



Stochastic modeling of the anaeorobic model AM2b

Fabien Campillo, Mohsen Chebbi, Salwa Toumi

► To cite this version:

Fabien Campillo, Mohsen Chebbi, Salwa Toumi. Stochastic modeling of the anaeorobic model AM2b: Models at different scales. 13th Africain Conference on Research in Computer Science and Applied Mathematics (CARI 2016), Oct 2016, Tunis, Tunisia. hal-01406450

HAL Id: hal-01406450

<https://inria.hal.science/hal-01406450>

Submitted on 1 Dec 2016

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.

Stochastic models recently gain more credibility and numerical efficiency in chemistry [7], biotechnology [8], system biology [12] where deterministic models have been extensively used. Taking the example of a biotechnological model, we explain how a stochastic modeling approach deepens the insights allowed by the deterministic classical models.

Wastewater treatment plant aims at reducing the volume of pollutants rejection, producing potential energy like CH_4 in anaerobic treatment, providing treated water for agriculture and industry. Among these technologies anaerobic membrane bioReactors are promising technologies provided that membrane fouling phenomenon could be reduced. AM2b is a mathematical model of anaerobic membrane bioreactors developed by [1, 2], it is a variant of model AM2 (2-steps Acidogenesis-Methanogenesis model, see [3]) with soluble microbial products (SMP) dynamics. The production and the degradation of SMP play an important role in the membrane fouling phenomenon.

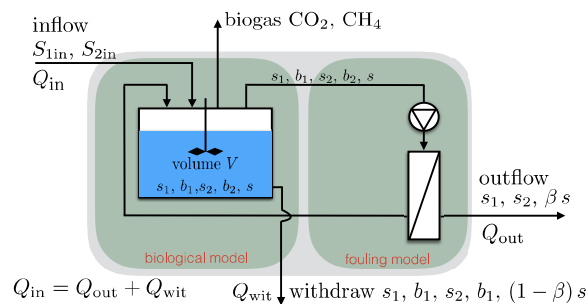
We present the original ODE model of the AM2b. Then we introduce a pure jump Markov model of the same device and a exact Monte Carlo simulation method of this process. Next we propose faster approximated simulation methods. Finally we present a stochastic differential equation (SDE) model of the AM2b. The validity of these models depends on the scale considered for the problem, this could be cleared with a law of large numbers/central limit theorem of the functional type presented in the last section.

1. The ODE model

The state variable of this model is:

$$x \stackrel{\text{def}}{=} \begin{pmatrix} s_1 \\ b_1 \\ s_2 \\ b_2 \\ s \end{pmatrix} \begin{array}{l} \text{concentration of organic matter} \\ \text{concentration of acidogenic biomass} \\ \text{concentration of volatile fatty acids (VFA)} \\ \text{concentration of methanogenic biomass} \\ \text{concentration of the soluble microbial products (SMP)} \end{array}$$

The AM2b model describes the dynamics of biological and anaerobic wastewater treatment



On the right of these scheme the membrane fouling model is represented; the separation of mater is as follow: the substrates s_1 and s_2 go through the membrane without retention (the size of their molecules is assumed to be smaller than pore diameter), the biomass

b_1 and b_2 is retained by the membrane, and a fraction βs of the SMP go through the membrane and leaves the reactor ($(1 - \beta) s$ will be considered as macromolecules).

In Appendix A.1 we describe the AM2b model as a reaction network of $J = 15$ reactions, and in Appendix A.2 we explain how this reaction network can be translated into an ODE system thanks to the laws of mass action and mass conservation. The ODE system reads:

$$\dot{s}_1 = D_{\text{in}} (S_{1\text{in}} - s_1) - k_1 \mu_1(s_1) b_1, \quad (1a)$$

$$\dot{b}_1 = (\mu_1(s_1) + \mu(s) - D_{\text{dec}} - D_{\text{wit}}) b_1, \quad (1b)$$

$$\dot{s}_2 = D_{\text{in}} (S_{2\text{in}} - s_2) - k_2 \mu_2(s_2) b_2 + (c_{12} \mu_1(s_1) + c_{02} \mu(s)) b_1, \quad (1c)$$

