



HAL
open science

On geometry and scale of a stochastic chemostat

Marc Joannides, Irène Larramendy-Valverde

► **To cite this version:**

Marc Joannides, Irène Larramendy-Valverde. On geometry and scale of a stochastic chemostat. *Communications in Statistics - Theory and Methods*, 2013, 42 (16), pp.2202-2211. 10.1080/03610926.2012.746983 . hal-00800908

HAL Id: hal-00800908

<https://hal.science/hal-00800908>

Submitted on 14 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON GEOMETRY AND SCALE OF A STOCHASTIC CHEMOSTAT

Marc Joannides and Irène Larramendy–Valverde

I3M UMR 5149

Université de Montpellier 2

Place Eugène Bataillon

34095 MONTPELLIER Cedex, France

Marc.Joannides@univ-montp2.fr

Irene.Larramendy-Valverde@univ-montp2.fr

Key Words: Chemostat; Diffusion approximation; Jump Markov process; Monte Carlo.

ABSTRACT

A chemostat is a fixed volume bioreactor in which micro–organisms are grown in a continuously renewed liquid medium. We propose a stochastic model for the evolution of the concentrations in the single species and single substrate case. It is obtained as a diffusion approximation of a pure jump Markov process, whose increments are comparable in mean with the deterministic model. A specific time scale, related to the noise intensity, is considered for each source of variation. The geometric structure of the problem, usable by identification procedures, is preserved both in the drift and diffusion term. We study the properties of this model by numerical experiments.

1. INTRODUCTION

1.1. THE CHEMOSTAT

The chemostat (chemical environment is static) is a laboratory device used to study the growth of micro–organisms like yeast or bacteria. It consists in a growth chamber populated with one or more species in a liquid medium of fixed volume that is continuously renewed, see Figure 1. The inflow contains the nutrient used by the bacteria to grow and reproduce while the outflow removes both the biomass and the substrate at the same rate, so that the volume

is kept unchanged. This leads naturally to a chemical equilibrium of the culture parameter (substrate concentration, pH, ...). When steady state is established, the biomass removed is exactly balanced by the increase of biomass due to the growth of micro-organisms. We consider here a single species B which uses the substrate S as its nutrient to grow and reproduce. The experimental conditions are determined by the substrate concentration in the influent S^{in} and by the dilution rate $D = \frac{Q}{V}$, where Q is the constant pump rate and V is the constant volume. The special case of *washout* occurs when the dilution rate is so fast that the increase of biomass within the growth chamber is not sufficient to balance the output, so that all micro-organisms are eventually removed from the chemostat. The quantities of interest chosen to characterize the system will be the concentrations of the substrate and of the biomass at each time. This state can only vary through the effect of the two mechanic actions (inflow and outflow) and one biological transformation (increase of biomass by the consumption of substrate).

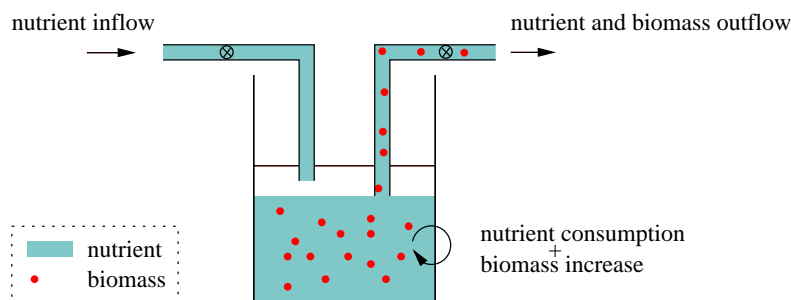


Figure 1: Operating principle of a chemostat: the growth chamber is provided with a sterile liquid medium with a constant substrate concentration. The biomass is partly evacuated with the outflow

