Manuel L. Esquível*, Paula Patrício, and Gracinda R. Guerreiro

From ODE to Open Markov Chains, via SDE: an application to models for infections in individuals and populations

Abstract: We present a methodology to connect an ordinary differential equation (ODE) model of interacting entities at the individual level, to an open Markov chain (OMC) model of a population of such individuals, via a stochastic differential equation (SDE) intermediate model. The ODE model here presented is formulated as a dynamic change between two regimes; one regime is of mean reverting type and the other is of inverse logistic type. For the general purpose of defining an OMC model for a population of individuals, we associate an Ito processes, in the form of SDE to ODE system of equations, by means of the addition of Gaussian noise terms which may be thought to model non essential characteristics of the phenomena with small and undifferentiated influences. The next step consists on discretizing the SDE and using the discretized trajectories computed by simulation to define transitions of a finite valued Markov chain; for that, the state space of the Ito processes is partitioned according to some rule. For the example proposed for illustration, the state space of the ODE system referred – corresponding to a model of a viral infection – is partitioned into six infection classes determined by some of the critical points of the ODE system; we detail the evolution of some infected population in these infection classes.

Keywords: Infection Modeling, Population Dynamics, Ordinary Differential Equations, Stochastic Differential Equations, Markov Chains

MSC: 92D30, 92B99, 60J20, 60J70

1 Introduction

In this work we have two generic goals. The first goal is to show how to associate an initial ODE model applied to a generic individual of a population to a Markov chain model for the evolution of an open population of these individuals. The second goal is to illustrate the details of the procedure by considering the case of a viral infection on a human population.We thus show how to build a finite state Markov chain model starting from an ordinary system of non linear differential equations. The method relies on an intermediary step that associates a system of stochastic differential equations to the ODE system.

In [15] there is a reference to the possibility of associating a Markov chain to a stochastic differential equation with the ultimate purpose of studying a model, by simulation, in a similar manner as the well known Markov Chain Monte Carlo method. The study of the properties of a stochastic differential equation of diffusion type that transfer to an associated Markov chain, by means of a discretization scheme – such as the Euler-Maruyama – has been made almost two decades ago, see [33]. Other important references that

*Corresponding Author: Manuel L. Esquível: Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Campus de Caparica, 2829-516, Caparica, Portugal & Centro de Matemática e Aplicações
Paula Patrício: Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Campus de Caparica, 2829-516, Caparica, Portugal & Centro de Matemática e Aplicações
Gracinda R. Guerreiro: Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Campus de Caparica, 2829-516, Caparica, Portugal & Centro de Matemática e Aplicações

study properties of an associated Markov chain, created from a diffusion by means of a discretization scheme, are [31], [29] and [19].

There are several examples in the literature of the type of models studied in this work, namely, identifying particular regimes. For instance, in [4] the authors focused on the *mean reverting* type of model for the SDE of interest. In [35], using *Kalman filter* approach for the parameters estimation the authors consider a diffusion with constant volatility and linear drift.

In [34] coupled systems of two and three SDE with constant volatility are considered. In [9] a modified Gompertzian SDE equation, with linear volatility function and an explicit solution similar to the Ornstein-Ulhenbeck process, is used to describe bacterial growth under bactericide. In [25] the approach is also from an ODE model to a SDE system model in order to account for the fact that (and we quote) *process is not smooth, subject as it is to a variety of metabolic and hormonal influences, which change over time*. The system has only one equation, with volatility term proportional to the product of the two processes and is coupled as the two processes appear in the equation with null volatility.

The most used estimation methods for SDE do not apply to the model here proposed, see [5] for a comprehensive review on the subject. We anticipate several difficulties developing such methods and we hope to address this issue in future works. For the moment we parametrize the SDE model to obtain illustrative examples for which the methodology here proposed can be applied.

The importance of linking within-host dynamics and between-host dynamics has been long recognized, see [14]. Heterogeneous population dynamics results from individual differences which produce the complex epidemiological patterns observed. For virus infections, in particular viral infection, authors have considered e.g. the connection between the two scales through the epidemiological reproduction number [3] and, using two ODE nested models to study the evolution of virulence [10].

Let us comment on some aspects of the example of viral infection we chose to illustrate our method. It is well known that there is an obvious relation between the number of viruses attacking leucocytes. As the viral load increases the leucocyte count is expected to decrease; moreover we should expect a regular variation in this coupled evolution. Being so, an ODE model for an average behavior seems appropriate. But, as there are many other factors conditioning the evolution of the mutual influence process, it is also expected that, for each individual data set variations on the average evolution pattern may appear. In [26] there is a reference to influences, in a particular process interfering with the interaction of viruses and leucocytes, that are classified as being stochastic in nature. Being so, a SDE version of the initial ODE model for the pure viruses–leucocyte interaction may be justified. Some other work on the same lines of thought is the analysis of a three dimensional model of virus competition that is presented in [32]; a five dimensional analysis, contemplating both infected and non infected leucocyte cells is given in [6]. Let us briefly describe the contents of this work.

- An overview of the methodology proposed in this work is presented in Section 2.
- In Section 3, the very important idea of competing regimes is presented and next we model the coupled evolution of two competing entities in an individual – for instance, a viral load and a leucocyte count – by a system of two ordinary differential equations. This ODE model is supposed to describe the average behavior of the quantities modeled.
- Secondly, in Section 4, we consider a stochastic differential equation model (SDE) built upon the ODE model by adding a noise term to each of the ODE. Each trajectory of this SDE models the evolution of the two competing and interacting entities for a particular individual. As usual, the interpretation of a SDE model built by adding a noise term to the equations of a ODE model is that we consider that the evolution of the modeled quantities for an individual – which in general is described by the ODE model - is perturbed, at the individual level, by unknown factors.
- The third step – developed in Section 5 – consists on associating a Markov chain transition matrix to the SDE model, by simulating trajectories of the solution process of the the SDE and counting the transitions between *Infection Classes*. These *Infection Classes* were defined by regions in the phase space of the ODE by means of critical points of the solutions of the ODE system.
- In the last step – presented in Section 6 – we consider an open Markov chain (OMC) model for the population with five transient states (the first five stages of infection) and an absorbing state corresponding

to the last stage of infection, from where we assume that it is not possible to move out without treatment or death. The OMC model is adequate since the population is open and constantly fed by the number of newly viral infected. For general conditions on the number of newly viral infected patients, we are able to determine the evolution of the expected numbers of patients of the subpopulations – the Poisson parameters – in each one of the *Infection Classes* – as done in [7]. We exploit the open Markov chain model by considering simulated data from a yearly viral newly infected population on a 23 year period.

Remark 1. As so, this sequence of models, each one built upon the preceding one – ODE, SDE and OMC models - allow to transfer information, in a coherent way, from an individual level to a whole population level.

2 From ODE to Markov chains via SDE: an overview

The general pathway from an ordinary differential equations model to open population Markov chain model via a stochastic differential equation model is schematically described in the following. Let us suppose that we have a model given by the solution of a Cauchy-Lipschitz ordinary differential equation (ODE) – satisfying some regularity assumptions – and an initial condition such as:

$$\boldsymbol{y}'(t) = f(t, \boldsymbol{y}(t)) \ , \ \boldsymbol{y}(t_0) = \boldsymbol{y}_0 \ ,$$

or, integrated form,

$$\boldsymbol{y}(t) = \boldsymbol{y}_0 + \int_{t_0}^{t} f(s, \boldsymbol{y}(s)) ds \ . \tag{2.1}$$

Now, suppose that we consider equation (2.1) as giving the evolution of the mean value of a population size and that, aiming at having an individual random model, we add a noise, given by an Ito integral with respect to a Brownian process $(\boldsymbol{B}_t)_{t \in [0,T]}$ – defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ – such that we get a stochastic differential equation (SDE) given, for $\omega \in \Omega$, by:

$$\widetilde{\boldsymbol{y}}(t, \omega) = \widetilde{\boldsymbol{y}}_0(\omega) + \int_{t_0}^{t} f(s, \widetilde{\boldsymbol{y}}(s, \omega)) ds + \int_{t_0}^{t} \sigma(t, \widetilde{\boldsymbol{y}}(t, \omega)) d\boldsymbol{B}_t(\omega) \ , \tag{2.2}$$

or in the usual differential form,

$$d\widetilde{\boldsymbol{y}}(t) = f(t, \widetilde{\boldsymbol{y}}(t)) dt + \sigma(t, \widetilde{\boldsymbol{y}}(t)) d\boldsymbol{B}_t \ , \ \widetilde{\boldsymbol{y}}(t_0) = \widetilde{\boldsymbol{y}}_0 \ .$$

Let us observe that under sufficient hypothesis guaranteeing that the Ito stochastic integral part of equation (2.2) is a martingale, with constant null mean, we will have,

$$\mathbb{E}\left[\widetilde{\boldsymbol{y}}(t, \omega)\right] = \mathbb{E}\left[\widetilde{\boldsymbol{y}}_0(\omega)\right] + \int_{t_0}^{t} \mathbb{E}\left[f(s, \widetilde{\boldsymbol{y}}(s, \omega))\right] ds \ ,$$

which, upon comparing with equation (2.1), allows the interpretation that the solution of equation (2.1) is the mean value with respect to the probability $\mathbb{P}$ – or the average – of the solution of equation (2.2).