$$\dot{b}_2 = (\mu_2(s_2) - D_{\text{dec}} - D_{\text{wit}}) b_2, \quad (1d)$$

$$\dot{s} = (c_{10} \mu_1(s_1) + D_{\text{dec}} - k_0 \mu(s)) b_1 + (c_{20} \mu_2(s_2) + D_{\text{dec}}) b_2 - M s \quad (1e)$$

or

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \quad (2)$$

where:

k_i degradation rate of s_i by b_i ,	$M = \beta D_{\text{out}} + (1 - \beta) D_{\text{wit}}$
k_0 degradation rate of s by b_1 ,	β SMP fraction passing through the membrane,
c_{12} production rate of s_2 by b_1 from s_1 ,	D_{in} dilution rate ($= Q_{\text{in}}/V$),
c_{02} production rate of s_2 by b_1 from s ,	D_{dec} decay rate of biomass,
c_{i0} production rate of s by b_i from s_i ,	D_{wit} withdrawal rate of biomass ($= Q_{\text{wit}}/V$),
S_{in} input concentrations of s_i ,	D_{out} outflow rate of the bioreactor ($= Q_{\text{out}}/V$)

($i = 1, 2$). To ensure a constant volume we state that $Q_{\text{in}} = Q_{\text{out}} + Q_{\text{wit}}$. The rate M states that a proportion β of the SMP will leave the bioreactor through the membrane, at rate D_{in} and a proportion $1 - \beta$ through the withdrawal process, at rate D_{wit} .

The growth functions are:

$$\mu_1(s_1) = m_1 \frac{s_1}{K_1 + s_1}, \quad \mu(s) = m \frac{s}{K + s}, \quad \mu_2(s_2) = m_2 \frac{s_2}{K_2 + s_2 + s_2^2/K_1}. \quad (3)$$

μ_1 and μ are of Monod type; μ_2 is of Haldane type to model the phenomenon of inhibition of the eventual accumulation of the volatile fatty acids in the bioreactor during the methanogenesis (the major problem of the anaerobic digestion).

Model (1) relies on the fact that the stochastic effects can be neglected or at least can be averaged out. Although this level of description is sufficient for a number of applications of interest, it could be a valuable way of accounting for the stochastic nature of the system. Indeed, at small population sizes the AM2b model could present stochastic behaviors. 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. In these cases, stochastic features may not be neglected. We aim to build a model that still relies on a mass balance principle and that encompasses the useful stochastic information.

2. Pure jump Markov model

Following the approach described in [4], we propose a representation of the AM2b model as a pure jump Markov process:

$$X(t) = (S_1(t), B_1(t), S_2(t), B_2(t), S(t))^* \quad (4)$$

taking values in \mathbb{R}_+^5 . This process will encompass the J reactions: each reaction j is now characterized by its intensity functions $\lambda_j(x)$ and its jump functions $\nu_j(x)$, see details in Section A.3. The dynamic of the process $X(t)$ is described as follows: $X(0) = x_0$ and conditionally on $X(t) = x$, we set

$$X(t + \Delta t) = \begin{cases} x + \nu_j(x) & \text{with probability } \lambda_j(x) \Delta t + o(\Delta t) \quad 1 \leq j \leq J \\ x & \text{with probability } 1 - \sum_{j=1}^J \lambda_j(x) \Delta t + o(\Delta t) \end{cases} \quad (5)$$

where $(\lambda_j(x), \nu_j(x))_{1 \leq j \leq J}$ is given by (12) and (13).

2.1. Simulation and representation of $X(t)$

The process $X(t)$ can be simulated according to the following SSA (stochastic simulation algorithm) [6]:

$$\begin{aligned} & X \leftarrow x_0, t \leftarrow 0 \\ & \textbf{while } t < T_{\max} \textbf{ do} \\ & \quad \tau \leftarrow \sum_{j=1}^J \lambda_j(X) \\ & \quad S \sim \text{Exp}(\tau) \\ & \quad \text{sample } j \text{ according to the distribution } (\frac{\lambda_1(X)}{\tau}, \dots, \frac{\lambda_J(X)}{\tau}) \\ & \quad t \leftarrow t + S \\ & \quad X \leftarrow X + \nu_j(X) \\ & \textbf{end while} \end{aligned} \quad (6)$$

This Monte Carlo procedure allows us to simulate *exact* trajectories of the process X_t , the only approximation resides in the algorithms used for simulating the basic probability distributions.

