Math. Model. Nat. Phenom.
Vol. 12, No. 5, 2017, pp. 63–77
DOI: 10.1051/mmnp/201712505

Modelling Stochastic and Deterministic Behaviours in Virus Infection Dynamics

I. Sazonov[1*], D. Grebennikov[2,3], M. Kelbert[1,4], G. Bocharov[3]

[1] Swansea University, Swansea, U.K.

[2] Moscow Institute of Physics and Technology (State University), Dolgoprudny, R.F.

[3] Institute of Numerical Mathematics of the RAS, Moscow, R.F.

[4] National Research University Higher School of Economics, Moscow, R.F.

Abstract. Many human infections with viruses such as human immunodeficiency virus type 1 (HIV–1) are characterized by low numbers of founder viruses for which the random effects and discrete nature of populations have a strong effect on the dynamics, e.g., extinction versus spread. It remains to be established whether HIV transmission is a stochastic process on the whole. In this study, we consider the simplest (so-called, 'consensus') virus dynamics model and develop a computational methodology for building an equivalent stochastic model based on Markov Chain accounting for random interactions between the components. The model is used to study the evolution of the probability densities for the virus and target cell populations. It predicts the probability of infection spread as a function of the number of the transmitted viruses. A hybrid algorithm is suggested to compute efficiently the dynamics in state space domain characterized by a mix of small and large species densities.

Keywords and phrases: mathematical model, virus infection, stochastic dynamics, Markov Chain, hybrid modelling

Mathematics Subject Classification: 35Q53, 34B20, 35G31

1. Introduction

Mathematical models are extensively used for studying the dynamics of virus infections in humans [25]. A vast majority of studies are based on a deterministic descriptions of the virus and immune cell population dynamics. Reviews of the existing approaches have been published recently [8, 9, 11, 29]. However, the initial infection with viruses such as human immunodeficiency virus type 1 (HIV–1) is characterized by low numbers for which the random effects and discrete nature of populations contribute significantly to the dynamics, e.g., extinction versus spread [24, 27].

There has been a substantial effort in the mathematical modelling of HIV infection primarily motivated by the AIDS epidemic (for reviews we refer to [3, 25, 28, 29]). Mathematical modelling of the virus-host interaction based on understanding of the underlying biological processes can provide deeper insights into the kinetic mechanisms of infection dynamics, cf. [1–3, 14, 16]. Most of the models are deterministic

*Corresponding author. E-mail: i.sazonov@swansea.ac.uk

and deal with continuous variables. This is correct when the numbers of all interacting components: uninfected cells and infected cells, virions, etc, are large enough. The dynamics of large populations is described by the chemical Langevin equations, which is a set of stochastic differential equations [40]. In the full thermodynamic limit, i.e. when the species populations and the system volume approach infinity in a way that species concentrations remain constant, the Langevin equations reduce to the traditional reaction-rate equations in the form of coupled, first-order, ordinary differential equations. The hydrodynamic approximation (i.e. dynamical form of the Law of Large Numbers) is the standard toolkit for studying stochastic systems with a large number of interacting components (see, e.g., [12] and the references therein). Virus infections are ideal systems for probing this approach and revealing its advantages and shortcomings versus stochastic kinetic analysis.

However, it typically occurs at the earlier stage of infection (eclipse phase in HIV) that the initial numbers of virions and infected cells are small [18]. During initial infection, a complex quasispecies of viruses has been shown to be restricted through a transmission bottleneck into a single or limited number of founder viruses that establish a systemic infection [43]. Depending on the route of infection the transition probability of systemic infection ranges from 0.2 to 0.5×10^{-3} [37]. It was pointed out that it remains to be determined if HIV transmission is largely a stochastic process. To clarify this type of questions, the continuous virus dynamics models are inefficient and it is necessary to account for the discrete nature of the components and the random interaction between them [24,27]. While a deterministic model predicts a single outcome for a given set of parameter values, a stochastic model would allow one to examine a spectrum (infinite set) of possible infection outcomes weighed by their likelihoods and probabilities. For example, if the deterministic models give the standard disease dynamics for any initial amount of viruses, in reality there is a probability that for small number of initial virions the disease can fade [22]. The smaller the number of virions, the higher should be this probability.

In this study, we consider the simplest virus dynamics model proposed in [25,28] and build an equivalent stochastic model based on Markov Chain accounting for random interaction between the components. The model is studied numerically, and the time-dependent probability densities for every components are calculated.

As in the case of developing infection, a small number in some components can occur only at the very beginning of the virus transmission followed by larger densities of the interacting components. To compute efficiently the process dynamics, we develop algorithm for hybrid modelling, where the discrete nature of the components is accounted at the initial stage of the infection only, and the deterministic model (with random initial conditions) is explored at the later stages. Such hybrid models have been previously developed for mathematical epidemiology; see, e.g., [31,32,34,35]. The use of the hybrid model essentially simplifies the analysis of the virus dynamics and calculation of the relevant moments characterizing the dynamics uncertainty.

In Section 2 we describe a simple deterministic model of HIV infection. The formulation of the stochastic model is presented in Section 3. In Section 4 the hybrid modelling algorithm is developed. We summarize the results in Section 5.