1.2. DETERMINISTIC MODEL

The system is classically described by a system of differential equations based on a mass-balance principle, see Smith and Waltman (1995). We denote by b_t and s_t the respective concentrations of the biomass and of the substrate. Writing the balance for each quantity

yields the system of ODE

$$\dot{b}_t = \mu(s_t) b_t - D b_t \quad (1)$$

$$\dot{s}_t = -k \mu(s_t) b_t + D S^{\text{in}} - D s_t, \quad (2)$$

where $\mu(s_t)$ is the (bounded) *specific growth rate* of the species with limiting factor S and k is a stoichiometric coefficient. In equation (1), the first (resp. last) term accounts for the increase of biomass due to growth (resp. biomass outflow). Similarly in equation (2) the first (resp. two last) term accounts for the quantity of substrate consumed by the micro-organisms (resp. substrate inflow and outflow). Numerous models have been proposed for the specific growth rate, among which the Monod model (uninhibited growth) and the Haldane model (inhibited growth) are the most commonly used. They read respectively

$$\mu(s) = \mu_{\max} \frac{s}{K_S + s} \quad \text{and} \quad \mu(s) = \bar{\mu} \frac{s}{K_S + s + \frac{s^2}{K_i}}$$

where μ_{\max} , K_S and K_i are unknown parameters, that can be either measured experimentally or estimated by identification procedures, see Bastin and Dochain (1990).

To emphasize the geometric structure of the system, we write system (1) and (2) under the vector form

$$\begin{pmatrix} \dot{b}_t \\ \dot{s}_t \end{pmatrix} = r(b_t, s_t) \begin{pmatrix} 1 \\ -k \end{pmatrix} + D \begin{pmatrix} 0 \\ S^{\text{in}} \end{pmatrix} - D \begin{pmatrix} b_t \\ s_t \end{pmatrix}. \quad (3)$$

with *reaction kinetics* $r(s_t, b_t) = \mu(s_t) b_t$, that exhibits the three vector fields corresponding to the three sources of variations, namely *biological transformation*, *nutrient input* and *biomass and nutrient output*.

To conclude this introduction, we notice that this deterministic description is based on the hypothesis that there is no stochastic fluctuation, or at least that it can be neglected.

2. PURE JUMP MARKOV MODEL

Although the deterministic approach is widespread, there is a need to take into account the stochastic fluctuations inescapable when living organisms are involved. Moreover, a stochastic model is required in order to exploit the information contained in the *demographic*

noise for parameter estimation purpose. Beside, it is questionable whether a continuously varying concentration is still appropriate as far as washout is concerned. Indeed, the population of micro-organism in the apparatus could become so small that its concentration would no longer be considered to be varying continuously. Various attempts to introduce a noise component in system (3) have been proposed, see e.g. Stephanopoulos *et al.* (1979) or Imhof and Walcher (2005). This can be done by adding a diffusion coefficient to (3), or by adding a noise term to its discretized version. However, the *geometric structure* of the problem should be preserved by the perturbed system. Moreover, we should have in mind the discrete nature of the real state. For these reasons, we propose to modelize the phenomenon by a pure jump Markov process with three types of transitions associated with the three sources of variation. In addition, we require that mass balance on which (3) is based should be satisfied *in mean*. Campillo *et al.* (2011) considered a similar approach to investigate the relation between models at different scales.

2.1. DYNAMIC DESCRIPTION

We introduce the pure jump Markov process $(B_t, S_t)_{t \geq 0}$ whose dynamics is described by

$$\mathbb{P}[(B_{t+h}, S_{t+h}) = (b', s') \mid (B_t, S_t) = (b, s)] = \begin{cases} h K^b r(b, s) + o(h) & \text{if } (b', s') = (b, s) + \frac{1}{K^b} (1, -k) \\ h K^{\text{in}} D + o(h) & \text{if } (b', s') = (b, s) + \frac{1}{K^{\text{in}}} (0, S^{\text{in}}) \\ h K^{\text{out}} D + o(h) & \text{if } (b', s') = (b, s) + \frac{1}{K^{\text{out}}} (-b, -s) \\ 1 - h [K^b r(b, s) + D (K^{\text{in}} + K^{\text{out}})] + o(h) & \text{if } (b', s') = (b, s) \\ 0 & \text{otherwise} \end{cases}$$

where K^b , K^{in} and K^{out} are scaling constants. These parameters control both the rate and the size of the jumps in such a way that the instantaneous mean equals the right hand side of (3). The process evolves only by jumps when an event E_Δ occurs, for $\Delta \in \{\text{b, in, out}\}$, corresponding respectively to the biological transformation, the inflow and the outflow. The rates $\lambda_\Delta(b, s)$ and the sizes $y_\Delta(b, s)$ of the jumps from a state (b, s) are summarized in table 1.