Algorithm (6) is an exact representation of the process $X(t)$, and it leads to the following representation of the process:

$$X_t = X_0 + \sum_{j=1}^J \int_{[0,t] \times \mathbb{R}_+} 1_{[0, \lambda_j(X_{s-})]}(v) \nu_j(X_{s-}) \mathcal{N}_j(ds, dv) \quad (7)$$

where $\mathcal{N}_j(ds, dv)$ are independent Poisson random measures of intensity measure $ds \times dv$ (the Lebesgue measure on \mathbb{R}_+^2).

3. Discrete time approximations

The SSA simulates each reaction of the ecosystem asynchronously in time. In many situations this detailed simulation is too cumbersome, this is why synchronous discrete time approximations have been proposed. Let $t_m = m \Delta t$, for $\Delta t > 0$ fixed.

Poisson approximation

We construct an approximation $(\tilde{X}(t_m))_{m \geq 1}$. On the interval $[t_m, t_{m+1})$ suppose that the different rate functions are approximated by:

$$\lambda_j(x) \simeq \lambda_j(\tilde{X}(t_m)), \quad \forall x \in \mathbb{R}_+^5$$

so that each of the J reactions are independent and occur at constant rates $\lambda_j(\tilde{X}(t_m))$, that is the number of reactions of type j is a Poisson process of intensity $\lambda_j(\tilde{X}(t_m))$.

Hence, on the time interval $[t_m, t_{m+1})$ the number of reactions of type j follows a Poisson distribution of parameter $\Delta t \lambda_j(\tilde{X}(t_m))$. We obtain the following approximation also called τ -leaping:

$$\tilde{X}(t_{m+1}) = \left[\tilde{X}(t_m) + \sum_{j=1}^J \nu_j(\tilde{X}(t_m)) \rho_{j,m} \right]_+ \quad (8)$$

where $\rho_{j,m}$ are independent Poisson distribution variables with parameter $\Delta t \lambda_j(\tilde{X}(t_m))$ and $[x]_+$ is the projection of \mathbb{R}^5 onto \mathbb{R}_+^5 (the positive part of each component).

Diffusion approximation

The Poisson distribution with parameter $\Delta t \lambda_j(\tilde{X}(t_m))$, for $\Delta t \lambda_j(\tilde{X}(t_m))$ large enough, can be approximated by a Normal distribution of mean $\Delta t \lambda_j(\tilde{X}(t_m))$ and variance $\Delta t \lambda_j(\tilde{X}(t_m))$. From (8) we get:

$$\tilde{\xi}(t_{m+1}) = \left[\tilde{\xi}(t_m) + \sum_{j=1}^J \nu_j(\tilde{\xi}(t_m)) \left\{ \Delta t \lambda_j(\tilde{\xi}(t_m)) + \sqrt{\Delta t \lambda_j(\tilde{\xi}(t_m))} w_{j,m} \right\} \right]_+$$

where $w_{j,m}$ are independent $N(0, 1)$ random variables. This last equation can be rewritten:

$$\tilde{\xi}(t_{m+1}) = \left[\tilde{\xi}(t_m) + F(\tilde{\xi}(t_m)) \Delta t + \sum_{j=1}^J \frac{1}{\sqrt{N_j}} g_j(\tilde{\xi}(t_m)) [W_j(t_{m+1}) - W_j(t_m)] \right]_+ \quad (9)$$

where $W_j(t)$ are independent standard Brownian motions so that $W_j(t_{m+1}) - W_j(t_m)$ are independent and $N(0, \Delta t)$; $F(x)$ is defined in (15) and:

$$g_j(x) \stackrel{\text{def}}{=} \sqrt{N_j} \sqrt{\lambda_j(x)} \nu_j(x) = \sqrt{\tilde{\lambda}_j(x)} N_j \nu_j(x).$$

Let

$$\tilde{g}_j(x) \stackrel{\text{def}}{=} \sqrt{\tilde{\lambda}_j(x)} \tilde{\nu}_j$$

so that:

$$g_j(x) - \tilde{g}_j(x) = 1_{x \in \mathcal{D}} (g_j(x) - \tilde{g}_j(x))$$

and $|g_j(x) - \tilde{g}_j(x)| \leq C 1_{x \in \mathcal{D}} \sqrt{1 + |x|}$.