2. Deterministic model

2.1. Governing equations

The following basic equations for the virus dynamics have been derived in [25, 28]:

$$\begin{aligned} \dot{x} &= \lambda - dx - \beta xv \\ \dot{y} &= \beta xv - ay \\ \dot{v} &= ky - uv \end{aligned} \tag{2.1}$$

where x and y be the number of uninfected cells and infected cells, respectively; v be the number of free virus particles (virions). The authors suppose that uninfected cells are produced at a constant rate, λ,

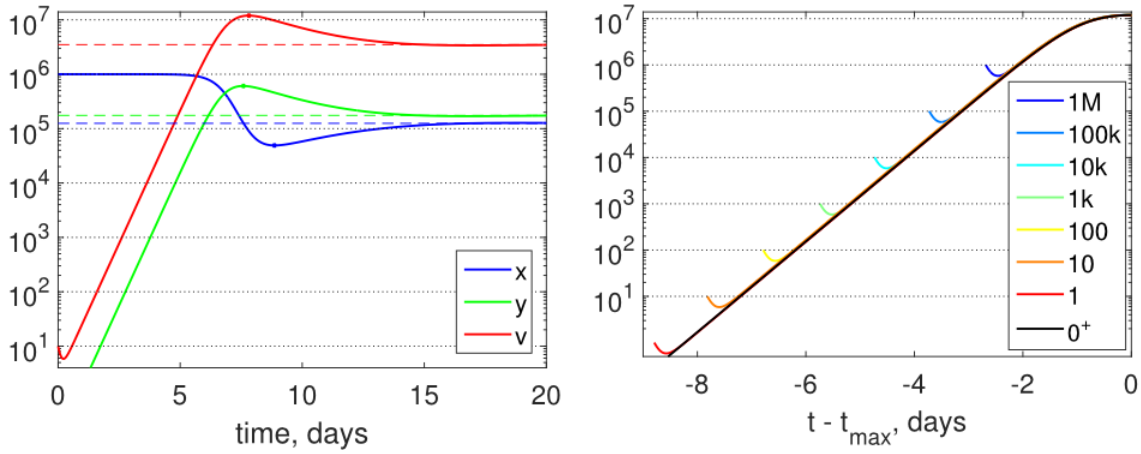

FIGURE 1. Left: Dynamics of uninfected cells (blue), infected cells (green), and virions (red). Dashed lines indicate the correspondent equilibrium values: x^*, y^*, v^*.
Right: Dynamics of the virions number for different initial numbers v_0 (is indicated in the legend).

and die at a rate dx; virions infect uninfected cells at a rate proportional to the product of their numbers, βxv; infected cells produce free virus at a rate ky; infected cells die at a rate ay; free virus particles are removed from the system at a rate uv. Apart from an infection-free equilibrium $\{x = x_0, y = 0, v = 0\}$, where

$$x_0 = \frac{\lambda}{d}, \tag{2.2}$$

the system has a globally stable equilibrium $\{x = x_*, y = y_*, v = v_*\}$ where

$$x_* = \frac{au}{\beta k}, \qquad y_* = \frac{\lambda}{a} - \frac{du}{\beta k}, \qquad v_* = \frac{\lambda\kappa}{au} - \frac{d}{\beta} \tag{2.3}$$

(see, e.g., [19]). We assume that $\frac{\lambda}{a} > \frac{du}{\beta k}$ and $\frac{\lambda\kappa}{au} > \frac{d}{\beta}$. It is natural to take $x_0 = \lambda/d$ as an initial number of uninfected cells in the virus dynamics and consider the following initial conditions

$$x(t_0) = \lambda/d = x_0, \qquad y(t_0) = 0, \qquad v(t_0) = v_0 \tag{2.4}$$

where t_0 is an instant when the infection begins.

2.2. Stages of virus dynamics

An example of numerical solution to initial value problem (2.1)–(2.4) for the case: $\lambda = 10^5$, $d = 0.1$, $a = 0.5$, $\beta = 2 \times 10^{-7}$, $k = 100$, $u = 5$ (the data are taken from [25]) and initial number of the virus particles $v_0 = 10$ is shown in Figure 1(left). Considering the dimensionless populations in the model, parameters d, a, β, k, u have dimensions of inverse days and $t_0 = 0$.

As it is seen from the plot, the total virus dynamics can be split into four distinguished stages:

Stage 1 is the initial stage. Details of this stage are better seen in Figures 1(right) and 2. At this stage the number of virions varies non-monotonically and reaches its minimum. For the data taken from [25] this stage lasts approximately 0.5 day.

At stage 2 there is an exponential growth of the numbers of virions and infected cells. At stage 2, plots of $y(t)$ and $v(t)$ are close to straight lines with positive slope (in the logarithmic scale) and, as it is seen in Figure 1, this stage occupies time interval approximately 0.5—6 day.

Stage 3 represents a developed disease when the number of virions and infected cells reach its maximum. Number of uninfected cells varies significantly at this stage. It is seen in Figure 1 that this stage occupies time interval approximately 6—13 days.

At stage 4 the numbers of cells and virions gradually approach their stationary state $\{x_*, y_*, v_*\}$. For the data taken from [25] and $v_0 = 1$ this stages begins after 13 days from the inserting the infection.

Observe that at stages 1 and 2, the number of uninfected cells varies insignificantly.

2.3. Approximate solution for small initial number of virus particles

Curves of virions dynamics $v(t)$ for different initial numbers of virions are shown in Figure 1(right). The curves are aligned to reach maximum at the same instant $t = 0$.

