that since W^1 and W^3 affect both components of ξ_t, the quadratic variation process $\langle\xi\rangle_t$ will not be a diagonal matrix. The diffusion term appears as the conjunction of three perturbations acting along the three vector fields determined by the sources of variation. Moreover, the intensity of the noise could be different for each type of perturbation. Considering this model could therefore be of interest, if the geometric interpretation of the noise is meaningful, see Joannides and Larramendy-Valverde (2010).

3.5.2. Comparison with the Imhof–Walcher model (Imhof and Walcher, 2005)

We finally mention that the diffusion model appearing in Imhof and Walcher (2005), is obtained from (19) by letting $K_1 = K_2 = K_3 = 0$ which leads to:

$$\mathrm{d}\beta_t = [\mu(\sigma_t) - D]\beta_t \mathrm{d}t + c^{\text{b}}\beta_t \mathrm{d}W_t^{\text{b}} \tag{20a}$$

$$\mathrm{d}\sigma_t = [-k\mu(\sigma_t)\beta_t + D(s^{\text{in}} - \sigma_t)]\mathrm{d}t + c^{\text{s}}\sigma_t \mathrm{d}W_t^{\text{s}} \tag{20b}$$

The choice of these coefficients is justified in Imhof and Walcher (2005) by constructing an approximating Markov chain, and then taking the limit as the sampling rate goes to 0. This model will be compared to the diffusion approximation (17) model on a simulation test in Section 5.3.

4. Simulation algorithms

We presented several models for the chemostat system: the pure jump model $(X_t)_{t\geq 0}$ could be considered as a detailed model at the microscopic scale. The Poisson approximation $(\tilde{X}_{t_n})_{n\in\mathbb{N}}$ given by (13) and the normal approximation $(\tilde{\xi}_{t_n})_{n\in\mathbb{N}}$ given (16) are constant time step approximation of the pure jump process. Finally the diffusion process $(\xi_t)_{t\geq 0}$ solution of the SDE (17) is a continuous time approximation of the pure jump process.

The now present the three associated simulation algorithms that will be valid at different scales.

4.1. Pure jump model

The pure jump model in continuous time described in Section 3.1 can be exactly simulated thanks to the Gillespie algorithm, also called stochastic simulation algorithm, described in Algorithm 4.1.

Algorithm 4.1. Gillespie algorithm (or stochastic simulation algorithm).

$t \leftarrow 0, x \leftarrow x_0$, save (t, x)
while $t \le T^{\max}$ **do**
  compute $\lambda_i(x)$ % see Table 1
  $\lambda = \sum_{i=1}^{5} \lambda_i(x)$
  $\Delta t \sim \mathrm{Exp}(\lambda)$ % exponential distribution
  $u \sim \mathrm{U}[0,1]$ % uniform distribution
  $t \leftarrow t + \Delta t$
  if $u \le \lambda_1(x)/\lambda$ **then**
    $x \leftarrow x + \bar{\nu}_1$ % biomass reproduction
  else if $u \le \{\lambda_1(x) + \lambda_2(x)\}/\lambda$ **then**
    $x \leftarrow [x - \bar{\nu}_2]_+$ % consumption
  else if $u \le \{\lambda_1(x) + \lambda_2(x) + \lambda_3(x)\}/\lambda$ **then**
    $x \leftarrow x + \bar{\nu}_3$ % substrate inflow
  else if $u \le \{\lambda_1(x) + \lambda_2(x) + \lambda_3(x) + \lambda_4(x)\}/\lambda$ **then**
    $x \leftarrow [x - \bar{\nu}_4]_+$ % biomass outflow
  else
    $x \leftarrow [x - \bar{\nu}_5]_+$ % substrate outflow
  end if
  save (t, x)
end while

Here $[x]_+$ is the projection on the positive quadrant: $[x]_+ = \left[\begin{pmatrix}\beta\\ \sigma\end{pmatrix}\right]_+ = \begin{pmatrix}\beta \vee 0\\ \sigma \vee 0\end{pmatrix}$.

When the rate coefficients $\lambda_i(x)$ are large the time increment will be small and the Gillespie algorithm is impractical. As the scale coefficients K_i are large, the $\lambda_i(x)$, $i \neq 3$, are large only when β and σ are small; $\lambda_3(x)$ will remain large as it does not depend on x.

4.2. Poisson approximation

Algorithm 4.2. Poisson approximation or tau-leap method.

$t \leftarrow 0, x \leftarrow x_0$, save (t, x)
while $t \le T^{\max}$ **do**
  compute $\lambda_i(x)$ % see Table 1
  $\lambda = \sum_{i=1}^{5} \lambda_i(x)$
  compute $\boldsymbol{m}_i(x), \boldsymbol{v}_i(x)$ % see (21)
  $\Delta t \leftarrow \min_{i=1\dots5}\{\varepsilon\lambda/|\boldsymbol{m}_i(x)|, \varepsilon^2\lambda^2/\boldsymbol{v}_i(x)\}$
  $t \leftarrow t + \Delta t$
  $\mathcal{P}_i \sim \mathrm{Poisson}(\lambda_i(x)\Delta t)$ for $i = 1 \dots 5$
  $x \leftarrow [x + \bar{\nu}_1\mathcal{P}_1 - \bar{\nu}_2\mathcal{P}_2 + \bar{\nu}_3\mathcal{P}_3 - \bar{\nu}_4\mathcal{P}_4 - \bar{\nu}_5\mathcal{P}_5]_+$
  save (t, x)
end while

Here $[x]_+$ is the projection on the positive quadrant: $[x]_+ = \left[\begin{pmatrix}\beta\\ \sigma\end{pmatrix}\right]_+ = \begin{pmatrix}\beta \vee 0\\ \sigma \vee 0\end{pmatrix}$.