Stochastic differential equation

Equation (9), is an Euler-Maruyama discrete time approximation of the following stochastic differential equation (SDE):

$$d\xi(t) = F(\xi(t)) dt + \sum_{j=1}^J \frac{1}{\sqrt{N_j}} g_j(\xi(t)) dW_j(t), \quad \xi(0) = x_0. \quad (10)$$

4. Scales and asymptotics

According to Section A.3, the scale parameters N_j are connected to the “size” of the jumps in the reactions ②. We can assume that the F_j ’s range from 10^4 to 10^9 . When a reaction involves only substrate molecules the corresponding F_j ’s range from 10^7 to 10^9 ; when a reaction involves only bacteria the corresponding F_j ’s range from 10^4 to 10^6 . Hence for reasonable concentrations, the simulation algorithm (6) will not be feasible as it simulates every single reaction.

First suppose that $N = N_j$ for all j and that N is large. The first well know result can be understood as a functional law of large numbers (originally proved in this context by Tom Kurtz [9, 10]), it states that:

$$\sup_{0 \leq t \leq T} |X(t) - x(t)| \xrightarrow[N \rightarrow \infty]{} 0 \quad (11)$$

in L^2 or in probability. It is clear that in (11) we can replace $X(t)$ by $\xi(t)$. So under mild conditions, when the population sizes are large and so the number of reaction, the ODE model (2) is adapted to this scale.

At an intermediate scale, a functional central limit theorem states that the process $\sqrt{N}(X(t) - x(t))$ can be approximated in law by $\sum_{j=1}^J \int_0^t g_j(x(s)) d\tilde{W}_j(s)$ where the $\tilde{W}_j(s)$ are independent standard Brownian motions, that is formally:

$$X(t) \simeq x(t) + \frac{1}{\sqrt{N}} \sum_{j=1}^J \int_0^t g_j(x(s)) d\tilde{W}_j(s).$$

This also proves that the SDE model (10) is adapted to this scale.

In many situation ODE and SDE models are not valid. This is the case when one of the bacterial population is of “small” size but still affects the global dynamic of the process. This so-called “molecular randomness” may influence the global dynamic even when the population sizes are not so small [5]. In this case we may adopt hybride approaches. We just present an example where we separate the dynamics of the substrates from the dynamics of the biomasses: The idea is to break down the reactions between substrate type reactions and biomass type reactions, then to describe the first ones as a system of ODE’s and to describe the second ones as a pure jump Markov process. For example we can obtain a system of ODE’s describing the continuous evolution of the substrates and the SMP concentrations:

$$\begin{aligned} \dot{s}_1 &= D_{in}(S_{1in} - s_1) - k_1 \mu_1(s_1) B_1, \\ \dot{s}_2 &= D_{in}(S_{2in} - s_2) - k_2 \mu_2(s_2) B_2 + (c_{12} \mu_1(s_1) + c_{02} \mu(s)) B_1, \\ \dot{s} &= (c_{10} \mu_1(s_1) + D_{dec} - k_0 \mu(s)) B_1 + (c_{20} \mu_2(s_2) + D_{dec}) B_2 - M s \end{aligned}$$

coupled to a 2-dimensional pure jump process describing the discrete evolution of the biomasses concentrations:

jump	rate	jump	rate
$B_1 \rightarrow B_1 + \delta_1$	$\mu_1(s_1) B_1 / \delta_1$	$B_2 \rightarrow B_2 + \delta_2$	$\mu_2(s_2) B_1 / \delta_1$
$B_1 \rightarrow B_1 + \delta_1$	$\mu(s) B_1 / \delta_1$	$B_2 \rightarrow B_2 - \delta_2$	$D_{dec} B_1 / \delta_2$
$B_1 \rightarrow B_1 - \delta_1$	$D_{dec} B_1 / \delta_1$	$B_2 \rightarrow B_2 - \delta_2$	$D_{wit} B_1 / \delta_2$
$B_1 \rightarrow B_1 - \delta_1$	$D_{wit} B_1 / \delta_1$		

This type of model, known as *Piecewise-deterministic Markov process*, is very promising and will be investigated in the near future.