From this point on we will not distinguish $\widetilde{\boldsymbol{y}}$ from $\boldsymbol{y}$. Now, for illustration purposes only [1], suppose that we discretize equation (2.2) with the Euler-Maruyama discretization scheme to get:

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + f(t_n, \boldsymbol{y}_n)(t_{n+1} - t_n) + \sigma(t_n, \boldsymbol{y}_n)(\boldsymbol{B}_{t_{n+1}} - \boldsymbol{B}_{t_n}) \ ,$$

---

1 Preservation of ergodic properties may require a higher order discretization scheme, but the main argument is preserved also in this case.

for $t_0 < t_1 < \cdots < t_N = T$ a subdivision of $[0, T]$. In the unidimensional case it is known that,

$$\boldsymbol{B}_{t_{n+1}} - \boldsymbol{B}_{t_n} \frown \mathcal{N}(0, (t_{n+1} - t_n)) \ ,$$

and we have that the sequence $(U_n)_{n=0,\ldots,N-1}$, defined by:

$$U_n := F_n^{\leftarrow}(\boldsymbol{B}_{t_{n+1}} - \boldsymbol{B}_{t_n}) \ ,$$

with $F_n^{\leftarrow}$ the inverse of the distribution function of a random variable distributed with the law $\mathcal{N}(0, (t_{n+1} - t_n))$, is a sequence of independent uniformly distributed random variables. In this way, if $t_{n+1} - t_n = h$ and $f(t, x) = f(x)$, $\sigma(t, x) = \sigma(x)$, that is, the diffusion given by (2.2) is homogeneous, with,

$$G(x, y) := x + f(x)h + \sigma(x)y \ ,$$

we may write that,

$$\boldsymbol{y}_{n+1} = G(\boldsymbol{y}_n, U_n) \ , \ \ n = 0, \ldots, N-1 \ ,$$

which is the functional definition of a Markov chain (see [28, p. 62]). As so, we could try to apply the methodologies used in [7] and [8] – and other references therein – with the purpose of identifying the possibility of stable behavior in the set of transient classes. Let us observe that, in principle, we will have that the chain $(\boldsymbol{y}_n)_{n=0,\ldots,N}$, for large N, may have a large number of states and so, in order to apply the referred methods of open Markov chain models, a natural situation would be to define a small number of classes C_i by considering a partition of the state space, S_i, with $S = \cup_{i=1}^d S_i$ and $S_i \cap S_j$ for $i \neq j$, such that at step n,

$$\omega \in C_i \Leftrightarrow \boldsymbol{y}_n(\omega) \in S_i \ .$$

In the context of epidemiological studies this is feasible if the state space is some set of values of some biological indicators or markers like a virus load and some cell count.

3 A non linear ODE 2–dimensional model

The competing regimes modeling we adopt in this work is rooted in *homeostasis* which is the maintenance of an internal stable environment for proper functioning of the body. As a consequence, we can suppose that body markers are subjected to variations that tend to conform to some alternating regimes, the alternation being driven by other processes in the body. For simplicity we suppose that, for the body marker we are going to model, the values can evolve in two different regimes. For illustration purposes we consider the usual mean reverting model given by the ODE,

$$w'(t) = 1 - w(t) \ , \ w(0) = w_0 \ , \tag{3.1}$$

which supposes that the marker has a tendency to evolve towards the value 1; and the inverted logistics model given by,

$$y'(t) = \frac{1}{(y(t) - 1)y(t)} \ , \ y(0) = y_0 \ , \tag{3.2}$$

which essentially says that the marker will evolve towards some larger value. We present some solutions of these ODE, corresponding to different initial values in Figure 3.1.

If in the *homeostasis* context one kind of evolution is changed in a different kind of evolution by the effect of external causes, we may consider that these two regimes are coupled by a regular function that shifts, in time, the evolution from one regime to another; the coupling will be provided by a convex combination of the two right-hand members of equations (3.1) and (3.2). For technical reasons we suppose that the coupling function, depends on two parameters α and λ [2] and is given by:

$$F(t, \alpha, \lambda) = \left(1 - e^{-(1-\alpha)\lambda t}\right)^{\frac{1}{1-\alpha}} \ . \tag{3.3}$$

---

2 F is given as a solution of the ODE $\Phi'(t) = \lambda\Phi(t)^\alpha(1 - \Phi(t)^{(1-\alpha)})$ with $\Phi(0) = 0$. This is important both for the qualitative study – see section 3.1 – and for a second stage of the modeling.
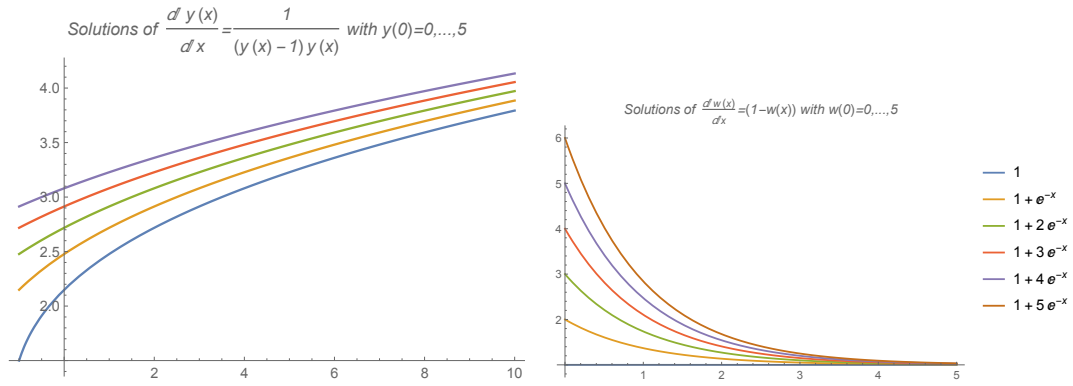
Figure 3.1: Solutions of deterministic *inverse logistics* and *mean reverting* models

The ODE with the two regimes, coupled by F, is given by (3.4) with additional parameters μ controlling the intensity of the *reversed logistics* drive, u_{Min}, which is a bound on the values of u and v, which is the mean reverting level.

$$u'(t) = \frac{1 - F(t, \alpha, \lambda)}{\mu \left(\frac{u(t)}{u_{\text{Min}}} - 1 \right) u(t)} + F(t, \alpha, \lambda)(v - u(t)) \ , \ \ u(t_0) = u_0 \ . \tag{3.4}$$

In Figure 3.2 we show, both F, for $\lambda = 1$ and for several values of the parameter α and a solution $u(t)$, $t \in [0, 30]$ – for $\alpha = 0.82$ in the coupling function f – of the ODE with the coupled regimes with some additional parameters given in Table 1.
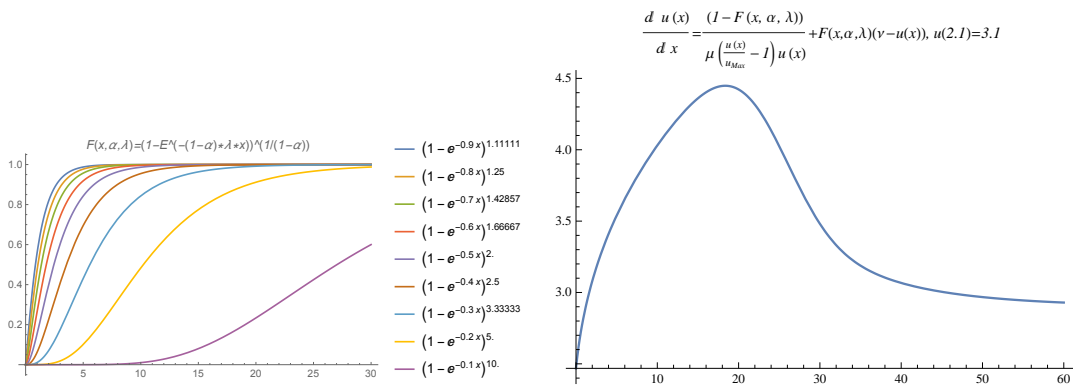

Figure 3.2: Coupling function and solution of coupled *inverse logistics* and *mean reverting* models

Remark 2. For modeling purposes we may use this particular kind of evolution, the one described by $u(t)$, $t \in [0, 30]$ in the right-hand side of Figure 3.2, for some biological internal marker of a normalized count that evolves with a continuous transition between a decreasing regime and an increasing one, according to the variation of the function f. We observe that we could also have a continuous transition between an increasing regime and a decreasing one by changing the *weights* f and $1 - f$, in the right-hand side of the ODE.