E_Δ	E_b	E_{in}	E_{out}
$\lambda_\Delta(b, s)$	$K^b r(b, s)$	$K^{\text{in}} D$	$K^{\text{out}} D$
$y_\Delta(b, s)$	$\frac{1}{K^b} \begin{pmatrix} 1 \\ -k \end{pmatrix}$	$\frac{1}{K^{\text{in}}} \begin{pmatrix} 0 \\ S^{\text{in}} \end{pmatrix}$	$\frac{1}{K^{\text{out}}} \begin{pmatrix} -b \\ -s \end{pmatrix}$

Table 1: Rates and jumps of the three type of events in vector form.

We also set to 0 the rates leading to a non admissible transition, that is a state outside the positive orthant.

We now give a description more suitable for comparison with deterministic model (3).

2.2. SEMIMARTINGALE REPRESENTATION

The infinitesimal generator of the process described above reads

$$Af(b, s) = \sum_{\Delta \in \{b, \text{in}, \text{out}\}} \lambda_\Delta(b, s) [f((b, s) + y_\Delta(b, s)) - f(b, s)] , \quad (4)$$

for any f in its domain. For such an f , we have a semimartingale representation thanks to the Dynkin's formula

$$f(B_t, S_t) = f(B_0, S_0) + \int_0^t Af(B_s, S_s) ds + M_t^f$$

where M_t^f is a martingale. This formula remains valid for a wider class of functions, even unbounded, provided some integrability condition holds, see e.g. Hamza and Klebaner (1995) or Theorem 1.19 of Klebaner (1998). Since the specific growth rate is bounded, there exists $C > 0$ such that

$$\sum_{\Delta \in \{b, \text{in}, \text{out}\}} \lambda_\Delta(b, s) |y_\Delta(b, s)| \leq C (1 + |(b, s)|) .$$

By Theorem 1.19 of Klebaner (1998) applied to the components of the identity function, we

get the semimartingale representation for the process itself

$$\begin{aligned} B_t &= B_0 + \int_0^t [r(B_s, S_s) - D B_s] ds + M_t^B \\ S_t &= S_0 + \int_0^t [-k r(B_s, S_s) + D S^{\text{in}} - D S_s] ds + M_t^S \end{aligned}$$

which is the integral form of the SDE

$$\begin{pmatrix} dB_t \\ dS_t \end{pmatrix} = \left[r(B_t, S_t) \begin{pmatrix} 1 \\ -k \end{pmatrix} + D \begin{pmatrix} 0 \\ S^{\text{in}} \end{pmatrix} - D \begin{pmatrix} B_t \\ S_t \end{pmatrix} \right] dt + \begin{pmatrix} dM_t^B \\ dM_t^S \end{pmatrix}. \quad (5)$$

We see that the dynamics of our process is now written as the sum of the drift appearing in (3) and of a martingale term carrying the stochastic perturbation. In particular, we can account for the additional information contained in the quadratic variation of this martingale in order to estimate the unknown parameters of the model.

3. DIFFUSION APPROXIMATION

When all the scaling parameters K^{b} , K^{in} and K^{out} are large, the process evolves by *small* but *frequent* jumps. In that case a diffusion approximation can be considered, see Ethier and Kurtz (1986) or Wilkinson (2006). Replacing the increments of f in (4) by a Taylor's expansion and dropping the terms of order greater than two gives, for $x = (b, s)$

$$\begin{aligned} \tilde{A}f(x) &:= \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_{\Delta}(x) \left[\nabla f(x)^* \cdot y_{\Delta}(x) + \frac{1}{2} y_{\Delta}^*(x) \cdot H_f(x) \cdot y_{\Delta}(x) \right] \\ &= \nabla f(x)^* \cdot \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_{\Delta}(x) y_{\Delta}(x) + \frac{1}{2} \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_{\Delta}(x) y_{\Delta}^*(x) \cdot H_f(x) \cdot y_{\Delta}(x) \end{aligned}$$

where $H_f(x)$ denotes the Hessian matrix of f . \tilde{A} is the generator of a diffusion process $\tilde{X}_t = (\tilde{B}_t, \tilde{S}_t)$ which is solution of the SDE:

$$d\tilde{X}_t = \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_{\Delta}(\tilde{X}_t) y_{\Delta}(\tilde{X}_t) dt + \sum_{\Delta \in \{\text{b, in, out}\}} \sqrt{\lambda_{\Delta}(\tilde{X}_t)} y_{\Delta}(\tilde{X}_t) dW_t^{\Delta}$$

with independent standard brownian motions W^{b} , W^{in} and W^{out} . Expanding the sums

yields the vector form

$$\begin{pmatrix} d\tilde{B}_t \\ d\tilde{S}_t \end{pmatrix} = \left[r(\tilde{B}_t, \tilde{S}_t) \begin{pmatrix} 1 \\ -k \end{pmatrix} + D \begin{pmatrix} 0 \\ S^{\text{in}} \end{pmatrix} - D \begin{pmatrix} \tilde{B}_t \\ \tilde{S}_t \end{pmatrix} \right] dt \\ + \sqrt{\frac{r(\tilde{B}_t, \tilde{S}_t)}{K^{\text{b}}}} \begin{pmatrix} 1 \\ -k \end{pmatrix} dW_t^{\text{b}} + \sqrt{\frac{D}{K^{\text{in}}}} \begin{pmatrix} 0 \\ S^{\text{in}} \end{pmatrix} dW_t^{\text{in}} + \sqrt{\frac{D}{K^{\text{out}}}} \begin{pmatrix} \tilde{B}_t \\ \tilde{S}_t \end{pmatrix} dW_t^{\text{out}} . \quad (6)$$

This last form is much comparable with (3). The trajectories are continuous (almost surely) and the drift term is the same. Moreover the geometric structure is preserved by the diffusion term. Indeed, the stochastic perturbation in the diffusion term appears as a sum of three independent gaussian noises, each one acting along a vector field corresponding to a source of variation. It should be noted that this diffusion model should be used away from the axis. In particular, studying extinction time would not make sense, see Pollett (2001). We see that the scaling parameters K^Δ can be reinterpreted as noise intensity on the sources. Notice also that the K^Δ could be of different magnitude order. Eventually, the diffusion coefficient vanishes as the scaling parameters tend to infinity, so that the deterministic model can be viewed as the small noise limit of the diffusion model (6). Since the drift term is nonlinear, the deterministic model (3) *is not* the mean of the diffusion model (6). Of course, it is possible to rewrite (6) as an SDE driven by a single two-dimensional brownian motion. However this would break the geometric understanding of the dynamics given by (6).

4. SIMULATION ALGORITHMS

The pure jump model is classically simulated with the stochastic simulation (Gillespie) algorithm, see Gillespie (1977), described below:

1. Initialization: let $(b, s) \leftarrow (b_0, s_0)$ and $t \leftarrow 0$
2. while $t < T_{\text{max}}$
 - compute global rate: $\lambda(b, s) = \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_\Delta(b, s)$
 - compute next time event (exponential): $t \leftarrow t + \mathcal{E}(\lambda(b, s))$

- jump: $(b, s) \leftarrow (b, s) + y_\Delta(b, s)$, where y_Δ is chosen with probability $\frac{\lambda_\Delta}{\lambda}(b, s)$.

For large scaling parameters, all rates may be so fast that the procedure described above becomes unnecessarily slow. A number of variants have been proposed to speed up the procedure (review in Wilkinson (2006)), mostly based on the approximation

$$(B_{t+h}, S_{t+h}) \simeq (B_t, S_t) + \sum_{\Delta \in \{\text{b, in, out}\}} N_\Delta y_\Delta(B_t, S_t)$$

where N_Δ denotes the number of event E_Δ that have occurred within $[t, t+h]$. We introduce

Assumption (i): The step size h is supposed to be small enough so that the rates λ_Δ do not vary significantly in the interval.

In that case, N_Δ are Poisson variables of respective parameters $\lambda_\Delta(B_t, S_t) h$. This leads to the *Poisson timestep method*, described by the algorithm:

1. Initialization: let $(b, s) \leftarrow (b_0, s_0)$ and $t \leftarrow 0$
2. while $t < T_{\max}$
 - For $\Delta \in \{\text{b, in, out}\}$, draw $N_\Delta \sim \mathcal{P}(\lambda_\Delta(b, s) h)$
 - Commit events: $(b, s) \leftarrow (b, s) + \sum_{\Delta \in \{\text{b, in, out}\}} N_\Delta y_\Delta(b, s)$
 - Increment time: $t \leftarrow t + h$

The timestep can also be adaptive, as in the *tau-leap method*, see Gillespie (2001).