The simulation of the previous model could be cumbersome for very high rates of event. In this case it is desirable to use the fixed time step Poisson approximation method (13) also called tau-leap (Gillespie, 2001). Recently many papers have addressed the numerical analysis of this approximation scheme (Rathinam et al., 2005; Li, 2007; Anderson et al., 2009). In this method the time step should be small enough so that it fulfills the following "leap condition": the state change in any leap should be small enough that no rate function $\lambda_i(x)$ will experience a macroscopically significant change in its value, that is: $|\lambda_i(x + \sum_{i'} \nu_{i'}(x)\mathcal{P}_n^{i'}(\Delta t\lambda_{i'}(x))) - \lambda_i(x)| \le \varepsilon\lambda(x)$ for $i = 1 \dots 5$, where $0 < \varepsilon \ll 1$ is an error control parameter.

For this method to be practicable, an automatic and simple way of determining the largest time step Δt compatible with the leap condition is proposed in Gillespie and Petzold (2003). Define:

$$\boldsymbol{m}_i(x) \overset{\text{def}}{=} \sum_{i'=1}^{5} \lambda_{i'}(x)(\nabla\lambda_i(x) \cdot \nu_{i'}), \quad \boldsymbol{v}_i(x) \overset{\text{def}}{=} \sum_{i'=1}^{5} \lambda_{i'}(x)(\nabla\lambda_i(x) \cdot \nu_{i'})^2 \tag{21}$$

for $i, i' = 1 \dots 5$, and let

$$\Delta t = \min_{i=1\dots5} \left\{ \frac{\varepsilon\lambda(x)}{|\boldsymbol{m}_i(x)|}, \frac{\varepsilon^2\lambda^2(x)}{|\boldsymbol{v}_i(x)|} \right\} \tag{22}$$

where ε is an error control parameter ($0 < \varepsilon \ll 1$), see Algorithm 4.2. Note that in the original context the jumps $\nu_i(x)$ do not depends on x, but in our situation they do not essentially depend on x, the dependence on x was introduced to handle the jump near the axes in order to avoid negative concentration (see Cao et al. (2005) for another possibility to overcome this difficulty).

4.3. Diffusion (normal) approximation

Algorithm 4.3. Normal approximation.

$t \leftarrow 0, (\beta, \sigma) \leftarrow (\beta_0, \sigma_0)$, save (t, β, σ)
while $t \le T^{\max}$ **do**
  $w^{\mathrm{b}} \sim \mathcal{N}(0,1), w^{\mathrm{s}} \sim \mathcal{N}(0,1)$
  $\beta' \leftarrow \beta + (\mu(\sigma) - D)\beta\Delta t + \sqrt{\frac{\mu(\sigma)\beta}{K_1} + \frac{D\beta}{K_4}}\sqrt{\Delta t}w^{\mathrm{b}}$
  $\sigma' \leftarrow \sigma + (-k\mu(\sigma)\beta + D(s^{\mathrm{in}} - \sigma))\Delta t + \sqrt{\frac{k\mu(\sigma)\beta}{K_2} + \frac{Ds^{\mathrm{in}}}{K_3} + \frac{D\sigma}{K_5}}\sqrt{\Delta t}w^{\mathrm{s}}$
  $\beta \leftarrow [\beta']^+$ % 0 is an absorbing state for β
  $\sigma \leftarrow |\sigma' - \sigma_{\min}| + \sigma_{\min}$ % reflection at $\sigma_{\min} = -\frac{K_5}{K_3}s^{\mathrm{in}}$ for σ
  $t \leftarrow t + \Delta t$
  save (t, β, σ)
end while

The normal approximation (16) can be slightly modified in order to take into account the qualitative behavior of the SDE (17) near the axes. We propose the following scheme:

$$\tilde{\beta}_{t_{n+1}} = \left[\tilde{\beta}_{t_n} + [\mu(\tilde{\sigma}_{t_n}) - (1 \wedge K_4\tilde{\beta}_{t_n})D]\tilde{\beta}_{t_n}\Delta t + \sqrt{\Delta t}\sqrt{\frac{\mu(\tilde{\sigma}_{t_n})\tilde{\beta}_{t_n}}{K_1} + \frac{(1 \wedge K_4\tilde{\beta}_{t_n})^2 D\tilde{\beta}_{t_n}}{K_4}}w_n^{\mathrm{b}} \right]^+, \tag{23a}$$

$$\tilde{\sigma}_{t_{n+1}} = \left| \tilde{\sigma}_{t_n} + [-(1 \wedge K_2\tilde{\sigma}_{t_n})k\mu(\tilde{\sigma}_{t_n})\tilde{\beta}_{t_n} + Ds^{\mathrm{in}} - (1 \wedge K_5\tilde{\sigma}_{t_n})D\tilde{\sigma}_{t_n}]\Delta t + \sqrt{\Delta t}\sqrt{\frac{(1 \wedge K_2\tilde{\sigma}_{t_n})^2 k\mu(\tilde{\sigma}_{t_n})\tilde{\beta}_{t_n}}{K_2} + \frac{Ds^{\mathrm{in}}}{K_3} + \frac{(1 \wedge K_5\tilde{\sigma}_{t_n})^2 D\tilde{\sigma}_{t_n}}{K_5}}w_n^{\mathrm{s}} - \sigma_{\min} \right| + \sigma_{\min}. \tag{23b}$$

Indeed as $\beta = 0$ is an absorbing state for the component β_t of the SDE, instead of the standard Euler–Maruyama (16a), we can use (23a) where $[\cdot]^+$ is the positive part and w_n^{b} are i.i.d. $\mathcal{N}(0,1)$ random variables.