2.3.1. Limiting solution

Observe that apart from the initial dynamics all the curves are close to each other and can be approximated by the limiting solution depicted by the black line in Figure 1(right). To calculate this solution, $v^{\lim}(t)$, we should take the initial amount of virions to be infinitesimal, then the infection peak will be reached in infinitely large time. Therefore to describe the limiting solution, it is natural to reference the time from the infection peak instant as it is done in Figure 1(right) i.e. $v^{\lim}(0) = \max_t v^{\lim}(t)$, and, hence, we have to tend the infection beginning instant, t_0, to $-\infty$. Thus the limiting solution can be obtained by solving the initial value problem (2.1),(2.4) in the following limit:

$$v_0 \to +0, \qquad t_0 \to -\infty, \qquad v(0) \to \max. \tag{2.5}$$

The limiting solution is independent of initial condition. It can be easily computed numerically, and can be used in building the hybrid stochastic model (see Section 4).

As it is seen in Figure 1(right), the limiting solution has growing asymptotic behaviour for large negative time. Its growing rate is computed below (see Eq. (2.7)). For the real solutions, it is an intermediate asymptotic behavior corresponding to stage 2.

The limiting solution satisfying (2.5) is similar to limiting solution introduced for epidemic models and called in [33] the Small Initial Contagion (SIC) approximation. It also has an exponential asymptotic behaviour for $t \to -\infty$. In the SIR epidemic model, the limiting solution can approximate all the stages of epidemic dynamics provided the initial number of infective species is small enough. In the virus dynamics, the limiting solution approximates the real solution only at stages 2,3,4. Stage 1 requires a separate consideration which is presented below.

2.3.2. Initial stage dynamics

The initial stage dynamics can be studied if we set $x(t) = const = x_0$: as we mention above at stages 1 and 2 the relative change in the number of uninfected cells is much smaller compared to virus and infected cell populations (see Figure 1). In contrast to the limiting solution, it is convenient to reference the time from the infection beginning, i.e. to set $t_0 = 0$. Substituting $x = x_0$ into (2.1) and solving the equations for initial conditions $y(0) = 0$, $v(0) = v_0$ by the Laplace transform method we obtain the following approximation of the virus dynamics at stage 1:

$$\begin{aligned} v(t) &= \frac{v_0}{2\rho}\left[(a-u+\rho)\,e^{\lambda_1 t} + (u-a+\rho)e^{\lambda_2 t}\right] \\ y(t) &= \frac{v_0\beta x_0}{\rho}\left[e^{\lambda_1 t} - e^{\lambda_2 t}\right] \end{aligned} \tag{2.6}$$

where $\rho = \sqrt{(a-u)^2 + 4k\beta x_0}$, $\lambda_{1,2} = \frac{1}{2}(-a-u\pm\rho)$. Since $\lambda_1 > 0$ and $\lambda_2 < 0$, then for $t \to +\infty$ solutions (2.6) grow exponentially:

$$v(t) \approx \frac{v_0}{2\rho}(a-u+\rho)\,e^{\lambda_1 t}, \qquad y(t) \approx \frac{v_0\beta x_0}{\rho}e^{\lambda_1 t}. \tag{2.7}$$

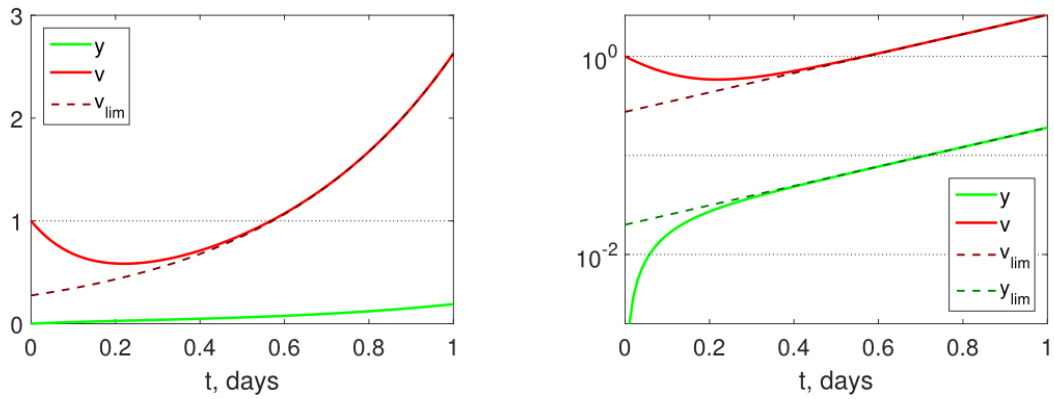

FIGURE 2. Initial virus dynamics (stage 1 and beginning of stage 2) for parameters indicated in [25] in the linear (left) and logarithmic (right) scales. Red: $v(t)/v_0$, green: $y(t)/v_0$, dark read and green dashed: limiting solution aligned with $v(t)/v_0$.

For any sufficiently small initial number of virus particles ($v_0 \ll x_0$) the initial dynamics looks very similar. The number of virions slightly decreases then starts to increase tending to almost the same exponentially growing curve.

For the parameter values indicated in [25] we have

$$v(t)/v_0 = 0.275\, e^{2.256t} + 0.725\, e^{-7.756t}$$
$$y(t)/v_0 = 0.0200\left(e^{2.256t} - e^{-7.756t}\right).$$

Thus the virus dynamics can be approximated by a combination of the initial solution (2.6) and the limiting solution. The limiting solution is independent of the initial condition, the same is true for $v(t)/v_0$ and $y(t)/v_0$ at the initial stage, i.e. the plots of the both solutions have universal shapes. The solutions are matched at stage 2 (cf Eq. (2.7)). Only the transition time from the initial stage solution to the limiting solution depends on the initial number of virions as it is seen in Figure 1(right).