Still following Gillespie (2000), we note that the diffusion process introduced above via a Taylor expansion, appears also as a numerical approximation of the jump process. Indeed, consider now

Assumption (ii): The timestep h is sufficiently large so that many events have occurred within $[t; t+h]$.

We can then use the normal approximation of the Poisson law, to get

$$\mathcal{P}(\lambda_\Delta(b, s) h) \simeq \lambda_\Delta(b, s) h + \sqrt{\lambda_\Delta(b, s)} \sqrt{h} \mathcal{N}^\Delta(0, 1)$$

k	μ_{\max}	K_S	D	S^{in}	K^{b}	K^{in}	K^{out}
10	3 h^{-1}	6 g/l	0.12 h^{-1}	0.5 g/l	10^7	10^5	10^5

Table 2: Parameter values

and

$$\begin{aligned}
(B_{t+h}, S_{t+h}) &\simeq (B_t, S_t) + h \sum_{\Delta \in \{\text{b, in, out}\}} \lambda_{\Delta}(b, s) y_{\Delta}(b, s) \\
&\quad + \sqrt{h} \sum_{\Delta \in \{\text{b, in, out}\}} \sqrt{\lambda_{\Delta}(b, s)} \mathcal{N}^{\Delta}(0, 1) y_{\Delta}(b, s)
\end{aligned}$$

which is nothing else but a Euler discretization scheme applied to the SDE (6). As a result, numerical solutions of $(\tilde{B}_t, \tilde{S}_t)$ obtained by such a scheme will have approximately same behaviour as (B_t, S_t) sampled with time step h . However, the choice of the time step remains problematic since it has to meet the two antagonist requirements (i) and (ii).

The numerical simulations presented below use a Monod model for the growth rate. Table 2 shows the values of the parameters. We use the Euler–Maruyama scheme to simulate the solutions of the SDE involved, see Kloeden and Platen (1992). Fig. 2 and Fig. 3 show that for sufficiently large K^{Δ} , the jump process and its diffusion approximation are very similar.

Fig. 4 illustrates the impact of the demographic noise on the equilibrium. It is hoped that statistical procedures can make use of the information contained in the variability.

5. CONCLUSION

In this paper we have presented a way to account for stochastic fluctuations in a simple chemostat, while preserving the geometric structure. Randomness is first considered through a pure jump Markov process whose infinitesimal increments agree with the classical deterministic model. A different time scale is possible for the different types of events. Using integrability conditions, we obtained the stochastic differential equation satisfied by this process.

Even if the fundamental structure of the system is discrete, it is reasonable to describe it by a process with continuous trajectories. We therefore introduced the diffusion approximation of the jump process, still preserving the geometry. The constants corresponding to the specific time scales can then be interpreted as the intensities of the independent noises affecting each source of variation. Numerical experiments showed that the approximation is safe provided that the system is far from washout. It is expected that the geometry of the process will lead to efficient statistical procedures.

BIBLIOGRAPHY

Bastin, G. and Dochain, D. (1990). *On-line estimation and adaptive control of bioreactors*. Amsterdam: Elsevier.

Campillo, F. and Joannides M. and Larramendy-Valverde, I., (2011). Stochastic modeling of the chemostat. *Ecol. Model.*, **222(15)**, 2676–2689.

Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes – Characterization and Convergence*. New York: John Wiley & Sons.

Gillespie, D. T. (2000). The chemical langevin equation. *J. Chem. Phys.*, **113**, 297–306.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81(25)**, 2340–2361.

Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. of Chem. Phys.* **115**, 1716–1733.

Hamza, K. and Klebaner, F. C. (1995). Conditions for integrability of markov chains. *J. Appl. Probab.* **32(2)**, 541–547.

Imhof, J. and Walcher, S. (2005). Exclusion and persistence in deterministic and stochastic chemostat models. *J. Differ. Equations* **217(1)**, 26–53.

Klebaner, F. C. (1998). *Introduction to Stochastic Calculus with Applications*. London:

Imperial College Press.

Kloeden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Volume 23 of *Applications of Mathematics*. New York: Springer Verlag.