Also, to take into account that the component σ is reflected in $\sigma_{\min} = -(K_5/K_3)s^{\mathrm{in}}$ we use the scheme (23b) where w_n^{s} are i.i.d. $\mathcal{N}(0,1)$ random variables. This discretization scheme was proposed in Diop (2003) in the context of the CIR diffusion process. In order to get a positive substrate concentration we can consider $\tilde{\sigma}_{t_n}^+ = \tilde{\sigma}_{t_n} \vee 0$ or let $\sigma_{\min} = 0$ in (23b). The simulation procedure is presented in Algorithm 4.3.

Remark 4.1 *(Scales and hybrid simulation).* The three algorithms proposed here are valid at different scales. In the Gillespie algorithm all the detailed microscopic jumps of the dynamics are simulated.

The idea of the Poisson approximation is to consider a time step Δ that should be small enough so that the different event rates barely evolve in the time interval $[t, t + \Delta t]$, but large enough for the approximation to be worthwhile. Starting in x at t, the time step Δt is given by (22) but if it is less than a few multiples of $1/\lambda(x)$ then the Gillespie algorithm should be preferred.

Table 2
Simulation cases.

Cases			K_1	K_2	K_3	K_4	K_5
0	Deterministic		∞	∞	∞	∞	∞
1	"Standard"	Case 1.1	10^4	10^6	10^6	10^4	10^6
		Case 1.2	10^5	10^7	10^7	10^5	10^7
		Case 1.3	10^7	10^9	10^9	10^7	10^9
2	"Unstirred inflow/outflows"	Case 2.1	10^6	10^6	10^4	10^4	10^4
		Case 2.2	10^7	10^7	10^5	10^5	10^5
		Case 2.3	10^9	10^9	10^7	10^7	10^7
3	"Fluid substrate"	Case 3.1	10^6	∞	∞	10^4	∞
		Case 3.2	10^7	∞	∞	10^5	∞
		Case 3.3	10^9	∞	∞	10^7	∞
4	"Biological only"	Case 4.1	10^4	10^6	∞	∞	∞
		Case 4.2	10^5	10^7	∞	∞	∞
		Case 4.3	10^7	10^9	∞	∞	∞

Here "$\infty = 10^{20}$".

Now the Poisson variables $\mathcal{P}_i$ of parameter $\lambda_i(x)\Delta t$ could be approximated by normal variables $\mathcal{N}(\lambda_i(x)\Delta t, \lambda_i(x)\Delta t)$ as soon as $\lambda_i(x)\Delta t \geq 20$.

The simulation method can automatically switch from one algorithm to another one according to the scale. We can also imagine that different components of the state vector are simulated with different algorithms.

5. Simulation study

We present simulation results of the discretized diffusion model (23) with Monod specific growth rates:

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

and parameters $k = 10$, $\mu_{\max} = 3$, $D = 0.12$, $s^{\text{in}} = 0.5$, $k_s = 6$. Simulations with the Haldane model are proposed in Campillo et al. (2010).

The ODE (1) is integrated with a Runge–Kutta[1] scheme but the Euler scheme, corresponding to (23) with $K_i = \infty$, gives very close results.

In addition to the deterministic case (case 0 with $K_i = \infty$ for all i), we consider 3 basic cases (see Table 2):

"Standard" scales: $K_{2,3,5} = 100 \times K_{1,4}$ corresponds to the "standard" case where the substrate concentration dynamics is closer to the deterministic case than the biomass concentration dynamics.
"Unstirred inflow/outflows" scales: $K_{1,2} = 100 \times K_{3,4,5}$ corresponds to the case where inflow and outflows are unstirred.
"Fluid substrate" scales: $K_{2,3,4} = \infty$, in this case the substrate Eq. (17b) is deterministic, i.e. the substrate dynamics is in fluid limit.
"Biological only" scales: $K_{3,4,5} = \infty$, in this case we consider that the randomness is only due to biological aspects of the system.

5.1. Law of the concentrations at a given time t

We perform Monte Carlo simulations to approximate the marginal densities of the biomass concentration B_t and of the substrate concentration S_t at a given time $t = 3$ (h). We consider the "standard" scales $K_{1,4} = 10^5$ and $K_{2,3,5} = 10^7$. Initial conditions are $B_0 = 0.026$ and $S_0 = 0.26$ which is quite far from the equilibrium state.