3. Stochastic model

3.1. Markov chain formulation

Now we account for the discrete nature of the populations and random interaction between different species.

Let $X, Y, V \in \mathbb{Z}^+$ be non-negative integer numbers of uninfected, infected cells and virus particles, respectively. Then we assume that their dynamics obeys the following continuous time Markov chain (MC)

Process #	Event	Rate	Description
1	$X \to X + 1$	λ	Uninfected cell birth
2	$X \to X - 1$	dX	Uninfected cell death
3	$X \to X - 1,\ Y \to Y + 1$	βXV	Cell infection
4	$Y \to Y - 1$	aY	Infected cell death
5	$V \to V + 1$	kY	Virus particle birth
6	$V \to V - 1$	uV	Virus particle death

(3.1)

The initial conditions for MC (3.1) are

$$X(0) = X_0, \qquad Y(0) = 0, \qquad V(0) = V_0 \in \mathbb{Z}^+ \tag{3.2}$$

where X_0 is a random integer number having the Poisson distribution with the rate $x_0 = \lambda/d$, as it will be shown in Section 3.2.

The PDF for time-dependent number $X(t)$ is calculated in Section 3.2 by considering the pre-infected dynamics and it asymptotic behaviour for large time.

3.2. Uninfected target cell random dynamics

In the absence of virus the random dynamics of the target cells is described by the simplified MC

Process #	Event	Rate	Description
1	$X \to X + 1$	λ	Uninfected cell birth
2	$X \to X - 1$	dX	Uninfected cell death

(3.3)

subjected to the initial condition

$$X(0) = \bar{X} \in \mathbb{Z}^+. \tag{3.4}$$

Let $P_m(t)$, $m \in \mathbb{Z}$, be a probability that the number of uninfected cells equals $X(t) = m$ and let $P_m \equiv 0$ if $m < 0$. Then Kolmogorov's/Master equation takes the form

$$\dot{P}_m = \lambda P_{m-1} - \lambda P_m + d(m+1)P_{m+1} - dmP_m \tag{3.5}$$

with the initial condition $P_m(0) = \delta_{m\bar{X}}$ where δ_{ij} is Kroneker's delta.

Introduce the probability generation function (PGF):

$$g(t,z) = \sum_{m=0}^{\infty} z^m P_m(t). \tag{3.6}$$

At time zero this function equals

$$g(0,z) = z^{\bar{X}}. \tag{3.7}$$

Multiplying the both sides of (3.5) by z^m and summing from 0 to infinity after some manipulations we obtain the following first order PDE for $g(t,z)$:

$$\frac{\partial g}{\partial t} = (1-z)\left(d\frac{\partial g}{\partial z} - \lambda g\right) \tag{3.8}$$

with initial condition (3.7). If we apply the Laplace transform with respect to time to Eq. (3.8) with account for (3.7) we obtain a first order ODE with separable variables which can be easily integrated. After applying the inverse Laplace transform to the result of integration of the ODE, we obtain the solution to initial value problem (3.8)–(3.7):

$$g(t,z) = \left[1 + (z-1)e^{-dt}\right]^{\bar{X}} \exp\left\{\frac{\lambda}{d}(z-1)\left(1 - e^{-dt}\right)\right\}$$

When $t \to \infty$, it tends to

$$g(+\infty, z) = \exp\left\{\frac{\lambda}{d}(z-1)\right\}$$

that is a PGF for the Poisson distribution with the rate $x_0 = \lambda/d$.

$$P(X = n) = \frac{x_0^n}{n!}e^{-x_0}. \tag{3.9}$$

Its mean and variance are both equal to x_0. It can be approximated by a normal distribution for $x_0 \gg 1$.

3.3. Fluid dynamic limit type analysis

In the case of developed infection, when the populations of all species tend to be large, the stochastic dynamics described by MC (3.1) being properly scaled tends (in probability) to a deterministic dynamics described by (2.1)–(2.4). This limiting transition is called fluid dynamics limit [10] or mean field limit [5], and Theorem 3.1 provides the formal mathematical framework.

Theorem 3.1. *Consider the Markov chain* $\{X(t), Y(t), V(t)\}$ *defined in Table 1 and subjected to initial conditions (3.2). The scaled MC* $\left\{\hat{X}(t), \hat{Y}(t), \hat{V}(t)\right\}$

$$\hat{X}(t) = \Lambda^{-1}X(t), \qquad \hat{Y}(t) = \Lambda^{-1}Y(t), \qquad \hat{V}(t) = \Lambda^{-1}V(t)$$

obtained by scaling the transition rates $\beta \to \Lambda\beta$*, and scaling of initial conditions as:*

$$\hat{X}(0) = \Lambda X(0), \qquad \hat{Y}(0) = \Lambda Y(0), \qquad \hat{V}(0) = \Lambda V(0)$$

converges in distribution as $\Lambda \to \infty$ *to the deterministic functions* $\{x(t), y(t), v(t)\}$ *satisfying (2.1) subjected to initial conditions (2.4).*

Sketch of the proof is given in Appendix A.

3.4. Direct numerical simulation

Analytical analysis based on Kolmogorov's/Master equation and PDE for probability generating function, analogous to that described in Section 3.2, results in a non-linear PDE that cannot be solved in a closed form. Therefore we will study the virus dynamics described by MC (3.1) numerically.