Pollett, P. (2001). Diffusion approximations for ecological models. In *Proceedings of the International Congress of Modelling and Simulation*.

Smith, H. L., Waltman, P. E. (1995). *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge University Press.

Solari, H. G. and Natiello, M. A. (2003). Stochastic population dynamics: The poisson approximation. *Phys. Rev. E* **67(3)**, 031918 .

Stephanopoulos, G., Aris, R. and Fredrickson, A. G. (1979). A stochastic analysis of the growth of competing microbial populations in a continuous biochemical reactor. *Math. Biosci.* **45**, 99–135

Wilkinson, D. J. (2006). *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC.

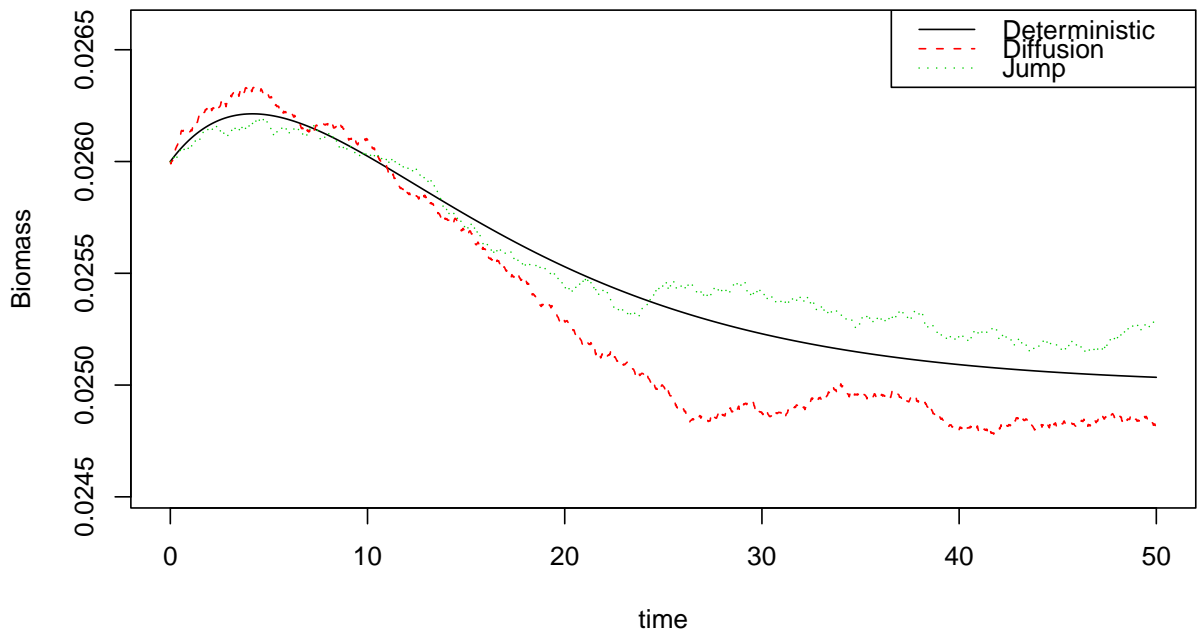


Figure 2: Evolution of the biomass concentration for the deterministic model b_t , the diffusion–approximation \tilde{B}_t and the jump process B_t .