We compute $(S_t^{(j)}, B_t^{(j)})$ for $t = 3$ (h) for $j = 1 \ldots 20,000$ independent Monte Carlo trials of the pure jump process (Gillespie method), with the Poisson approximation (tau-leap method) and with the normal approximation. For the tau-leap method we choose a constant time step. For "Poisson 1" and "Normal 1" we use a step of 0.05, for "Poisson 2" and "Normal 2" we use a step of 0.5. We check, using a two-sample Kolmogorov–Smirnov test, if the sample from each of these four last cases matches the sample from the pure jump process (null hypothesis), the results of the test are:

Test	Δt	p-Value	Statistic
B_t Poisson	0.5	6×10^{-7}	0.02735
S_t Poisson	0.5	5×10^{-250}	0.1693
B_t Poisson	0.05	0.2851	0.00985
S_t Poisson	0.05	0.074897	0.0128
B_t normal	0.5	5×10^{-7}	0.0275
S_t normal	0.5	2×10^{-245}	0.1677
B_t normal	0.05	0.56716	0.00785
S_t normal	0.05	0.045172	0.01375

The statistic is the Kolmogorov–Smirnov distance $\sup_x |F_1(x) - F_2(x)|$ where F_i are the empirical cumulative distribution function of the corresponding sample. We also perform a single sample Kolmogorov–Smirnov test to check if the sample from the pure jump process matches a Gaussian distribution (null hypothesis), the result of the test is:

Test	p-Value	Statistic
B_t Gillespie	0.26235	0.0071138
S_t Gillespie	0.74484	0.0047973

For each of the previous cases we also compute the approximate PDF's (probability density functions) of S_t and B_t from the sample with a kernel method; and the Normal PDF's corresponding to the mean and variance of the pure jump process sample. We plot the different PDF's in Figure 2.

The conclusions are: The two approximations (Poisson and normal) are very close to the exact simulation of the pure jump process; the approximation with a larger step 0.5 is slightly different. The computation times[2] are: 5 h 45 min 32.6 s for the exact simulation of the pure jump process; 33.2 s (with the time step 0.05) and 4.6 s (with the time step 0.5) for the Poisson approximation; 0.7 s (with the time step 0.05) and 0.1 s (with the time step 0.5) for the normal approximation. Hence, in the present situation, where the parameters K_i are rather high, and for non-small concentration of the biomass and the substrate, *the exact simulation of the pure jump process (Gillespie method) should be avoided*. The resulting empirical densities are very close to normal densities and the solution of the ODE coincide with the mean of these normal densities.

5.2. About the scales parameters

As we have seen, for large populations, the diffusion approximation $\tilde{\xi}_{t_n} = (\tilde{\beta}_{t_n}, \tilde{\sigma}_{t_n})$ given by (23) is very close to the reference pure jump model $X_t = (B_t, S_t)$. So we now propose simulations of the diffusion approximation in the case of a Monod specific growth rate according to the scales scenarios of Table 2.

Figure 3 represents a simulation of a single trajectory in the 3 levels of scale: cases m, 1 to m, 3 (for $m = 1 \ldots 4$). Figure 4 represents the result of 10,000 Monte Carlo trials in the 3 levels of scale: cases i.1 and i.2 (for $i = 1 \ldots 4$). We represent the mean trajectory and the empirical law of $\tilde{\xi}_T$ at final time T.

We can conclude that, *at this level of population and scale*: The stochasticity is negligible only in the case 4 ("biological only") and at the highest scale level (cases i.3). The ODE solution $x(t)$ matches the (empirical) mean of the stochastic process at these scales (as the stochastic process is solution of a nonlinear equation, there is no reason for the mean of the stochastic process to coincide with the solution of the deterministic equations). Equivalent results have been obtained for the Haldane case.

[1] The routine `ode45` of `Matlab`, an explicit Runge–Kutta (4,5) formula.

[2] CPU time on a 2.13 GHz Intel Core 2 Duo with a RAM of 2 GB.

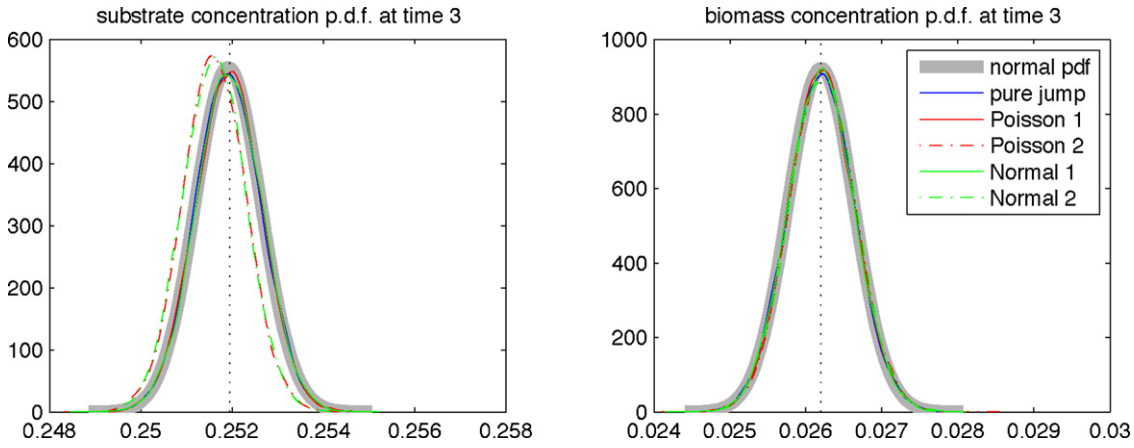

Fig. 2. Empirical densities for the substrate and the biomass concentrations at time $t = 3$ obtained with the exact simulation of the pure jump process (S_t, B_t) with the Gillespie method (blue line), with the Poisson approximation $(\tilde{S}_t, \tilde{B}_t)$ with constant time step (red solid line for a step 0.05, red dash line for a step 0.5), with the normal approximation $(\bar{\sigma}_t, \bar{\beta}_t)$ with constant time step (green solid line for a step 0.05, green dash line for a step 0.5). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