Remark 3. We stress that further combinations of regimes are possible: either by considering other coupling functions performing a convex combination of more than two terms or by considering a second convex combination – with another coupling function – of the existing convex combination. This could be done to alter

the asymptotic regime. Also, we observe that the parameter α defines, in some sort, the speed of the regime change.

There is a natural perspective to study epidemics namely, the population dynamics angle (see, for instance, [18] and references therein). To study the infection at the individual level and on the average, the approach relies on the modeling of the dynamics – by means of ODE – of some body internal markers, such as viral charges or leucocyte counts (see [1]) or leucocyte percentages or ratios.

That is the general approach we pursue in the following considering the first part of the proposed methodology: a non-linear ODE model for infections in which we may have two interacting entities – for instance some *viral load* and for instance the *leucocyte count*. This model, which aims at describing an average general behavior, has two coupled ODE, one for u – the *leucocyte count* – and another for z – *viral load*. In each equation there are two evolution regimes and the transition in time, between these regimes, is achieved by a coupling function. We stress that the model chosen is not of mechanistic type in the sense that it does not explicitly model the interaction of the two populations as in [23] and [27].

$$\begin{cases} u'(t) = \dfrac{1 - F(t, \alpha_1, \lambda_1)}{\mu_1 \left(\frac{z(t)}{u_{\text{Min}}} - 1 \right) z(t) + \epsilon_1} + F(t, \alpha_1, \lambda_1)(v_1 - u(t)) \,, \ u(t_1) = u_0 \\[4mm] z'(t) = \dfrac{F(t, \alpha_2, \lambda_2)}{\mu_2 \left(\frac{z(t)}{z_{\text{Min}}} - 1 \right) z(t) + \epsilon_2} + \left(1 - F(t, \alpha_2, \lambda_2) \right) (v_2 - z(t)) \,, \ z(t_2) = z_0 \,. \end{cases} \tag{3.5}$$

Let us detail some general features of this model:

1. Both the equations have two regimes, of the same type in the two equations – *inverse logistics* and *mean reverting* – coupled by the coupling function $F(t, \alpha, \lambda)$; this function of logistic type goes from zero – at time zero – to values very close to one in finite time.
2. The initial regime for u, controlled by $1 - F(t, \alpha_1, \lambda_1)$, for small values of t, is the *inverse logistics* term for the function z, which is also the final regime, controlled by the values of $F(t, \alpha_2, \lambda_2)$ for the function z. Being so, in this model, the *viral load* exerts some influence in the initial behavior of *leucocyte count*.

3.1 The qualitative study of the ODE model

In this section we state some questions related to the qualitative study of the set of ODE given in (3.5). We observe first that this system can be seen as an autonomous system by considering:

$$\begin{cases} u'(t) = \dfrac{1 - \Phi_{\alpha_1, \lambda_1}(t)}{\mu_1 \left(\frac{z(t)}{u_{\text{Min}}} - 1 \right) z(t) + \epsilon_1} + \Phi_{\alpha_1, \lambda_1}(t)(v_1 - u(t)) \,, \ u(t_1) = u_0 \\[4mm] \Phi'_{\alpha_1, \lambda_1}(t)(t) = \lambda_1 \Phi_{\alpha_1, \lambda_1}(t)^{\alpha_1} \left(1 - \Phi_{\alpha_1, \lambda_1}(t)^{(1-\alpha_1)} \right) \,, \ \Phi_{\alpha_1, \lambda_1}^{(1-\alpha_1)}(0) = 0 \\[4mm] z'(t) = \dfrac{\Phi_{\alpha_2, \lambda_2}(t)}{\mu_2 \left(\frac{z(t)}{z_{\text{Min}}} - 1 \right) z(t) + \epsilon_2} + \left(1 - \Phi_{\alpha_2, \lambda_2}(t) \right) (v_2 - z(t)) \,, \ z(t_2) = z_0 \\[4mm] \Phi'_{\alpha_2, \lambda_2}(t)(t) = \lambda_2 \Phi_{\alpha_2, \lambda_2}(t)^{\alpha_2} \left(1 - \Phi_{\alpha_2, \lambda_2}(t)^{(1-\alpha_2)} \right) \,, \ \Phi_{\alpha_2, \lambda_2}^{(1-\alpha_2)}(0) = 0 \,. \end{cases} \tag{3.6}$$

In fact, let us consider, for $\alpha \in]0, 1[$, the differential equation:

$$u'(t) = \lambda u(t)^{\alpha} \left(1 - u(t)^{(1-\alpha)} \right) \,, \ u(0) = u_0 \,. \tag{3.7}$$

With the change of variable $v = u^{(1-\alpha)}$ which implies $dv = (1 - \alpha)u^{-\alpha} du$, we have that,

$$\frac{du}{u^{\alpha}(1 - u^{(1-\alpha)})} = \frac{dv}{(1 - \alpha)(1 - v)} \,,$$

and so the equation in (3.7) may be written as

$$\frac{dv}{(1 - v)} = \lambda(1 - \alpha)dt \,,$$

with the immediate solution $v = 1 - \gamma e^{-\lambda(1-\alpha)t}$, with $\gamma > 0$ some integration constant, which gives,

$$u(t) = \left(1 - \gamma e^{-\lambda(1-\alpha)t}\right)^{\frac{1}{1-\alpha}}, \quad \gamma = 1 - u_0^{(1-\alpha)},$$

as the general solution of equation (3.7).

The qualitative study of a system such as the one in (3.6) – a system of four non linear ODE - may require some special delicate analysis. An alternative path could be to study directly the system (3.5) by means of some ad-hoc procedures. Examples of such studies are [11], [17] and [2].

Remark 4. The linearized part of the ODE system (3.6) is given by:

$$\begin{pmatrix} -\frac{(1-r)\left(\frac{z\mu_1}{a} + \left(\frac{z}{a}-1\right)\mu_1\right)}{\left(\epsilon_1 + z\left(\frac{z}{a}-1\right)\mu_1\right)^2} & -r & -u + v_1 - \frac{1}{\epsilon_1 + z\left(\frac{z}{a}-1\right)\mu_1} & 0 \\ 0 & 0 & r^{\alpha_1-1}\left(1 - r^{1-\alpha_1}\right)\alpha_1\lambda_1 - (1-\alpha_1)\lambda_1 & 0 \\ -\frac{s\left(\frac{z\mu_2}{b} + \left(\frac{z}{b}-1\right)\mu_2\right)}{\left(\epsilon_2 + z\left(\frac{z}{b}-1\right)\mu_2\right)^2} & s-1 & 0 & u - v_2 + \frac{1}{\epsilon_2 + z\left(\frac{z}{b}-1\right)\mu_2} \\ 0 & 0 & 0 & s^{\alpha_2-1}\left(1 - s^{1-\alpha_2}\right)\alpha_2\lambda_2 - (1-\alpha_2)\lambda_2 \end{pmatrix}$$

Remark 5. The linearized part of the ODE system (3.6) when $u = z = r = s = 0$ is given by:

$$\begin{pmatrix} \frac{\mu_1}{\epsilon_1^2} & 0 & v_1 - \frac{1}{\epsilon_1} & 0 \\ 0 & 0 & (\alpha_1 - 1)\lambda_1 & 0 \\ 0 & -1 & 0 & \frac{1}{\epsilon_2} - v_2 \\ 0 & 0 & 0 & (\alpha_2 - 1)\lambda_2 \end{pmatrix}$$

Remark 6. The eigenvalues of the linearized part at $(\mathbf{0}, \mathbf{0})$ of the system given in (3.6), given as function of the parameters, are the following.

$$\left\{ -\sqrt{-(\alpha_1 - 1)\lambda_1}, \sqrt{-(\alpha_1 - 1)\lambda_1}, (\alpha_2 - 1)\lambda_2, \frac{\mu_1}{\epsilon_1^2} \right\}$$

The first and the third ones are negative and the second and the fourth ones are positive. As so, by the theorem of Hartman-Grobman, see [24, pp. 120–127], the trajectories of the linearized system are homeomorphic images of the original system in a neighborhood of $\mathbf{0}$ and so, the point $(\mathbf{0}, \mathbf{0})$ is not a stable point of system (3.6). This property may have relevance for the question treated in subsection 10.

3.2 An illustrative numerical application

The system of equations (3.5) was solved numerically using the parameter values given in Table 1. These parameters were chosen in order to provide a solution with the desired characteristics.