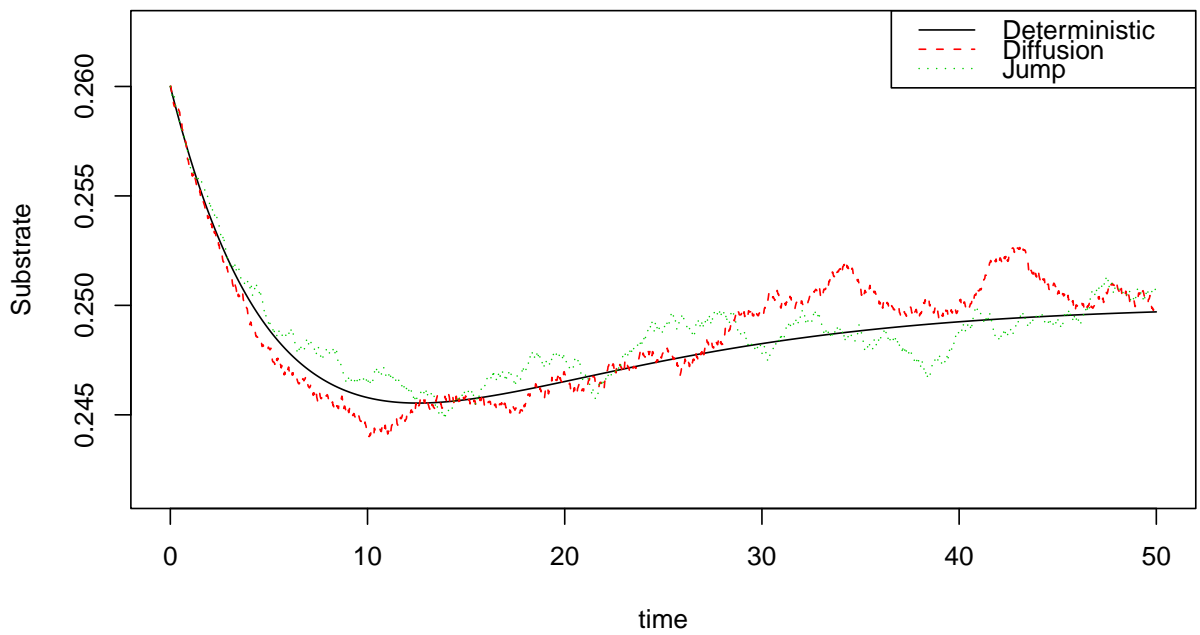


Figure 3: Evolution of the substrate concentration for the deterministic model s_t , the diffusion–approximation \tilde{S}_t and the jump process S_t .

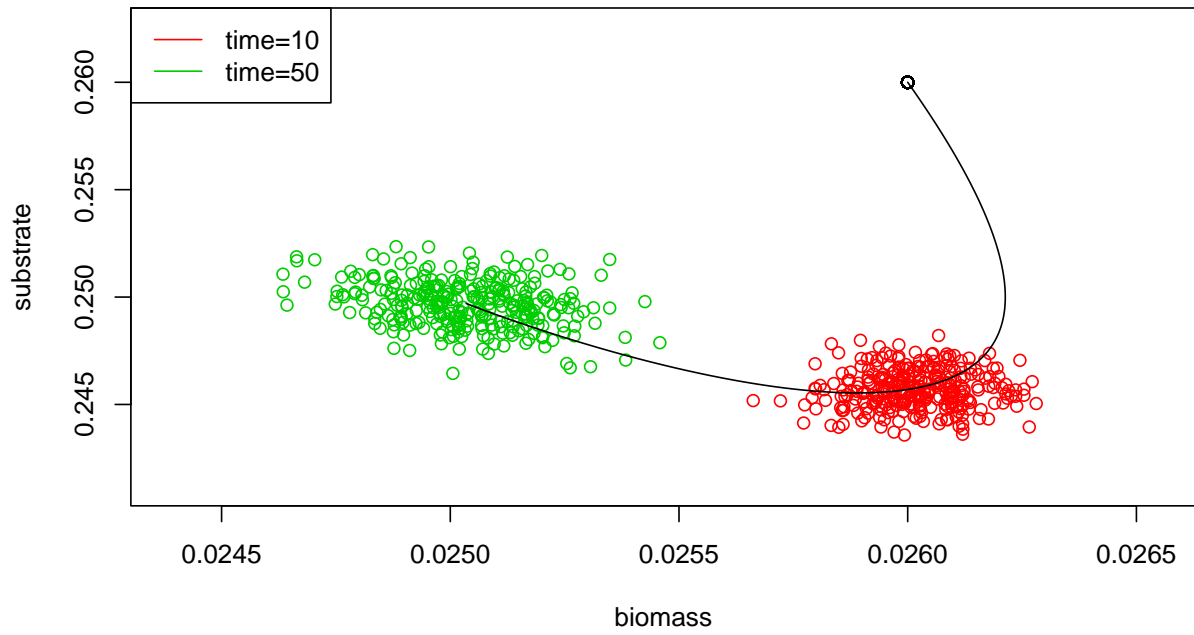


Figure 4: Particle cloud of 300 Monte Carlo runs of the diffusion–approximation $(\tilde{B}_t, \tilde{S}_t)$ at different time.