5.3. Comparison with the Imhof–Walcher model (Imhof and Walcher, 2005)

We compare the processes $\xi_t = (\beta_t, \sigma_t)$ given by the diffusion approximation model (17) with the one given by the ad hoc model (20). The parameter are: $K = 1$, $\mu_{\max} = 1$, $D = 0.5$, $S^{\text{in}} = 8$, $k_s = 0.5$, final time $T = 20$, $\Delta t = 0.02$, 20,000 Monte Carlo trials; $K_i = 10^5$ for (17) and $c^{\text{b}} = c^{\text{s}} = 0.02$ for (20). The parameters are chosen so that the biomass concentration evolves from 0.5 to about 7.5, and the substrate concentration from 5 to about 0.5. Also the final time distribution is lesser than 1 in the substrate and greater than 1 in the biomass. Indeed one of the main difference between (17) and (20) is than for state values less than 1 (resp. more greater than 1) the noise variance for the first model is greater (resp. lesser) than the noise variance for the second model. This example illustrates clearly that the two models differ substantially (see Fig. 5).

6. Discussion

We started from a reference pure jump model X_t, described by rates/jumps structure of Table 1 or as a solution of the stochastic differential equation (8). The martingale decomposition (11) clearly describes that the dynamics of X_t is the combination of the classical deterministic dynamics of the chemostat (1) plus martingale terms with coefficients $1/\sqrt{K_i}$ and with explicitly known quadratic variations, see (12). These quadratic variation terms allow us to assess the difference between the stochastic model and the deterministic one.

We presented the explicit Monte Carlo simulation procedure, called Gillespie method, for the process X_t. In standard cases, that is for large population sizes (i.e. K_i large), this procedure is not feasible as it requires us to simulate too many events. In this case, we presented the Poisson approximation (23b) and the normal approximation (23), both in discrete-time. These approximations are valid only for large populations, i.e. about the axes, it is necessary to return to the pure jump process representation. In the application discussed here, the Poisson approximation is of little interest: it is more time-consuming than the diffusion approximation and valid only on a very limited scale range between the pure jump model and the normal approximation model.

In contrast with previous stochastic chemostat models (Stephanopoulos et al., 1979; Imhof and Walcher, 2005) where the stochasticity was introduced according to an ad hoc approach, in the present work we propose a family of models where the structure of the noise emerges from the very dynamics and where the scale parameters can be tuned according to the problem under interest. In particular it allows us to propose hybrid models where the cell population dynamics features stochasticity as the substrate is in fluid dynamics (ODE), corresponding to the case 3 of Table 2. This kind of model has already been proposed in Grasman and De Gee (2005) in a three trophic levels case where the stochasticity appears only in the top level trophic as a stochastic logistic model and with fluid limit dynamics for the two other levels; it also has been proposed in Crump and O'Young (1979) with a pure jump process for the biomass dynamics and a fluid limit for the substrate.

The approach proposed here can be applied to any model of population dynamics especially in cases of difference of scale between the different dynamics (e.g. cell/substrate). The dynamics of interacting populations cannot be modeled by a single model but rather by a family of models whose domain of validity depends on the scale at which the dynamics are considered. For example the normal approximation model represented as stochastic differential equations (17) or the ODE model (1) are valid in high population sizes, hence using such models to infer extinction characteristics like extinction time and extinction probabilities is not valid. This was already noticed by Pollett (2001) and Wilcox and Possingham (2002).

In most standard population scales of the chemostat the ODE model is justified. Also, the ODE framework proposes analysis, control and optimization tools that are more accessible than the one of the SDE context. Though, as seen, the stochasticity cannot be neglected in many situations. This stochasticity could be of small intensity in the present single species/single substrate situation but could deeply perturb multiple species/multiple substrates situations. The SDE model could be simulated at a small extra computational cost and offers a more realistic prediction tool. Indeed, as it can account for the variability of the experiments, the simulation of the SDE offers the possibility to explore in depth the potentialities of the dynamical systems.

Stochastic models are also more adapted for the confrontation to the data as they allow us to build a statistical model and the associated likelihood function (Ross et al., 2006, 2009). For example, the growth curves are usually obtained by measuring equilibrium concentrations and by fitting these measurements to a given growth law via a least square procedure; the demographic variability obtained from the proposed stochastic models (e.g. the quasi-stationary distribution in Figure 4) should result in a more relevant fit. One of the next important steps, that we will investigate in coming work, will be to propose an adapted