Remark 7. For practical application of this model, the estimation of the parameters for equation system (3.5) should be done using the corresponding set of Ito processes in unidimensional case, in order the methodology to be applied with a set of significant data. Efficient estimating methods are being developed. Also, the qualitative study of the system of equations (3.5) should provide an insight into the desired properties of the model and the corresponding range of the parameters.

Table 1: The parameters values for the solution's computation

t_1	$u(t_1)$	α_1	T	μ_1	v_1	u_{Min}	λ_1	ϵ_1
0	4.1	0.94	70	17.1	1.9	2	1.8	0.1
t_2	$z(t_2)$	α_2	T	μ_2	v_2	z_{Min}	λ_2	ϵ_2
0	3	0.78	30	4.1	2.3	1.9	2.3	0.2

In Figure 3.3 there is a representation of the solution (u, z) of the system of equations (3.5) with parameters given in Table 1. This solution was obtained by numerical integration [3].
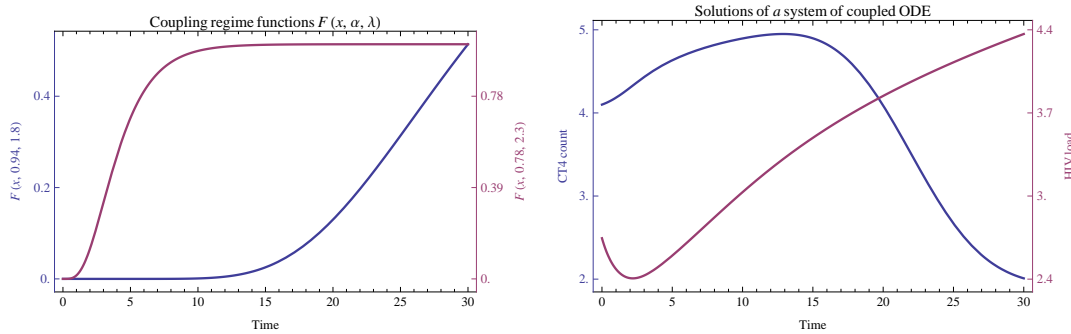

Figure 3.3: Solution of the coupled equations model for *viral load* and *leucocyte count*

Preliminary conclusions 3.1. Let us state some first conclusions about the presented solution of the model.

1. The general pattern of significant periods in time of contra-variation, of the two functions of time of *viral load* and of *leucocyte count*, is manifest in Figure 3.3.
2. The present report may be thought as a *proof of concept* exposition so, no homogeneity analysis for the units choice has been performed. We must stress that, in principle, the values of the *viral load* must be thought to be given in log scale. At this point of the exposition the time scale presented was not adjusted to any real data; see Remark 19 for an adjusted time scale proposal.
3. The pattern of the initial decreasing of the *viral load* that appears in Figure 3.3 was first modeled in the works of M. Novak and R. May. See, for instance, [22] where several models, with complex dynamics explaining this pattern, are studied.

We must emphasize that the model here studied does **not** take into consideration the existence of some kind of disease treatment; for such a purpose, within the approach developed in this work, we would add a third regime describing the average response of both variables to such a treatment. One example, within a different context, of such model is presented in [16].

4 The associated SDE 2–dimensional model

We now develop the second part of the proposed methodology: to model the two entities interaction – for instance, *viral load* and *leucocytes count* – at the individual level, by adding a Wiener process type noise to

---

[3] For the detailed computations see file *OnTheSecondIdea-E.nb* available upon request.

the ODE model introduced in Section 3. The SDE system we consider is given by:

$$\begin{cases} dX_t = F_1(t, X_t, Y_t)dt + \sigma_1 \sqrt{X_t} dB_t^1 \ , \ X_{t_0} = X_0 \ , \\ dY_t = F_2(t, X_t, Y_t)dt + \sigma_2 \sqrt{Y_t} dB_t^2 \ , \ Y_{t_0} = Y_0 \ , \end{cases} \tag{4.1}$$

where $(B_t^1)_{t \geq 0}$ and $(B_t^2)_{t \geq 0}$ are two independent Wiener processes and where, for notational simplicity for the coupled system of SDE, we have defined the drift functions taken from the right-hand side of the ODE system in formulas (3.5), by:

$$\begin{cases} F_1(t, u(t), z(t)) = \dfrac{1 - F(t, \alpha_1, \lambda_1)}{\mu_1 \left(\frac{z(t)}{u_{\text{Min}}} - 1 \right) z(t) + \epsilon_1} + F(t, \alpha_1, \lambda_1)(v_1 - u(t)) \ , \\ F_2(t, u(t), z(t)) = \dfrac{F(t, \alpha_2, \lambda_2)}{\mu_2 \left(\frac{z(t)}{z_{\text{Min}}} - 1 \right) z(t) + \epsilon_2} + \left(1 - F(t, \alpha_2, \lambda_2) \right) (v_2 - z(t)) \ . \end{cases} \tag{4.2}$$

Remark 8. We note that the additive noises used – having an intensity proportional to the square root of the SDE equations unknowns – X_t (respectively Y_t), corresponding to u (respectively v) in the ODE model – were chosen in order to ensure, if possible, that the solutions of the SDE system are always non negative.

Remark 9. Although it would be possible to consider some non identity fixed covariance matrix between $(B_t^1)_{t \geq 0}$ and $(B_t^2)_{t \geq 0}$, in formulas (4.1), we have assumed independence by lack of an interpretation of this possible covariance structure. This covariance structure may be added, in a real data model, upon the development of the parameter estimation procedure.

The system of SDE given by formulas (4.1) was discretized in order to obtain a simulated sample of trajectories. The parameter values used for the simulation are shown in Table 2. The trajectories of both processes defined by the SDE system (4.1), $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$, which were obtained by the simulation procedure, can be seen in the left-hand side of Figure 4.1.

Table 2: The parameters values for the simulation of sample trajectories

t_0	ϵ_1	α_1	T	μ_1	v_1	u_{Min}	λ_1	X_{t_0}	σ_1
0	0.1	0.94	30	17.1	1.9	2	1.8	4.1	0.3

t_0	ϵ_2	α_2	T	μ_2	v_2	z_{Min}	λ_2	Y_{t_0}	σ_2
0	0.2	0.78	30	4.1	2.3	1.9	2.3	3	0.25

Remark 10. A far easier approach to the conjectured result on positivity of the kind of diffusions used in this work is to show that the discretized sequence will only take positive values. This amounts to use properties of Gaussian random variables and can be done by induction. The stability analysis of the ODE system (3.6), in Subsection 3.1, has consequences for the proof of a technical result ensuring that the solutions of the SDE system (4.1) will remain positive almost surely. The general idea for such a result is that as the processes approach zero, with positive values, the *volatility* part of the equation becomes less and less important and so the *drift* part of the equations – that is the part corresponding to the ODE system given by (3.5)– dominates; as the point $(\mathbf{0}, \mathbf{0})$ is not stable (see remark 6), by force of two positive eigenvalues of the linearized part of the SDE system, the solution should be forced to stay on the positive values. The papers [21] and[20] contain recent work that may be a starting point to the necessary technical proof on the positivity of trajectories of the SDE of the model here presented.
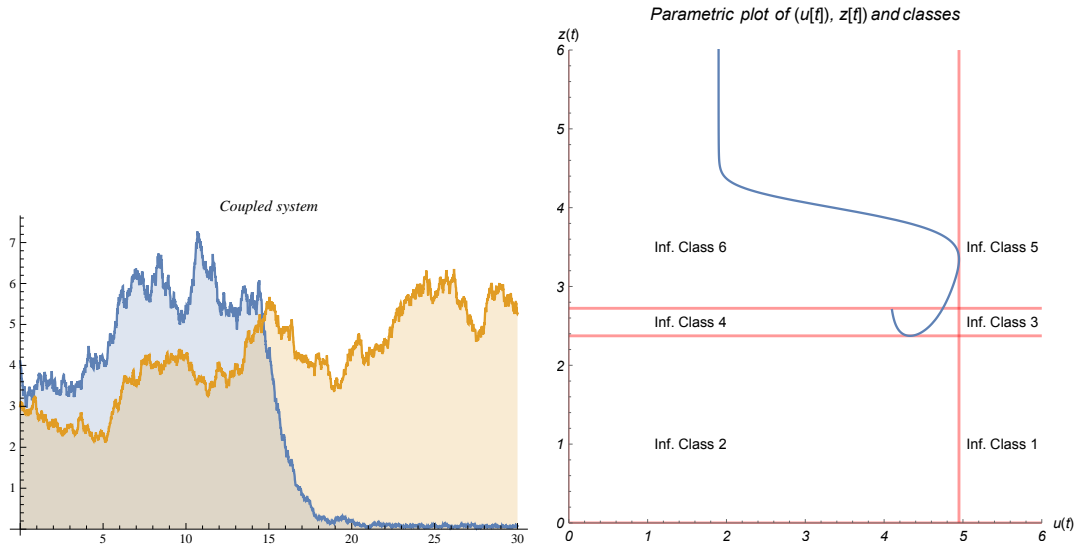
Figure 4.1: Simulated trajectories for X_t, the *viral load* (orange), and for Y_t, the *leucocyte count* (blue), SDE model and, on the right, the parametric plot of $(u(t), z(t))$ together with the images by u (respectively z) of the critical points of u, u' (respectively z, z')