Consider a Poisson process with the constant rate ν. Let the previous event occurs at instant t_0. Probability that a new event occurs before the time $t \geq t_0$ can be estimated as

$$P(t|t_0, \nu) = 1 - e^{-\nu(t-t_0)}. \tag{3.10}$$

This gives a clue to model this process using a random number generator which generates a random number r uniformly distributed in the segment [0,1]. To find time of the next event t_r we equate the l.h.s. of the (3.10) to r and solve the equation with respect to t. This gives

$$t_r = t_0 + \frac{1}{\nu}\ln\frac{1}{1-r}. \tag{3.11}$$

For variable rate Eq. (3.10) should be modified

$$P(t|t_0, \nu) = 1 - \exp\{-\int_{t_0}^{t} \nu(t')\,dt'\}. \tag{3.12}$$

Equation for evaluation of the next event time t_r becomes more complicated.

In the case of Markovian process, the rate $\nu(t)$ is a piecewise constant function. This is because the rate of any process in the MC depends on the number of species: it remains constant in between the events and then 'jumps' to the next value when a next event occurs. So it is possible to evaluate the next event time t_r. Let the value of integral in (3.12) be known $\int_{t_0}^{t_1} \nu(t)\,dt = J$ and for $t > t_1$ the rate remains constant $\nu(t > t_1) = \nu_1$ then

$$t_r = t_1 + \frac{1}{\nu_1}\Big(\ln\frac{1}{1-r} - J\Big). \tag{3.13}$$

that simplifies computation of t_r.

To describe the numerical algorithm modelling MC (3.1) we introduce the state vector $\mathbf{X} = [X_1, X_2, X_3]^T \equiv [X, Y, V]^T$ and matrix of events:

$$\mathbf{M} = \begin{bmatrix} +1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & +1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & +1 & -1 \end{bmatrix} \tag{3.14}$$

If an nth process occurs at a certain instant, then the nth column of matrix $\mathbf{M}$ should be added to the state vector $\mathbf{X}$.

Also we introduce a vector of event rates for all the processes $\nu = [\nu_1, \ldots, \nu_6]$, a vector of expected event time for all the processes $\mathbf{t} = [t_1, \ldots, t_6]$, a vector of integrals $\mathbf{J} = [J_1, \ldots, J_6]$ and a vector of generated random numbers $\mathbf{r} = [r_1, \ldots, r_6]$ uniformly distributed on $[0, 1]$.

Algorithm 3.2. *The numerical algorithm modelling stochastic virus dynamics.*

1. *Set the final time of the process t_f and set current time $t = 0$;*
2. *Initialize: $\mathbf{X}(0) = [X_0, 0, V_0]$ and set $\mathbf{J} = [0, \ldots, 0]$;*
3. *Compute the rate vector ν using (3.1);*
4. *Generate vector of random numbers $\mathbf{r}$;*
5. *Compute the expected times of next event in all the processes: $t_n = \dfrac{1}{\nu_n} \ln \dfrac{1}{1 - r_n}$, $n = 1, \ldots, 6$;*
6. *Find process p with the minimal expected time: $t_p = \min\{t_1, \ldots, t_6\}$;*
7. *Update the state vector $X_m \leftarrow X_m + M_{mp}$, $m = 1, 2, 3$.*
8. *Update the $\mathbf{J}$ vector: $J_n \leftarrow J_n + \nu_n(t_p - t), n = 1, \ldots, 6, n \neq p$ and set $J_p = 0$;*
9. *Update the rate vector ν using (3.1);*
10. *Update the expected times: $t_n = t_p + \dfrac{1}{\nu_n}\Big(\ln \dfrac{1}{1 - r_n} - J_n\Big)$, $n = 1, \ldots, 6$, $n \neq p$;*
11. *Generate a random number $r_p \in [0, 1]$;*
12. *Update the next event time of the pth process: $t_p \leftarrow t_p + \dfrac{1}{\nu_p} \ln \dfrac{1}{1 - r_p}$;*
13. *Set current time $t = t_p$ and store the current state and time*
14. *If $t < t_f$ go to 6 otherwise terminate the computation.*

The proposed algorithm is similar to Gillespie's "First Reaction" algorithm [13].

3.5. Full stochastic numerical simulations results

Algorithm 3.2 for the direct numerical simulations of MC (3.1) has been implemented in the C language with the use of the PCG library [26] for random number generator. The script calling the program in loop to obtain 10k non-degenerate realizations (in which there is no extinction of disease) for various V_0 is written in bash. It utilizes the GNU parallel tool [38] to parallelize executions of different realizations on a maximum available number of threads. Typical CPU time to execute a single realization is about 260 s, i.e., 4 min 20 s.

An example of numerical simulation is presented in Figure 3(left). One can see that the stochastic behaviour is very typical for the earlier stage of the infection.

Note that in contrast to the deterministic case, the infection dynamics can instinct in some realizations, i.e., it can reach the state $V = 0, Y = 0$ at a certain time after which the virus dynamics is terminated. As usual this occurs at a small time when values V, Y are still not too large. We refer to such realizations as degenerate ones. A realization in which numbers of virions and infected cells can reach their peak values is a non-degenerate realization and corresponds to a developed infection disease. The probability for infection to develop strongly depends on the initial number of virions V_0. The numerical simulations enables computation of the probability to develop infection depending on initial number of virus particles V_0. The plot is shown in Figure 3(right). From the plot it is seen that for $V_0 \approx 20$ the probability of extinction is around 0.5.