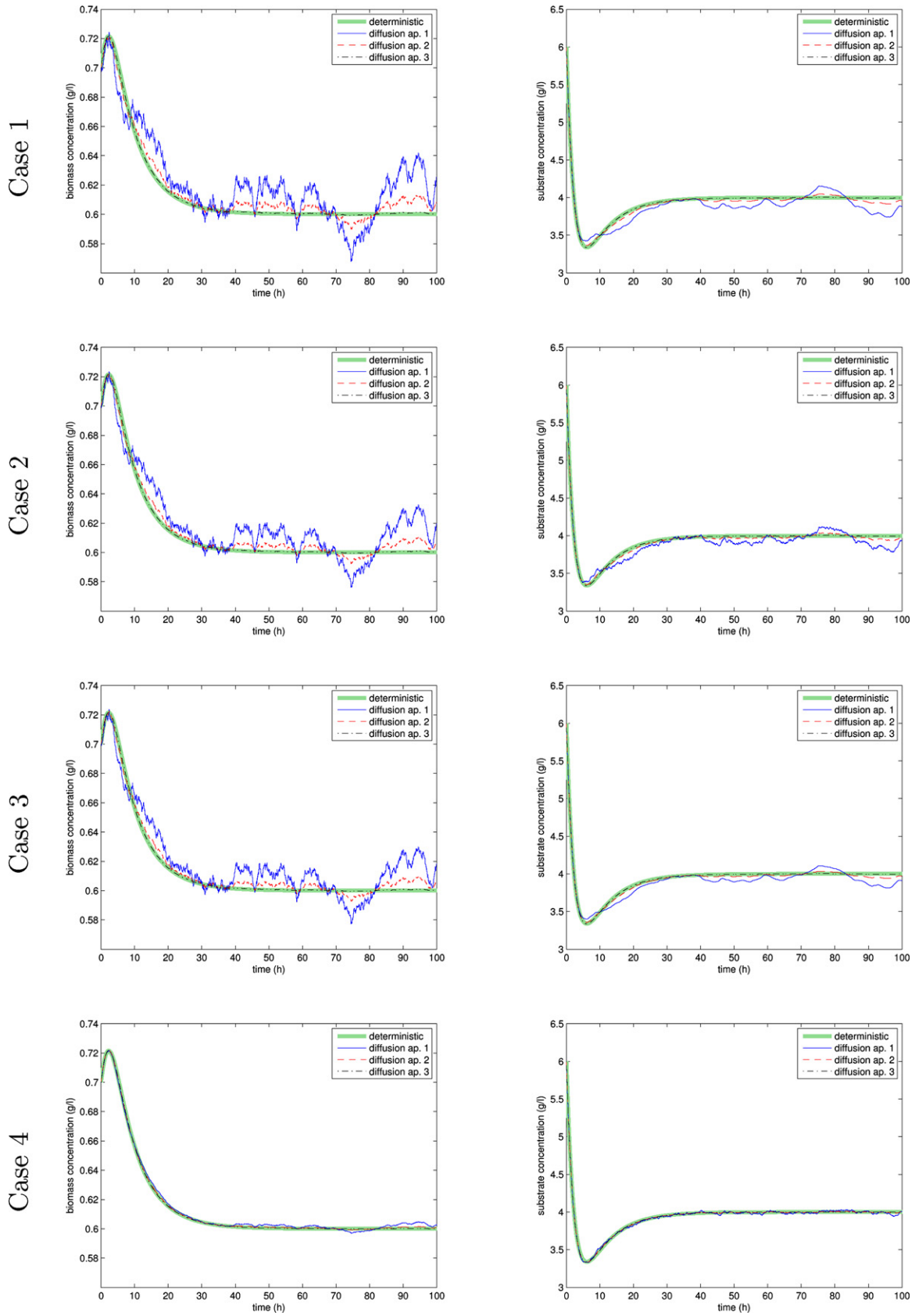

Fig. 3. Biomass and substrate concentration evolution/diffusion approximation, cases 1–4, Table 2/simulation of (23) – Time evolution of the biomass concentration (left), time evolution of the substrate concentration (right) according to 4 cases: case 0, case *i*.1, case *i*.2, case *i*.3 (see Table 2). Cases 0 (deterministic) and *i*.3 are almost identical.