5 The associated Markov chain model

We now detail the procedure for the third part of the proposed methodology: associating an OMC model to our SDE model. We follow this procedure in two steps:

Step 1 We partition the state space of the SDE in six regions, corresponding to six *Infection Classes*, by means of the critical points of the solutions u and v of the ODE system (3.5) (see Figure 4.1 for a quick identification of these regions and Table 4 for the precise definition of these regions). The determination of a set of critical points of u, u' and z, z' and the corresponding images by the functions u and z, are presented in Table 3. These points correspond, in the right-hand side of Figure 4.1, to a vertical line – for $u(t_a)$ –

Table 3: Images of critical points for classes determination

$\{u(t_a) : u'(t_a) = 0\}$	$\{z(t_b) : z'(t_b) = 0\}$	$\{z(t_c) : z''(t_c) = 0\}$
4.94971	2.37373	2.72319

and two horizontal lines – corresponding to $z(t_b)$ and $z(t_c)$ – allowing for the definition of six regions in the phase space of (u, z) that are indicated in Table 4. These regions admit qualitative descriptions as

Table 4: Regions in the phase space of (u, v) for *Infection Classes* determination

Region 1	Region 2	Region 3
$[u(t_a), +\infty[\times[0, z(t_b)[$	$[0, u(t_a)[\times[0, z(t_b)[$	$[u(t_a), +\infty[\times[z(t_b), z(t_c)[$
Region 4	Region 4	Region 6
$[0, u(t_a)[\times[z(t_b), z(t_c)[$	$[u(t_a), +\infty[\times[z(t_c), +\infty[$	$[0, u(t_a)[\times[z(t_c), +\infty[$

Infection Classes. *Region 6* is the terminal infection state with a higher viral load and a reduced leucocyte

count. *Region 4* corresponds to the infection establishment followed by a noticeable increase of the viral load and the corresponding immune response. *Region 2* is the most reasonable initial infection state due to its definition in the phase space of the SDE and so on and so forth.

Step 2 We simulate trajectories of the SDE (4.1) and we count the transitions between the different regions corresponding to the six *Infection Classes*. That is, considering every pair of regions and by counting the transitions from the first region of the pair to the second region of the pair, for every trajectory of a sample – of (only)16 – simulated trajectories, we get the transition matrix T presented in (5.1). Increasing the number of trajectories does not noticeably change the accuracy of the estimation of the Markov chain transition probabilities.

$$
T = \begin{pmatrix}
0.921265 & 0.0189807 & 0.059051 & 0.000702988 & 0. & 0. \\
0.00867303 & 0.952154 & 0.000578202 & 0.038595 & 0. & 0. \\
0.0458584 & 0.00143308 & 0.881055 & 0.0171969 & 0.053597 & 0.000859845 \\
0.000438837 & 0.03017 & 0.00680197 & 0.914427 & 0.000329128 & 0.0478332 \\
0. & 0. & 0.0187966 & 0.000420036 & 0.971333 & 0.0094508 \\
0. & 0. & 0.0000380105 & 0.00849535 & 0.00157744 & 0.989889
\end{pmatrix} \tag{5.1}
$$

We stress some general methodological remarks for the building up of the Markov chain model associated to the SDE extension of the ODE model:

1. The determination of the state space partition for the determination of the states of the Markov chain model could be done, *a priori*, by using, firstly, the critical points of u and v and after the inflection points, that is the critical points of u' and v'. The interpretation of (u, v) as the mean value functions of the SDE solution trajectories may give a justification for this choice. The computation of these critical points is feasible after the parameters are estimated. This is the approach in the following developments.

2. Another possible way to determine the partition of the phase space, perhaps more adequate for the Markov chain model definition, could be the analysis of the two dimensional histogram of sample of trajectories obtained by discretization of the SDE model as done in section 4. See the right-hand side of Figure 5.1 for a first idea of the occupation of the phase space by the trajectories. Also, a cluster analysis in the plane phase space could give some indications of the classes to be defined *a posteriori*.
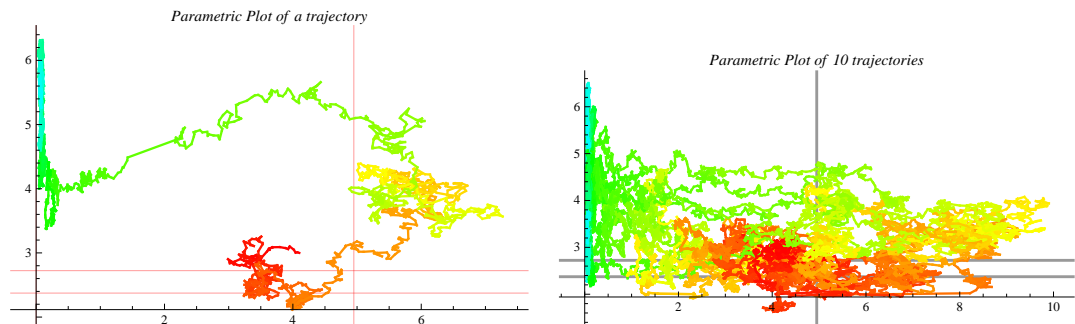

Figure 5.1: One trajectory and 10 superimposed trajectories in the state space of (X_t, Y_t)

3. The simulated trajectories must be determined in a way that preserves the – possible – ergodicity, and other asymptotic properties – of the diffusions discretized solutions of the SDE model.

A quick analysis over matrix T, reveals that the Markov chain corresponding to this transition matrix is irreducible and therefore all states are recurrent. So, it may be that the initial proposition of considering the partition of the state space using the critical points is not adequate in order to obtain a matrix with transient states and possibly one only recurrent state. In any case, the partition of the state space should be performed according to some justification drawn from the phenomena under study.

Remark 11. Besides the classes definition, there are other choices that may have an influence on the values of the transition matrix; these are the discretization step – which in this case was taken as 0.01 – and the discretization scheme chosen. As already said, the choice of the discretization scheme for the SDE has an influence on the properties of the diffusion that are transmitted to the Markov chain, such as the ergodicity.

5.1 The altered Markov chain model associated to the SDE system model

In this subsection, we will justify the assumption that the sixth state – corresponding to the sixth *Infection Class* – is the sole recurrent state and that all the other states are transient. This assumption, besides being justified, is also necessary for the OMC model implementation (see [12] for a full development of these properties). The assumption leads us to change the transition matrix T into $\widetilde{T}$, in (5.2), by simply considering non existing the transitions from class 6 to classes 5, 4 and 3; we stress that in T these transitions have small probabilities, respectively 0.00157744, 0.00849535 and 0.0000380105.

$$
\widetilde{T} = \begin{pmatrix}
0.921265 & 0.0189807 & 0.059051 & 0.000702988 & 0. & 0. \\
0.00867303 & 0.952154 & 0.000578202 & 0.038595 & 0. & 0. \\
0.0458584 & 0.00143308 & 0.881055 & 0.0171969 & 0.053597 & 0.000859845 \\
0.000438837 & 0.03017 & 0.00680197 & 0.914427 & 0.000329128 & 0.0478332 \\
0. & 0. & 0.0187966 & 0.000420036 & 0.971333 & 0.0094508 \\
0. & 0. & 0. & 0. & 0. & 1.
\end{pmatrix}
\tag{5.2}
$$

Remark 12. We again recall that we are not taking into consideration the existence of disease treatments. However, even if we did so, there would always be a recurrent state, namely death, and so this highlights that an OMC model would be adequate to describe viral infected populations.

Remark 13. We could, justifiably, alter matrix T by considering that all entries smaller than a certain threshold – say 0.01, for instance – would be considered zero and the corresponding probability would be summed either to the greatest probability in the row or distributed proportionally through the non zero entries of the row. Any of these choices should be done according to the interpretation of the classes and respective transitions. An experience performed during our study did not show any appreciable changes in the conclusions.