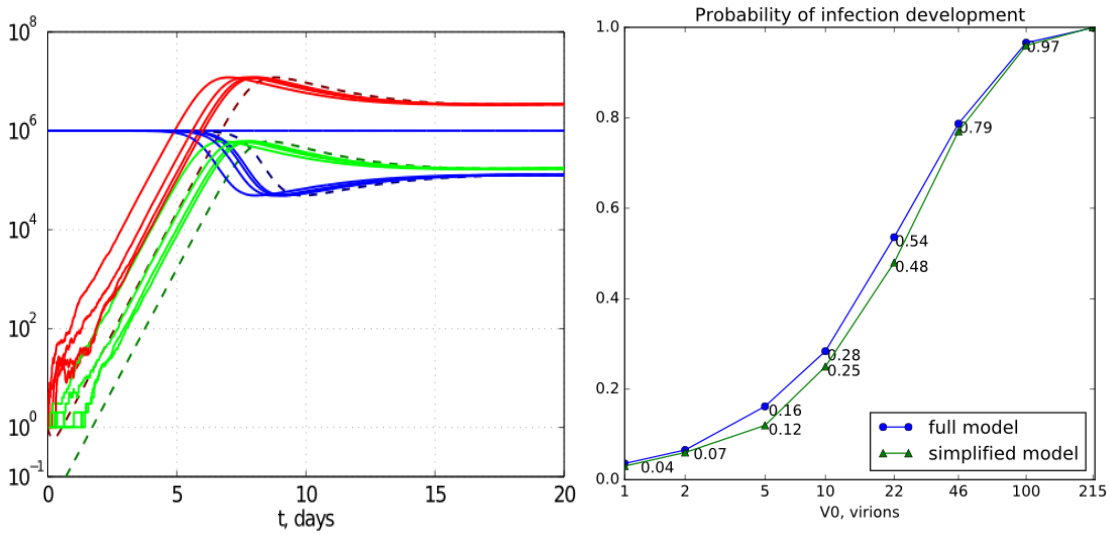

FIGURE 3. Left: Examples of non-degenerated realizations obtained by the MC direct numerical simulation (solid curves). Dashed lines indicate the deterministic solution of virus dynamics.
Right: Probabilities of infection development (non-degenerate cases) for different initial number of viruses V_0. Blue points represent the calculations obtained from full stochastic model (3.1) simulations, green points – simplified model (4.2) simulations.

The time-dependent sample means and sample STDs for ensemble of non-degenerate realizations are plotted, respectively, in Figures 4 and 5 for various initial number of viruses V_0. The Figure 4 shows the differences between the stochastic and deterministic solutions. One can see that the smaller number of the initial virions—the more is discrepancy between the deterministic and the mean stochastic dynamics, i.e., averaged over non-degenerate realizations.

The time dependence of the standard deviation (STD) is shown in Figure 5. Plots of STD time dependence have peculiar, two-hump shapes with a local minimum between the humps. The hump maxima and the STD in general are reduced in magnitude with the increasing number of initial virions. This confirms the importance to model the viral dynamics with the stochastic approach as in reality the initial number of virions is small down to a single virion [18, 37, 43].

Smoothed histograms for the number of viruses V at different stages of infection are shown in Figure 6. Stage 3 corresponds to the time of mean viral load peak $t_{\max} : V(t_{\max}) \to \max$, stage 2 and stage 4 are 4 days before and after $t_{\max}$, respectively. The histograms are normalized in such the way that they approximate PDF of the process: they should tend to the real PDF when the number of realizations tends to infinity.

In fact, the distribution of MC (3.1) is discrete and three-dimensional. We are dealing here with the smoothed PDF over wide enough intervals and also integrated over variables X and Y. One can see that for small initial number of virions, the PDFs are very wide, this fact indicates the necessity to consider random dynamics of infection in those cases. Also observe that PDF shapes are far from being Gaussian, but become narrower and close to Gaussian with the growth of initial virions for stages 2 and 4. As for the maximal infection load at stage 3, the PDF shapes have an steep decay after the maximum and do not tend to Gaussian with increase of V_0.