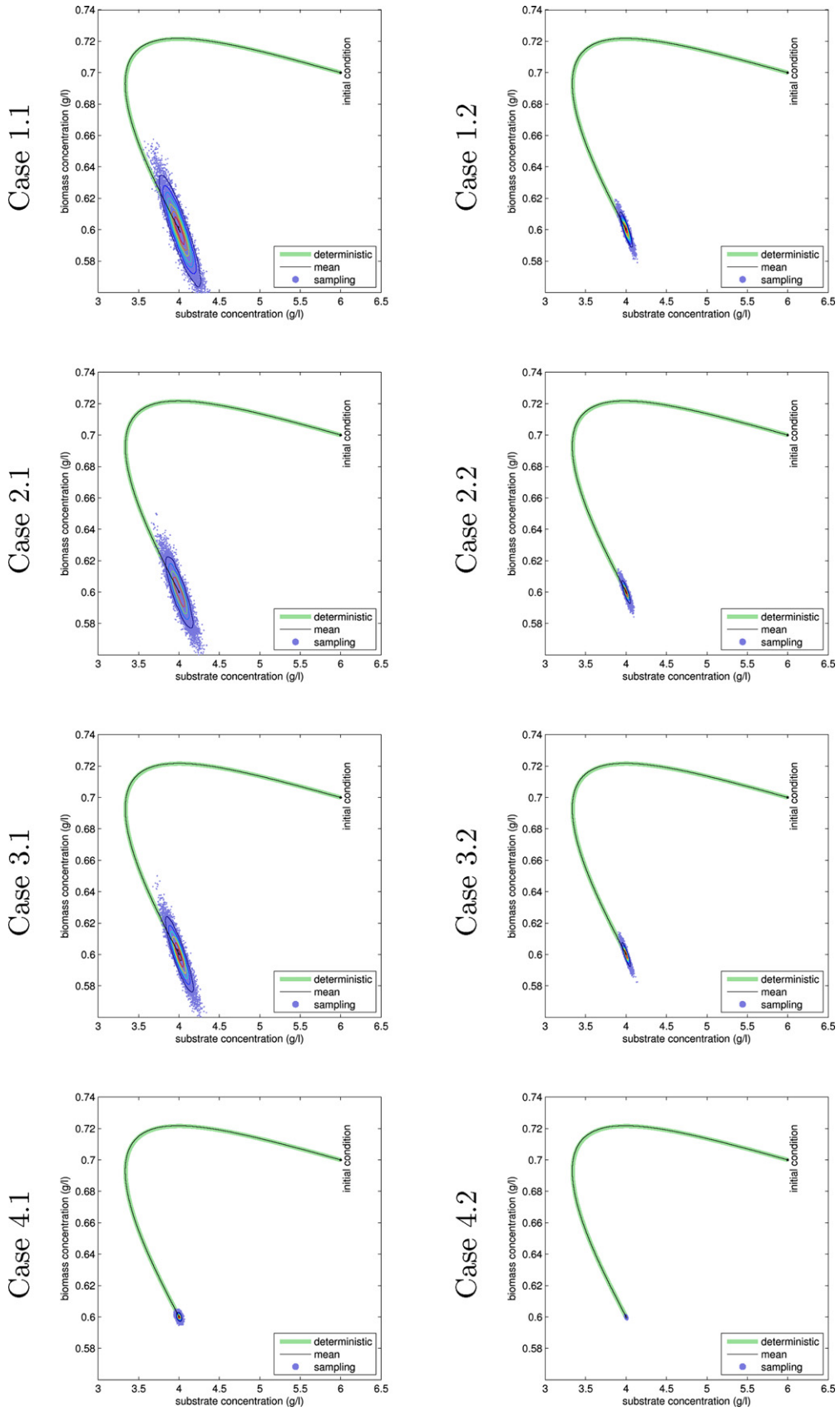

Fig. 4. Phase portraits/sampling 10,000 Monte Carlo trials of the law of $(\tilde{\beta}_{t_n}, \tilde{\sigma}_{t_n})$ for $t_n = 100$ according to the cases of Table 2 – the deterministic solution and the mean of the sampled trajectories coincide – the final time law is represented by the sample and by the contour plot of the corresponding kernel approximation of the p.d.f.

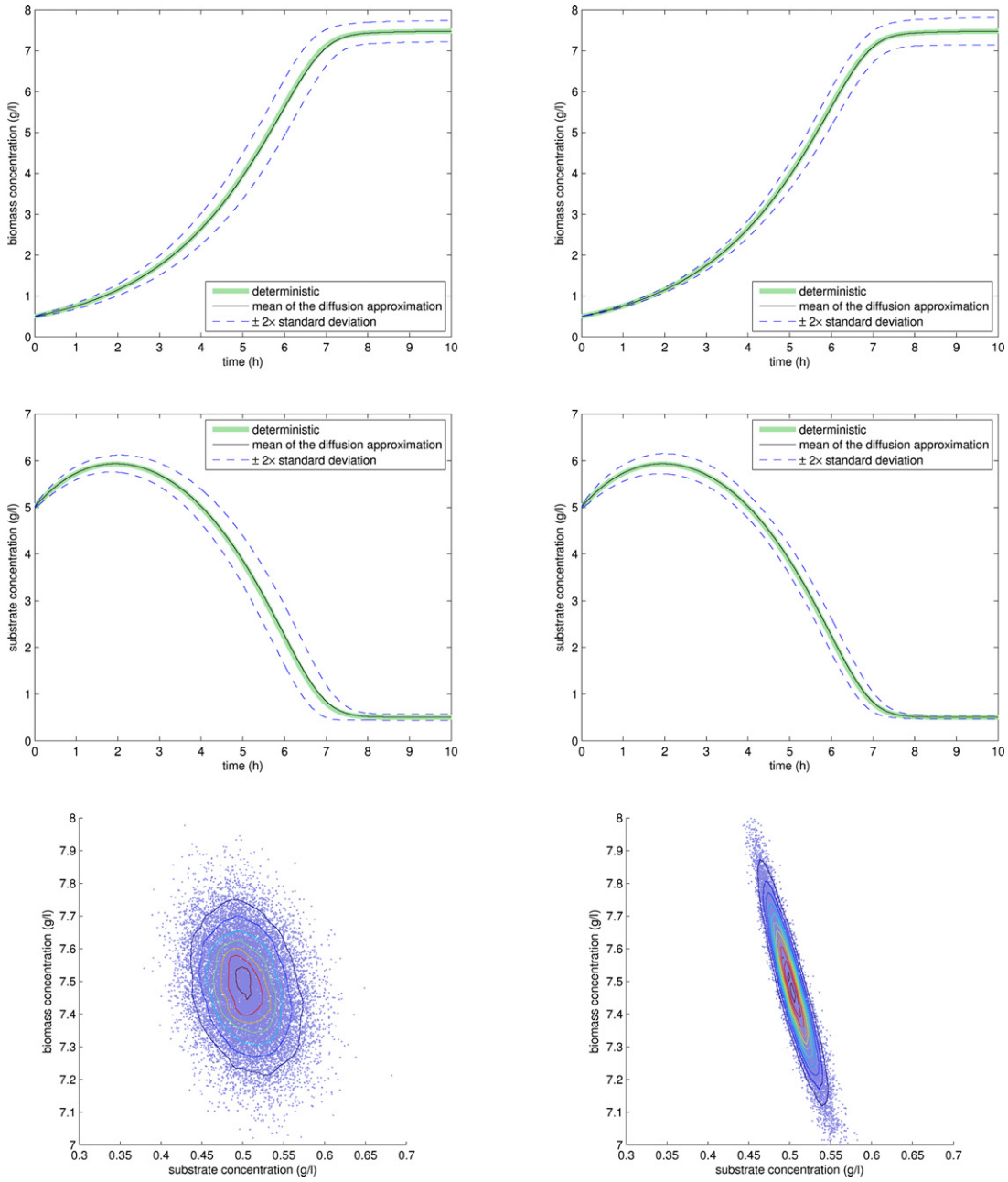

Fig. 5. Comparison of the processes $\xi_t = (\beta_t, \sigma_t)$ given by the diffusion approximation model (17) and by the ad hoc model (20). Evolution of the biomass concentration (top) and substrate concentration (center) during time; final joint density approximation of concentration (bottom). The two models differ substantially: compared to the diffusion approximation, the ad hoc model overestimates (resp. underestimates) the noise variance for concentration greater (resp. lesser) than 1.

statistical procedure to estimate the scale parameters K_i, and in a second step to estimate the parameters $(D, s^{\text{in}} \ldots)$. In the future we will also investigate the long-term behavior of these models as well as their optimal command.