With this altered matrix we now have a Markov chain model with the first 5 states being transient states and the sixth being a recurrent state. The evolution of the proportions of the elements of some closed population – an unique cohort – with initial distribution given by:

$$
\{0.15, 0.75, 0.05, 0.05, 0., 0.\} ,
\tag{5.3}
$$

is presented left-hand side of Figure 6.1 (see [8, p. 280] for a quick mathematical justification). As expected, the proportions in the transient classes – *Infection Classes* 1 to 5 – decrease while the proportion in the *Infection Class* 6 goes to one. This is a natural asymptotic behavior of an usual **closed** population Markov chain model.

Remark 14. The initial distribution (5.3) was chosen according to the following idea: in principle, in a country with a wide reaching health care system, the newly viral infected will be mostly detected whenever belonging to *Infection Class* 2; the one class where there is already a shortening of leucocytes and the presence of viral load. The results of the evolution, that we show, depend, of course on the initial distribution.

Remark 15. With the purpose of showing the quantitative effects of the initial distribution, on the evolution patterns of the population, we considered a second initial distribution given by:

$$
\{0.2, 0.5, 0.15, 0.1, 0.05, 0.\} .
\tag{5.4}
$$

This corresponds to more weight in the *Infection Classes* 1, 3, 4 and 5 and corresponding less weight in *Infection Class* 2. The evolution dependence on the initial distribution highlights the need for an accurate estimation of this initial distribution.

6 Open Markov chain modeling: a review

We now detail and comment the results that will be used in this paper to get the evolution of the newly viral infected in the different *Infection Classes*. The study of open Markov chain models we will present next relies on results and notations that were presented recently in [7] and that we reproduce next, for the readers convenience. We will suppose that, in general, the transition matrix of the Markov chain model may be written in the following form.

$$\mathbf{P} = \begin{bmatrix} \mathbf{K} & \mathbf{U}_1 \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \tag{6.1}$$

where $\mathbf{K}$ is a $k \times k$ transition matrix between transient states, $\mathbf{U}_1$ a $k \times (r - k)$ matrix of transitions between the transient and the recurrent states and $\mathbf{V}$ a $(r - k) \times (r - k)$ matrix of transitions between the recurrent states. A straightforward computation then shows that,

$$\mathbf{P}^{(n)} = \begin{bmatrix} \mathbf{K}^{(n)} & \mathbf{U}_n \\ \mathbf{0} & \mathbf{V}^{(n)} \end{bmatrix} \quad , \quad n \in \mathbb{N}$$

with $\mathbf{U}_n = \mathbf{U}_{n-1}\mathbf{V} + \mathbf{K}^{(n-1)}\mathbf{U}_1 = \sum_{i=0}^{n-1} \mathbf{K}^{(i)} \mathbf{U}_1 \mathbf{V}^{(n-1-i)}$. We write the vector of the initial classification, for a time period i, as

$$\mathbf{c}_i^t = \begin{bmatrix} \mathbf{t}_i^t & \mathbf{r}_i^t \end{bmatrix} \quad , \quad i \in \mathbb{N} \tag{6.2}$$

with $\mathbf{t}_i$ the vector of the initial allocation probabilities for the transient states and $\mathbf{r}_i$ the vector of the initial allocation probabilities for the recurrent states. With (6.1) and (6.2), we now notice that the vector of the Poisson parameters, for the population sizes in each state at time T, may be written as

$$\boldsymbol{\lambda}_T^{++t} = \left[\sum_{i=1}^{T} \lambda_i \mathbf{t}_i^t \mathbf{K}^{(T-i)} \,\middle|\, \sum_{i=1}^{T} \lambda_i \left(\mathbf{t}_i^t \mathbf{U}_{T-i} + \mathbf{r}_i^t \mathbf{V}^{(T-i)} \right) \right] . \tag{6.3}$$

We observe that the first block corresponds to the transient states and the second block, the one in the right-hand side, corresponds to the recurrent states. From now on, as a first restricting hypothesis, we will also suppose that the transition matrix of the transient states, $\mathbf{K}$, is diagonalizable and so,

$$\mathbf{K} = \sum_{j=1}^{k} \eta_j \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t ,$$

with $(\eta_j)_{j \in \{1,\dots,k\}}$ the eigenvalues, $(\boldsymbol{\alpha}_j)_{j \in \{1,\dots,k\}}$ the left eigenvectors and $(\boldsymbol{\beta}_j)_{j \in \{1,\dots,k\}}$ the right eigenvectors of matrix $\mathbf{K}$. We observe that $j \in \{1, \dots, k\}$ corresponds to a transient state if and only if $| \eta_j | < 1$. We may write,

$$\mathbf{K}^{(n)} = \sum_{j=1}^{k} \eta_j^n \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t , \tag{6.4}$$

and so, as a consequence of (6.3), for the vector of the Poisson parameters corresponding only to the transient states, $\boldsymbol{\lambda}_T^{+t}$, we have:

$$\boldsymbol{\lambda}_T^{+t} = \sum_{i=1}^{T} \lambda_i \, \mathbf{t}_i^t \, \mathbf{K}^{(T-i)} = \sum_{j=1}^{k} \sum_{i=1}^{T} \lambda_i \, \eta_j^{T-i} \, \mathbf{t}_i^t \, \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t . \tag{6.5}$$

The main result we will use next is the following.

Theorem 6.1. *Let a Markov chain driven system have a diagonalizable transition matrix between the transient states $\mathbf{K} = \sum_{j=1}^{k} \eta_j \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t$, written in its spectral decomposition form. Suppose the system to be fed by Poisson inputs with intensities $(\lambda_i)_{i \in \mathbb{N}}$ and such that the vector of initial classification of the inputs in the transient states converges to a fixed value, that is, $\lim_{i \to +\infty} \mathbf{t}_i^t = \mathbf{t}_\infty^t \neq \mathbf{0}$. Then, with $\boldsymbol{\lambda}_n^{+t}$ the vector of Poisson parameters of the transient sub-populations, at date $n \in \mathbb{N}$, we have the following:*

1. If $\lim_{n \to +\infty} \lambda_n = \lambda \in \mathbb{R}_+$ then

$$\lambda_\infty^+ = \lim_{n \to +\infty} \lambda_n^{+t} = \sum_{j=1}^{k} \frac{\lambda}{1 - \eta_j} \mathbf{t}_\infty^t \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t \; . \tag{6.6}$$

2. If $\lim_{n \to +\infty} \lambda_n = +\infty$ and there exists a constant $C > 0$ such that

$$\max_{1 \le i \le n} \left| \frac{\lambda_i - \lambda_{i+1}}{\lambda_n} \right| \le C$$

then

$$\lim_{n \to +\infty} \frac{\lambda_n^{+t}}{\lambda_n} = \sum_{j=1}^{k} \frac{1}{1 - \eta_j} \mathbf{t}_\infty^t \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t \; . \tag{6.7}$$

Remark 16. We observe that proportions in the Markov chain transient classes, on both statements of the Theorem 6.1, only depend on the eigenvalues $\eta_j, j = 1, \ldots, k$. In fact, whenever using formula (6.6) to compute proportions these proportions do not depend on the value of λ as we have that,

$$\sum_{j=1}^{k} \frac{\lambda}{1 - \eta_j} \mathbf{t}_\infty^t \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t = \lambda \left[\mathbf{t}_\infty^t \cdot \left(\sum_{j=1}^{k} \frac{1}{1 - \eta_j} \boldsymbol{\alpha}_j \boldsymbol{\beta}_j^t \right) \right] \; ,$$

and the term in the right-hand side multiplying λ is a vector with the dimension equal to the number of transient classes k, which is equal to the dimension of the square matrix $\mathbf{K}$. As so, when computing proportions, by normalizing this vector with the sum of its components, $\lambda \ne 0$ disappears.

6.0.1 The main theorem formula interpretation

This subsection aims at clarifying some aspects of the implementation of the main theorem 6.1.