5. Conclusion

We show that an ODE model of microbial dynamics (1) contains all the ingredients that can be used to establish a pure jump Markov model, see Section 2. This pure jump Markov model can be exactly simulated with the Monte Carlo technique (6). This exact Monte Carlo method is not feasible in large population size cases and we proposed a Poissonian discrete time approximation (8) (also called τ -leaping) and a diffusion discrete time approximation (9). This last equation is the Euler-Maruyama time discretization of the SDE (10). This SDE is valid in high population size and different recalling so that alternative formulations of this SDE can be established. In Section 4, we describe the validity of these different models according to the scales at which the process should be simulated. All these models share the same ingredient, however they have very different qualitative properties.

Acknowledgments

This work was funded by the NuWat LIRIMA Inria project which funded this work.

6. References

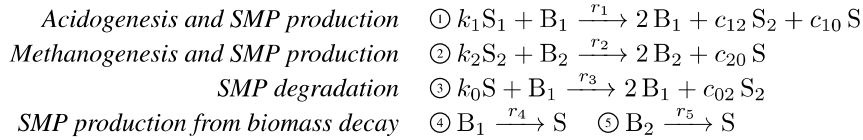
- [1] Boumediène Benyahia, Tewfik Sari, Brahim Cherki and Jérôme Harmand. Anaerobic membrane bioreactor modeling in the presence of Soluble Microbial Products (SMP) – the Anaerobic Model AM2b. *Chemical Engineering Journal*, 228(0):1011–1022, 2013.
- [2] Boumediène Benyahia. *Modélisation et observation des bioprocédés à membranes: application à la digestion anaérobie*. PhD thesis, Université de Tlemcen et Université de Montpellier 2, 2012.
- [3] Olivier Bernard, Zakaria Hadj-Sadock, Denis Dochain, Antoine Genovesi, and Jean-Philippe Steyer. Dynamical model development and parameter identification for an anaerobic wastewater treatment process. *Biotechnology and Bioengineering*, 75:424–438, 2001.
- [4] Fabien Campillo, Marc Joannides, and Irène Larramendy-Valverde. Stochastic modeling of the chemostat. *Ecological Modelling*, 222(15):2676–2689, 2011.
- [5] Fabien Campillo and Claude Lobry. Effect of population size in a Predator-Prey model. *Ecological Modelling*, 246:1–10, 2012.
- [6] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [7] Desmond J. Higham. Modeling and simulating chemical reactions. *SIAM Review*, 50(2):347–368, 2008.
- [8] Sebastiaan Kops, Krist Gernaey, Olivier Thas and Peter A. Vanrolleghem. The modelling of noise processes in stochastic differential equations: Application to biotechnological processes. In *Proceedings 7th IFAC Conference on Computer Applications in Biotechnology CAB7. Osaka, Japan, May 31 - June 4*, pages 67–72, 1998.
- [9] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

- [10] Thomas G. Kurtz. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8:344–356, 1971.
- [11] Mukhtar Ullah and Olaf Wolkenhauer. *Stochastic approaches in systems biology*. Springer, 2011.
- [12] Darren J. Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC, 2006.

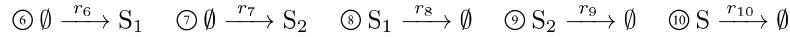
A. Appendix: the AMB2b models

A.1. AMB2b as a reaction network

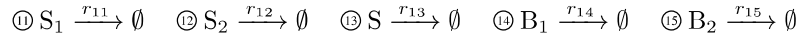
The first set of reactions describes the biochemical reactions:



The second set of reactions describes the *substrate inflow* and the *substrate outflow through the membrane*:



The third and last set of reactions describes the *biomass and substrate withdrawal* with :



The second and third set of reactions are not biochemical reactions they just describe the inflows and outflows in the AM2B process. In reaction $\textcircled{10}$ only a proportion β of the SMP goes through the membrane, and in reaction $\textcircled{13}$ a proportion $1 - \beta$ of the SMP is withdrawn, this mechanism will be explicated in the definition of the rates. Let j the index of the reaction and $J = 15$ the number of reactions. $\{(\textcircled{j}, r_j); j = 1, \dots, J\}$