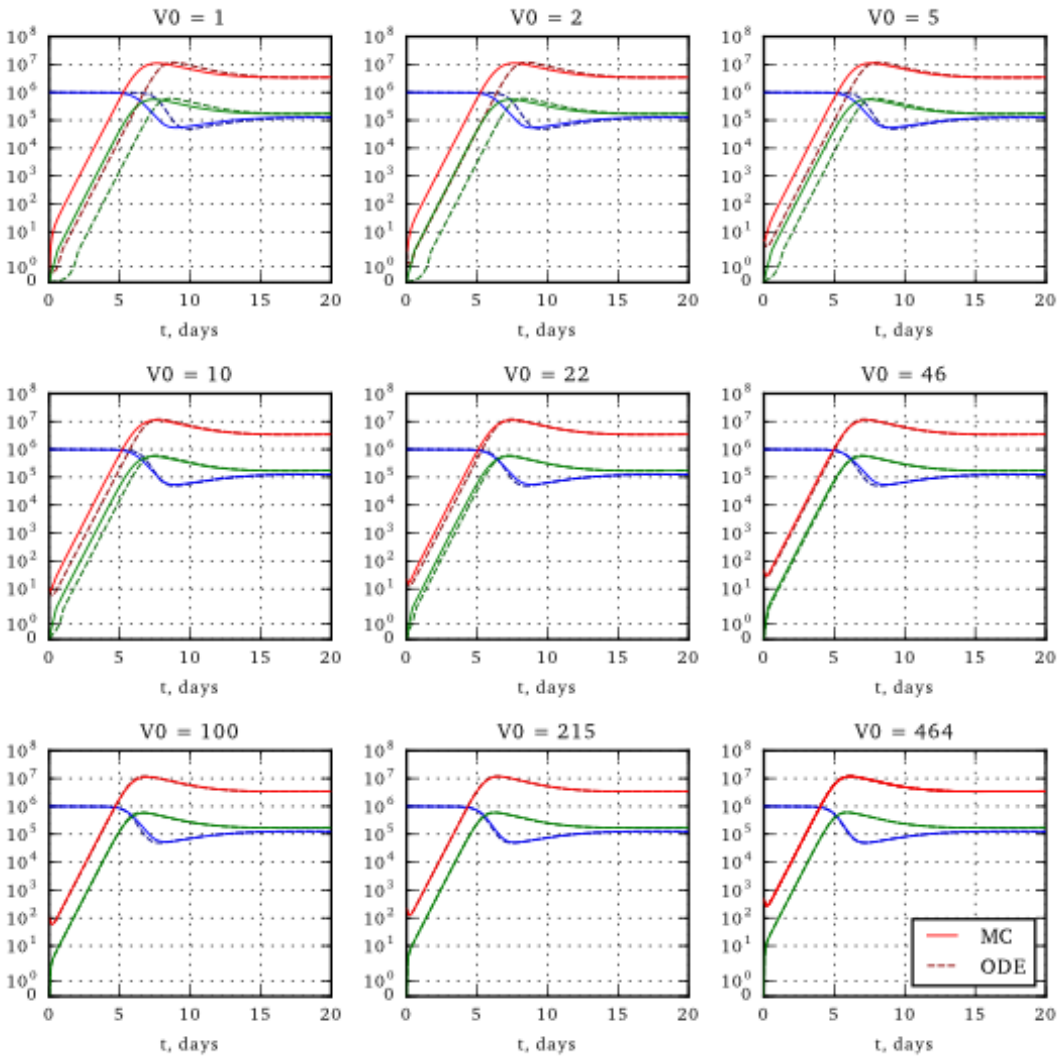

FIGURE 4. Means of 10k infection development realizations of the MC numerical simulations of virus (red), uninfected (blue) and infected (green) target cells are shown by solid curves. Results of the deterministic dynamics are plotted by dashed lines.

4. The hybrid stochastic model

Plots in Figure 3(left) show that the relatively large random fluctuations are apparent at the initial phase only (stage 1 and beginning of stage 2). In subsequent stages the curves are rather smooth, the numbers of all dynamic participants are large (several orders), and the fluid dynamics limit works rather well. The randomness in realizations at the latest stages is determined by the stochasticity at the early stages. This gives rise to an idea to split dynamics into two main stages or phases. These two phases in hybrid model should not be confused with four stages of the deterministic dynamics described in Section 2.2.

At the initial phase when the numbers of virions and infected cells are small and, hence, the stochasticity of interaction between species should be taken into account. However at this phase, the number of infected cells varies little and can be regarded as constant. This fact enables us to simplify the earlier phase model

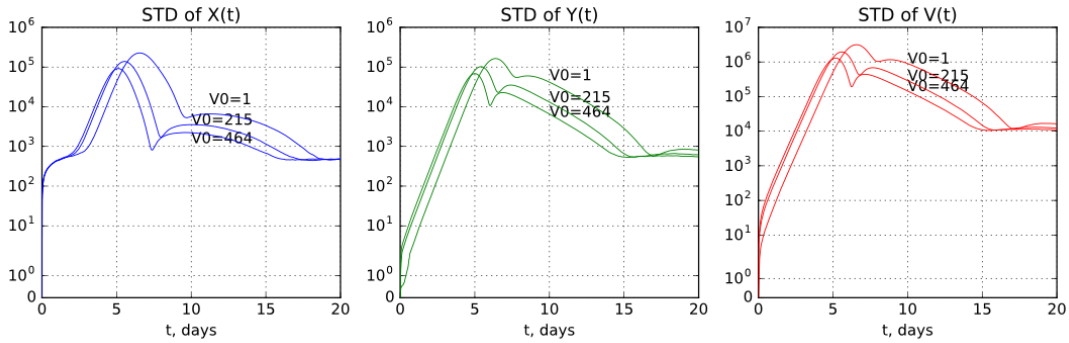

FIGURE 5. STDs of populations numbers for 10k infection development realizations of MC numerical simulations. Colors: blue – uninfected cells, green – infected cells, red – viruses.