Let us suppose that we have three square matrices of the same order, $\boldsymbol{B} = [b_{ij}]$ $\boldsymbol{C} = [c_{ij}]$ and a diagonal one $\boldsymbol{D} = [d_{ij}]$. We then have that:

$$\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{D} \cdot \boldsymbol{C} = [a_{ij}]$$

which is, as $\boldsymbol{D} = [d_{ij}]$ is a diagonal matrix and so $d_{kl} = 0$, for $k \ne l$

$$a_{ij} = \sum_k \sum_l b_{ik} d_{kl} c_{lj} = \sum_k b_{ik} d_{kk} c_{kj} = \sum_k d_{kk} \left(b_{ik} c_{kj} \right) \; . \tag{6.8}$$

In [30, p. 111] it is shown that if $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ is any set of orthonormal eigenvectors corresponding, respectively, to distinct eigenvalues $\eta_1, \ldots, \eta_m$ of a matrix $\boldsymbol{A}$ then,

$$\boldsymbol{A} = \sum_{i=1}^{m} \eta_i \boldsymbol{x}_i \boldsymbol{x}_i^t \; .$$

Also, if the eigenvectors are not an orthonormal set, see [30, pp. 169, 170], if the matrix $\boldsymbol{A}$ is diagonalizable with eigenvalues $\eta_1, \ldots, \eta_m$ and the correspondent independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, then:

$$\boldsymbol{A} = \boldsymbol{X} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{X}^{-1} = \sum_{k=1}^{m} \eta_k \boldsymbol{x}_k \cdot \boldsymbol{y}_k^t \; , \tag{6.9}$$

with $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$ and $(\boldsymbol{X}^{-1})^t = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m) := \boldsymbol{Y}$. Now, with $\boldsymbol{X} = [x_{ij}]$ and $\boldsymbol{Y} = [y_{ij}]$ we have that, for a diagonalizable matrix $\boldsymbol{A}$, by formula (6.8),

$$\boldsymbol{A} = \sum_k \eta_{kk} \left(x_{ik} y_{kj} \right)$$

and so $\boldsymbol{x}_k \cdot \boldsymbol{y}_k^t = [x_{ik} y_{kj}]$. This observation is important to compute the relevant values with theorem 6.1.

6.1 The OMC model associated to the SDE system

In an OMC model we consider a permanent inflow of new members in the population which has an evolution driven by a Markov transition matrix. In an OMC model the asymptotic behavior of subpopulations in the transient classes become non trivial and may be quantitatively described. So, in order to study the evolution of the whole population of newly viral infected individuals spread in a long period of time – by an OMC model – we will consider, as an example, simulated data of newly detected with some infection on a period of, say, 33 years, for instance from 1983 to 2016. In the right-hand side of Figure 6.1 we present the fitting of this data with a function such as,

$$g_{a,b,c,d,k}(x) := (x + k)^b e^{a-cx+dx^2}$$ (6.10)

with a, b, c, d, k parameters. The fitted parameters give us the function:

$$g_{\text{fitted}}(x) := \exp\left(0.00356302x^2 - 0.674046x + 13.4741\log(x + 6.2) - 23.9968\right)$$ (6.11)
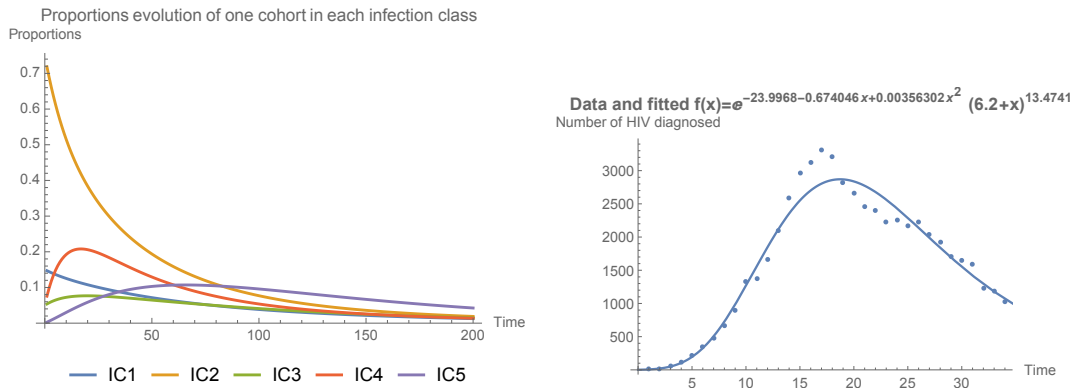

Figure 6.1: One cohort proportions evolution and Portuguese reported infection numbers

Remark 17. The parameters a, b, c, d, k for the best fitting of the function $g_{a,b,c,d,k}$ given in formula (6.10), may be estimated – in a more rigorous way – by maximum likelihood, as done in [7] for a sigmoid type function or in [13] for an exponential type function.

Remark 18. The function g_{fitted} used to fit the evolution of the newly viral diagnosed has a stationary point $x_{\text{stationary}} \approx 70$ and after it becomes increasing. A natural alternative choice in the absence of further information and for illustration purposes lead us to consider the function G, which coincides with the fitted function – see (6.10) – until this stationary point and remains constant for values of the variable greater than this stationary point.

$$G(x) := g_{a,b,c,d,k}(x)\mathbf{1}_{[0,x_{\text{stationary}}[}(x) + g_{a,b,c,d,k}(x_{\text{stationary}})\mathbf{1}_{[x_{\text{stationary}},+\infty[}(x) \,,$$

where the relevant stationary point of $g_{a,b,c,d,k}(x)$ is given, for $k = 6.2$, by

$$x_{\text{stationary}} = \frac{0.25\left(\sqrt{-8.bd + c^2 + 24.8cd + 153.76d^2} + c - 12.4d\right)}{d} = 69.66579746554228 \,.$$

A first outcome of the approach inherent to the OMC model theorem allow us to draw Figure 6.2 which represents the expected number of patients in each of the transition infection classes over time, as well as the proportion of patients in the *Infection Classes* considered as treatable.

Remark 19. We supposed that the time units $t = 30$ presented in Figures 3.3, 4.1, 6.1 correspond to a period of roughly 11 years. This is an intermediate value between approximately 6 and 18 years and is coherent with the period 1983–2016 in the reported newly viral infected data. Being so, the time period in all the Figures presented next, corresponds roughly to a third of a year, that is 4 months.

As an application of Theorem 6.1, namely of formula (6.6) we obtain the estimates of Table 5 and Table 6.
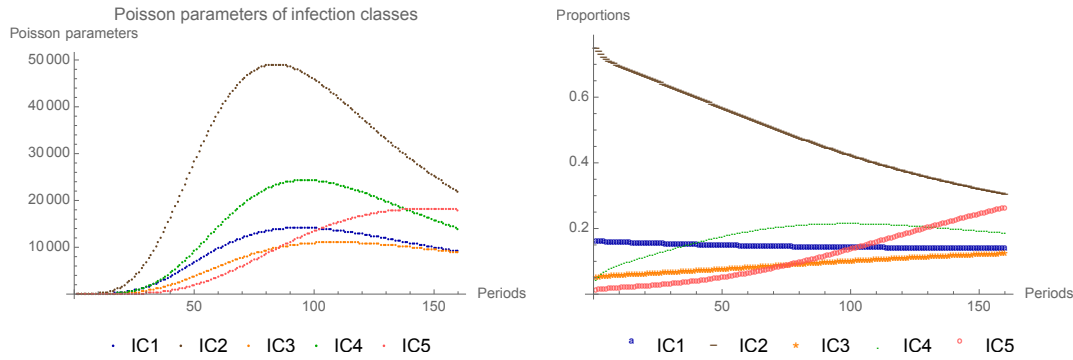

Figure 6.2: Open Markov chain evolution with Portuguese viral newly diagnosed infected numbers: expected numbers of patients – the Poisson parameters – and the correspondent proportions.

Table 5: The expected numbers of patients – the Poisson parameters – and the corresponding proportions for the populations of the *Infection Classes* IC1 to IC5 in case of function G at date $t \to +\infty$.

	IC1	IC 2	IC3	IC4	IC5
Expected numbers	1213.62	3313.8	1091.9	1796.24	2062.05
Proportions	0.13189	0.360128	0.118662	0.195206	0.224094

Table 6: The expected numbers of patients – the Poisson parameters – and the corresponding proportions for the populations of the *Infection Classes* IC1 to IC5 in case of function $g_{a,b,c,d}$ at date $t \to +\infty$.

	IC1	IC 2	IC3	IC4	IC5
Expected numbers	11.4119	31.1603	10.2673	16.8904	19.3899
Proportions	0.13189	0.360128	0.118662	0.195206	0.224094

When the number on newly diagnosed increases without upper bound from date 70 – but under the condition stated in the second thesis of Theorem 6.1 – formula (6.7) shows that nevertheless the mean value of number of elements in each *Infection Class*, divided by mean value of the total newly diagnosed at each date, also stabilizes as a constant value. This can be seen in Table 6 where we present the correspondent results.

Preliminary conclusions 6.1. Let us observe the following preliminary conclusions on the application results of the open Markov chain model that are represented in Figure 6.2 and in Tables 5 and 6.

1. As a consequence of the assumption made that the newly diagnosed viral infected come from a Poisson distributed independent populations in each year, the populations in the different *Infection Classes* 1 to

5, are also Poisson distributed and the evolution of the Poisson parameters for these populations - and so also the expected number of these populations - is what is shown in the right-hand side of Figure 6.2.

2. The first obvious conclusion is that the evolution of the proportions in the transient classes corresponding to the different cohorts – dates 1983–2016 – is quite different for the proportions evolution of one only cohort as shown in Figure 6.1.