A.2. AMB2b as an ODE system

The AMB2b reaction network described in Section A.1 is “translated” to an ODE system thanks to the laws of mass action and conservation of mass. The state variables of the ODE system are the concentration $s_i = [S_i]$, $b_i = [B_i]$ ($i = 1, 2$) and $s = [S]$. For example, for the $\textcircled{1}$, the *rate of reaction* also called *speed of reaction*, is defined by:

$$r_1 = \frac{b_1(t+\Delta) - b_1(t)}{\Delta} = -k_1 \frac{s_1(t+\Delta) - s_1(t)}{\Delta} = c_{12} \frac{s_2(t+\Delta) - s_2(t)}{\Delta} = c_{10} \frac{s(t+\Delta) - s(t)}{\Delta}$$

(the equalities are due to the mass conservation). This reaction, like reactions $\textcircled{2}$ and $\textcircled{3}$, is of order two and the mass action law states that $r_1 = s_1 b_1$, but saturation/inhibition phenomena suggest to replace in this last expression s_1 by $\mu_1(s_1)$ indeed:

Biochemical reactions: The mass action law applied to the second order reactions $\textcircled{1} \textcircled{2} \textcircled{3}$ states that:

$$r_1 = \mu_1(s_1) b_1, \quad r_2 = \mu_2(s_2) b_2, \quad r_3 = \mu(s) b_1$$

where the growth functions μ_i and μ are chosen as (3), indeed for low substrat concentration, these growth functions are linear accordingly to the mass action law, but for higher

substrate concentrations, saturation and inhibition phenomena have to be taken into account. The mass action law applied to first order reactions ④ ⑤ gives $r_4 = D_{\text{dec}} b_1$ and $r_5 = D_{\text{dec}} b_2$ where D_{dec} is decay rate of biomass.

Inflow and outflows: Inflow is done at rate D_{in} , outflow through the membrane at rate D_{out} and only a proportion β of the SMP is affected by the outflow, so the rates of these reactions are:

$$r_6 = D_{\text{in}} S_{1\text{in}}, \quad r_7 = D_{\text{in}} S_{2\text{in}}, \quad r_8 = D_{\text{out}} s_1, \quad r_9 = D_{\text{out}} s_2, \quad r_{10} = \beta D_{\text{out}} s.$$

Withdrawal: The withdrawal is done at rate D_{wit} and only a proportion $1 - \beta$ of the SMP affected by the withdrawal, so the rates of these reactions are:

$$r_{11} = D_{\text{wit}} s_1, \quad r_{12} = D_{\text{wit}} s_2, \quad r_{13} = (1 - \beta) D_{\text{wit}} s, \quad r_{14} = D_{\text{wit}} b_1, \quad r_{15} = D_{\text{wit}} b_2.$$

Summing up these expressions and applying the mass conservation law lead to the system of differential equations (1).

A.3. AMB2b as a pure jump Markov process

The AMB2b reaction network described in Section A.1 is “translated” into a pure jump Markov process thanks to the stochastic law of mass action [12]. Now $X(t) = [S_1(t), B_1(t), S_2(t), B_2(t), S(t)]^*$ is a pure jump Markov process defined by (5) where each reaction ⑦ is described as an instantaneous jump $X(t) \rightarrow X(t) + \nu_j(X(t))$ occurring with intensity $\lambda_j(X(t))$ defined respectively by:

$$\lambda_j(x) \stackrel{\text{def}}{=} N_j \tilde{\lambda}_j(x), \quad \nu_j(x) \stackrel{\text{def}}{=} [x + \frac{1}{N_j} \tilde{\nu}_j]_+ - x \quad (12)$$