Acknowledgements

The work was partially supported by the French National Research Agency (ANR) within the SYSCOMM project ANR-09-SYSC-003.

References

Anderson, D.F., Ganguly, A., Kurtz, T.G., 2009. Error analysis of tau-leap simulation methods. Annals of Applied Probability. arXiv:math.PR/0909.4790.

Campillo, F., Joannides, M., Larramendy, I., 2010. Stochastic models of the chemostat. Rapport de Recherche RR-7458, INRIA. http://hal.archives-ouvertes.fr/inria-00537886/.

Cao, Y., Gillespie, D.T., Petzold, L.R., 2005. Avoiding negative populations in explicit Poisson tau-leaping. Journal of Chemical Physics 123, 054104.

Crump, K.S., O'Young, W.-S.C., 1979. Some stochastic features of bacterial constant growth apparatus. Bulletin of Mathematical Biology 41 (1), 53–66.

Diop, A., 2003. Sur la discrétisation et le comportement à petit bruit d'eds unidimensionnelles dont les coefficients sont à dérivées singulières. Ph.D. thesis, Université de Nice-Sophia-Antipolis.

Ethier, S.N., Kurtz, T.G., 1986. Markov Processes – Characterization and Convergence. John Wiley & Sons.

Gillespie, D.T., 2000. The chemical Langevin equation. Journal of Chemical Physics 113 (1), 297–306.

Gillespie, D.T., 2001. Approximate accelerated stochastic simulation of chemically reacting systems. Journal of Chemical Physics 115 (4), 1716–1733.

Gillespie, D.T., Petzold, L.R., 2003. Improved leap-size selection for accelerated stochastic simulation. Journal of Chemical Physics 119, 8229–8234, http://dx.doi.org/10.1063/1.1613254.

Grasman, J., De Gee, M., 2005. Breakdown of a chemostat exposed to stochastic noise volume. Journal of Engineering Mathematics 53 (3), 291–300.
Herbert, D., Elsworth, R., Telling, R.C., 1956. The continuous culture of bacteria; a theoretical and experimental study. Journal of General Microbiology 14 (3), 601–622.
Ikeda, N., Watanabe, S., 1981. Stochastic Differential Equations and Diffusion Processes. North-Holland/Kodansha, Amsterdam.
Imhof, L., Walcher, S., 2005. Exclusion and persistence in deterministic and stochastic chemostat models. Journal of Differential Equations 217 (1), 26–53.
Joannides, M., Larramendy-Valverde, I., 2010. On geometry and scale of a stochastic chemostat. In: Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA).
Kurtz, T.G., 1970. Solutions of ordinary differential equations as limits of pure jump Markov processes. Journal of Applied Probability 7 (1), 49–58.
Kurtz, T.G., 1971. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. Journal of Applied Probability 8, 344–356.
Lamberton, D., Lapeyre, B., 1996. Introduction to Stochastic Calculus Applied to Finance. Chapman & Hall/CRC.
Li, T., 2007. Analysis of explicit tau-leaping schemes for simulating chemically reacting systems. Multiscale Modeling & Simulation 6 (2), 417–436.
Monod, J., 1950. La technique de la culture continue, thorie et applications. Annales de l'Institut Pasteur 79 (4), 390–410.
Norris, J.R., 1998. Markov Chains. Cambridge University Press.
Novick, A., Szilard, L., 1950. Description of the chemostat. Science 112 (2920), 715–716.
Pollett, P., 2001. Diffusion approximations for ecological models. In: Proceedings of the International Congress of Modelling and Simulation.
Rathinam, M., Petzold, L.R., Cao, Y., Gillespie, D.T., 2005. Consistency and stability of tau-leaping schemes for chemical reaction systems. Multiscale Modeling & Simulation 3, 867–895.
Ross, J.V., Taimre, T., Pollett, P.K., 2006. On parameter estimation in population models. Theoretical Population Biology 70 (4), 498–510.
Ross, J.V., Taimre, T., Pollett, P.K., 2009. On parameter estimation in population models II: multi-dimensional processes and transient dynamics. Theoretical Population Biology 75 (2–3), 123–132.
Smith, H.L., Waltman, P.E., 1995. The Theory of the Chemostat: Dynamics of Microbial Competition. Cambridge University Press.
Stephanopoulos, G., Aris, R., Fredrickson, A., 1979. A stochastic analysis of the growth of competing microbial populations in a continuous biochemical reactor. Mathematical Biosciences 45, 99–135.
Wilcox, C., Possingham, H., 2002. Do life history traits affect the accuracy of diffusion approximations for mean time to extinction? Ecological Applications 12 (4), 1163–1179.
Wilkinson, D.J., 2006. Stochastic Modelling for Systems Biology. Chapman & Hall/CRC.