3. The evolution of population in *Infection Class* 2, clearly, reflects the best the evolution pattern of the newly diagnosed viral infected as shown in Figure 6.1. This a direct consequence of the initial distribution in which we assumed that 75% of the newly infected would be in this *Infection Class* 2.

4. As a consequence of Remark 16, the proportions of the elements in each of the *Infection Classes* are the same either when the newly diagnosed are in a constant expected number from date 70 on or when the expected number of the newly diagnosed grows without bound from date 70 on. And so, the bottom half of both Tables 5 and 6 are equal, despite being computed by averaging different vectors.

7 Conclusions

In this work we presented a method for associating an evolution model at individual level – given by an ODE system – to a Markov chain evolution model for an open population of the aforementioned individuals. We illustrated the methodology with an example at the individual level, we modeled a viral infection and considered the viral load and a leucocyte count as variables of interest. We showed how to extend the study, based on individuals measurements feeding parameter estimations, to the evolution of a varying population of newly infected individuals divided in five transient *Infection Classes*; for that purpose we used simulated data. We achieved a coherent link between the evolutions at individual level and at whole population level.

Author Statement
Conflict of interest: Authors state no conflict of interest.

References

[1] Jose G. Castro, Gabriel Manzi, Luis Espinoza, Michael Campos, and Catherine Boulanger. Concurrent pcp and tb pneumonia in hiv infected patients. *Scandinavian Journal of Infectious Diseases*, 39(11-12):1054–1058, 2007.

[2] A.D.D. Craik and H. Okamoto. A three-dimensional autonomous system with unbounded 'bending' solutions. *Physica D*, 164(3-4):168–186, 2002.

[3] S. Debroy and M. Martcheva. Immuno-epidemiology and hiv/aids: A modeling prospective. In L.B. Wilson, editor, *Mathematical Biology Research Trends*, chapter 6, pages 175–192. Nova Science Publishers, New York, 2008.

[4] Susanne Ditlevsen, Kay-Pong Yip, and Niels-Henrik Holstein-Rathlou. Parameter estimation in a stochastic model of the tubuloglomerular feedback mechanism in a rat nephron. *Mathematical Biosciences*, 194(1):49 – 69, 2005.

[5] Sophie Donnet and Adeline Samson. A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. *Advanced Drug Delivery Reviews*, 65(7):929 – 939, 2013. Mathematical modeling of systems pharmacogenomics towards personalized drug delivery.

[6] A.M. Elaiw. Global properties of a class of hiv models. *Nonlinear Analysis: Real World Applications*, 11(4):2253 – 2263, 2010.

[7] Manuel L. Esquível, José M. Fernandes, and Gracinda R. Guerreiro. On the evolution and asymptotic analysis of open Markov populations: application to consumption credit. *Stoch. Models*, 30(3):365–389, 2014.

[8]   M.L. Esquível, G.R. Guerreiro, and J.M. Fernandes. Open Markov chain scheme models. *REVSTAT*, 15(2):277–297, 2017.

[9]   L. Ferrante, S. Bompadre, L. Leone, and M. P. Montanari. A stochastic formulation of the gompertzian growth model for in vitro bactericidal kinetics: Parameter estimation and extinction probability. *Biometrical Journal*, 47(3):309–318, 2005.

[10] Michael A. Gilchrist and Akira Sasaki. Modeling host–parasite coevolution: A nested approach based on mechanistic models. *Journal of Theoretical Biology*, 218(3):289 – 308, 2002.

[11] A. F. Gribov, I. K. Volkov, and A. P. Krishchenko. Qualitative analysis of a three-dimensional population evolution model with possible nonequilibrium size preservation. *Computational Mathematics and Modeling*, 11(2):119–134, Apr 2000.

[12] Gracinda Rita Guerreiro and Jo ao Tiago Mexia. Stochastic vortices in periodically reclassified populations. *Discuss. Math., Probab. Stat.*, 28(2):209–227, 2008.

[13] Gracinda Rita Guerreiro, João Tiago Mexia, and Maria de Fátima Miguens. Statistical approach for open bonus malus. *ASTIN Bulletin*, 44(1):63–83, 2014.

[14] Barbara Hellriegel. Immunoepidemiology – bridging the gap between immunology and epidemiology. *Trends in Parasitology*, 17(2):102 – 106, 2001.

[15] Radu Herbei, Rajib Paul, and L. Mark Berliner. Applying diffusion-based Markov chain Monte Carlo. *PLOS ONE*, 12(3):1–14, 03 2017.

[16] AVM Herz, S Bonhoeffer, RM Anderson, RM May, and MA Nowak. Viral dynamics in vivo: Limitations on estimates of intracellular delay and virus decay. *Proceedings of the National Academy of Sciences of the USA*, 93(14):7247–7251, JUL 9 1996.

[17] Thorsten Hüls and Christian Pötzsche. Qualitative analysis of a nonautonomous Beverton-Holt Ricker model. *SIAM J. Appl. Dyn. Syst.*, 13(4):1442–1488, 2014.

[18] João Sollari Lopes, Paula Rodrigues, Suani TR Pinho, Roberto FS Andrade, Raquel Duarte, and M. Gabriela M. Gomes. Interpreting measures of tuberculosis transmission: a case study on the portuguese population. *BMC Infectious Diseases*, 14(1):340, Jun 2014.

[19] Jonathan C. Mattingly, Andrew M. Stuart, and M. V. Tretyakov. Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM Journal on Numerical Analysis*, 48(2):552–577, 2010.

[20] Yuliia Mishura, S.V. Posashkova, and Georgiy Shevchenko. Properties of solutions of stochastic differential equations with nonhomogeneous coefficients and non-lipschitz diffusion. *Theor. Probability and Math. Statist. (Teoriya Jmovirnostej ta Matematychna Statystyka)*, 79:117–126, 12 2009.

[21] Yuliia Mishura and Svitlana V. Posashkova. Positivity of solution of nonhomogeneous stochastic differential equation with non-lipschitz diffusion. *Theory of Stochastic Processes*, 14(3-4):77–88, 2008.

[22] MA Nowak, RM May, RE Phillips, S Rowlandjones, DG Lalloo, S Mcadam, P Klenerman, B Koppe, K Sigmund, CRM Bangham, and AJ Mcmichael. Antigenic Oscillations and Shifting Immunodominance in HIV-1 Infections. *Nature*, 375(6532):606–611, JUN 15 1995.

[23] Alan S. Perelson, Avidan U. Neumann, Martin Markowitz, John M. Leonard, and David D. Ho. Hiv-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586, 1996.

[24] Lawrence Perko. *Differential equations and dynamical systems. 3rd ed.* New York, NY: Springer, 3rd ed. edition, 2001.

[25] Umberto Picchini, Susanne Ditlevsen, and Andrea De Gaetano. Modeling the euglycemic hyperinsulinemic clamp by stochastic differential equations. *Journal of Mathematical Biology*, 53(5):771–796, Nov 2006.

[26] Elizabeth L. Read, Allison A. Tovo-Dwyer, and Arup K. Chakraborty. Stochastic effects are important in intrahost hiv evolution even when viral loads are high. *Proceedings of the National Academy of Sciences*, 109(48):19727–19732, 2012.

[27] Roland R. Regoes, Dominik Wodarz, and Martin A. Nowak. Virus dynamics: the effect of target cell limitation and immune responses on virus evolution. *Journal of Theoretical Biology*, 191(4):451 – 462, 1998.

[28] Sidney I. Resnick. *Adventures in stochastic processes.* Boston, MA: Birkhäuser, 1992.

[29] Gareth O. Roberts and Richard L. Tweedie. Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363, 12 1996.

[30] James R. Schott. *Matrix analysis for statistics. 3rd edition.* Hoboken, NJ: John Wiley & Sons, 3rd edition edition, 2017.

[31] T. Shardlow and A. M. Stuart. A perturbation theory for ergodic markov chains and application to numerical approximations. *SIAM Journal on Numerical Analysis*, 37(4):1120–1137, 2000.

[32] Hal L. Smith and Patrick De Leenheer. Virus dynamics: A global analysis. *SIAM Journal on Applied Mathematics*, 63(4):1313–1327, 2003.

[33] Denis Talay. Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics and Stochastic Reports*, 29(1):13–36, 1990.

[34] Christoffer W. Tornøe, Judith L. Jacobsen, Oluf Pedersen, Torben Hansen, and Henrik Madsen. Grey-box modelling of pharmacokinetic /pharmacodynamic systems. *Journal of Pharmacokinetics and Pharmacodynamics*, 31(5):401–417, Oct 2004.

[35] Christoffer Wenzel Tornøe, Judith L. Jacobsen, and Henrik Madsen. Grey-box pharmacokinetic/pharmacodynamic modelling of a euglycaemic clamp study. *Journal of Mathematical Biology*, 48(6):591–604, Jun 2004.