($[x]_+$ the orthogonal projection of x onto \mathbb{R}_+^5) with

j	$\tilde{\lambda}_j(x)$	$\tilde{\nu}_j^*$
①	$\mu_1(s_1) b_1$	$[-k_1 \quad +1 \quad +c_{12} \quad 0 \quad +c_{10}]$
②	$\mu_2(s_2) b_2$	$[0 \quad 0 \quad -k_2 \quad +1 \quad +c_{20}]$
③	$\mu(s) b_1$	$[0 \quad +1 \quad +c_{02} \quad 0 \quad -k_0]$
④	$D_{\text{dec}} b_1$	$[0 \quad -1 \quad 0 \quad 0 \quad +1]$
⑤	$D_{\text{dec}} b_2$	$[0 \quad 0 \quad 0 \quad -1 \quad +1]$
⑥	$D_{\text{in}} S_{1\text{in}}$	$[+1 \quad 0 \quad 0 \quad 0 \quad 0]$
⑦	$D_{\text{in}} S_{2\text{in}}$	$[0 \quad 0 \quad +1 \quad 0 \quad 0]$
⑧	$D_{\text{out}} s_1$	$[-1 \quad 0 \quad 0 \quad 0 \quad 0]$
⑨	$D_{\text{out}} s_2$	$[0 \quad 0 \quad 0 \quad -1 \quad 0]$
⑩	$\beta D_{\text{out}} s$	$[0 \quad 0 \quad 0 \quad 0 \quad -1]$
⑪	$D_{\text{wit}} s_1$	$[-1 \quad 0 \quad 0 \quad 0 \quad 0]$
⑫	$D_{\text{wit}} b_1$	$[0 \quad -1 \quad 0 \quad 0 \quad 0]$
⑬	$D_{\text{wit}} s_2$	$[0 \quad 0 \quad 0 \quad -1 \quad 0]$
⑭	$D_{\text{wit}} b_2$	$[0 \quad 0 \quad 0 \quad -1 \quad 0]$
⑮	$(1 - \beta) D_{\text{wit}} s$	$[0 \quad 0 \quad 0 \quad 0 \quad -1]$

(13)

About the second equation of (12): Basically the jumps are $\frac{1}{N_j} \tilde{\nu}_j$, but near the border of \mathbb{R}_+^5 to avoid jumps that can lead to negative concentration values, we adopt truncated jumps so that $x + \nu_j(x) \in \mathbb{R}_+^5$ for all $x \in \mathbb{R}_+^5$. Indeed, note that $-\tilde{\nu}_{\min} \leq \tilde{\nu}_{ij} \leq \tilde{\nu}_{\max}$, so that if $x_i \geq \tilde{\nu}_{\min} / \min_j N_j$ then $\nu_{ij}(x) = \tilde{\nu}_{ij} / N_j$, where $\tilde{\nu}_{ij}$ and $\nu_{ij}(x)$ denote the i th component of $\tilde{\nu}_j$ and $\nu_j(x)$ respectively. Define:

$$\mathcal{D} \stackrel{\text{def}}{=} \{x \in \mathbb{R}_+^5; x_i \geq \tilde{\nu}_{\min} / \min_j N_j\} \quad (14)$$

so that $x \in \mathcal{D}$ implies that $\nu_j(x) = \frac{1}{N_j} \tilde{\nu}_j$.

The drift coefficient: Given $X(t) = x$, the expectation of $X(t + \Delta t)$ is

$$\begin{aligned} \mathbb{E}[X(t + \Delta t) | X_t = x] &= \sum_{j=1}^J (x + \nu_j) \mathbb{P}[\text{reaction } j | X_t = x] + x \mathbb{P}[\text{no reaction} | X_t = x] \\ &\simeq \sum_{j=1}^J (x + \nu_j) (\lambda_j(x) \Delta t) + x (1 - \sum_{j=1}^J \lambda_j(x) \Delta t) \\ &\simeq x + \sum_{j=1}^J \lambda_j(x) \nu_j(x) \Delta t = x + F(x) \Delta t \end{aligned}$$

where

$$F(x) \stackrel{\text{def}}{=} \sum_{j=1}^J \lambda_j(x) \nu_j(x) \quad (15)$$

So locally in time, $\mathbb{E}(X(t))$ evolves according to the drift coefficient $F(x)$ (note that $\mathbb{E}(X(t))$ is not solution of an ODE as the function F is non linear). We can easily check that:

$$F(x) = 1_{x \in \mathcal{D}} \tilde{F}(x) + 1_{x \notin \mathcal{D}}(x) F(x)$$

where

$$F(x) \stackrel{\text{def}}{=} \sum_{j=1}^J \tilde{\lambda}_j(x) \tilde{\nu}_j \quad (16)$$

does not depend on the N_j 's. Finally from $|F(x)| + |\tilde{F}(x)| \leq C(1 + |x|)$ and $F(x) - \tilde{F}(x) = 1_{x \notin \mathcal{D}}(F(x) - \tilde{F}(x))$, we get:

$$|F(x) - \tilde{F}(x)| \leq C 1_{x \notin \mathcal{D}}(1 + |x|).$$