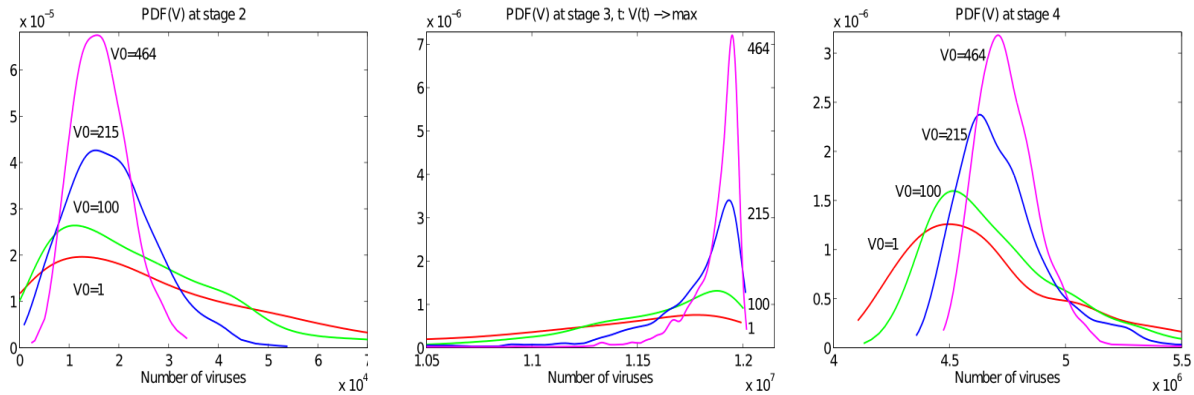

FIGURE 6. Smoothed normalized histograms for the number of virions at stage 2 (left), stage 3 (middle), stage 4 (right) for different initial number of viruses V_0. The histograms are calculated for ensemble of 10k infection development realizations.

by excluding the uninfected cells dynamics from the total process. The number of all species becomes large enough to switch the total process to a deterministic behaviour somewhere in the middle of the stage 2. This is a main idea of hybrid model in which the stochastic and deterministic phases are separated and can be matched at any instant $t = t_*$ taken from the interval in which two following conditions are satisfied simultaneously

$$X \approx x_0, \qquad Y, V \gg 1. \tag{4.1}$$

These conditions can be satisfied in the middle of the stage 2 provided the initial number of virions is much smaller than the number of virions taking place near the maximum infection load: $V_0 \ll V(t = t_{\max})$. The later condition is the condition of the Small Initial Contagion (SIC) approximation introduced in [33] for epidemic modelling. Note that the term hybrid model [31, 32] is referred to stochastic processes whereas the SIC approximation [33–35] can be applied to a deterministic process as well to simplify its computation and analysis (see Section 2.3).

The hybrid model has been developed first for epidemic modelling [31, 32, 34, 35] but one can see that it can be also applicable for the viral dynamics simulation. In this simulation, the stochastic phase can be modelled by Markov chain which is simplified compared with the full model (3.1). In the deterministic phase of the hybrid model, the deterministic equations (2.1) should be integrated many times with the

random initial conditions defined by the stochastic phase. In the framework of the SIC approximation, the once computed limiting solution can be used for all the realizations, e.g., [34].

Thus, in the framework of the hybrid model we are keeping the number of uninfected cells equal to $x_0 = \lambda/d$. Introducing parameter $\beta' = \beta x_0$ we can write down the following approximate MC which contains only four processes

Process #	Event	Rate	Description
1	$Y \to Y + 1$	$\beta' V$	Cell infection
2	$Y \to Y - 1$	aY	Infected cell death
3	$V \to V + 1$	kY	Virus particle birth
4	$V \to V - 1$	uV	Virus particle death

(4.2)

MC (4.2) has been implemented in the same way as MC (3.1). The computation of the stochastic process (4.2) is performed until the time t_* which was selected to be 4 days before the expected averaged peak of the infection process: $t_* = t_{\max} - 4$, i.e. in the middle of stage 2. At this time numbers of species reach the values $V \sim 10^4$, $Y \sim 10^3$, i.e. they are large enough for application of the deterministic model. After that time, the deterministic ODE (2.1) with the initial condition $v(t_*) = V(t_*)$, $y(t_*) = Y(t_*)$, $x(t_*) = x_0$ are integrated for the total time of realization which is 20 days. This is true for non-degenerate realizations. In some realizations which are degenerate, the values $V(t) = 0, Y(t) = 0$ can be reached before time t_* and the stochastic process is terminated.

The CPU time of a single realization in the framework of the hybrid model is about 5 s, i.e., 50 fold faster than the CPU time for the full model (see Section 3.4). This is mainly because the dynamics of uninfected cells is disregarded. The number of uninfected cells is very large: 10^7–10^8 and interactions between cells are very frequent, therefore an average time interval between the events is much shorter than in the hybrid model. At the same time relatively small fluctuations of X weakly effect the dynamics of V and Y.

The probability to the infection process to develop versus the number of initial virions computed in by the hybrid model is shown in Figure 3(right). Observe a very small discrepancy between the values obtained by the full and hybrid models.

In Figure 7, the computed smoothed normalized histograms are presented for the virion numbers in stage 2 (left) and stage 4 (right), i.e., before and after the infection peak. The curves with the darker color indicate the histograms for the full model, curves with the same but lighter color indicate the histograms for the hybrid model. The plots show that the discrepancy of the histograms is reasonably small except in the case $V_0 = 1$. This case needs explanation and further study.

5. Conclusion

The dynamics of infectious diseases in humans is subject to many random effects resulting from stochastic fluctuations in biochemical and cellular processes underlying the infection of target cells, within-the-cell replication and effect of the immune responses. The potential of the stochastic models has been appreciated in recent studies based on single scale models formulated with SDEs [21], discrete stochastic models [24, 27], genetic algorithms [4] as well as multi-scale hybrid models [6, 7, 23, 30].

The kinetics of the virus and immune responses is characterized by a large variation in scale, i.e., from few species to millions of virions and cells, respectively. In order to correctly model the corresponding evolution, computational algorithms which take into account the transition of the system from noise-dominated behaviour to a deterministic dynamics need to be developed. In our study we implemented a Markov chain approach to generalize a deterministic set of equations, representing a basic model of HIV infection, to a stochastic discrete populations model. The convergence in distribution of the stochastic model realizations to the solution of the deterministic model has been proved. We explored the computational aspects of the stochastic behaviour of virus infection spread across the population of target

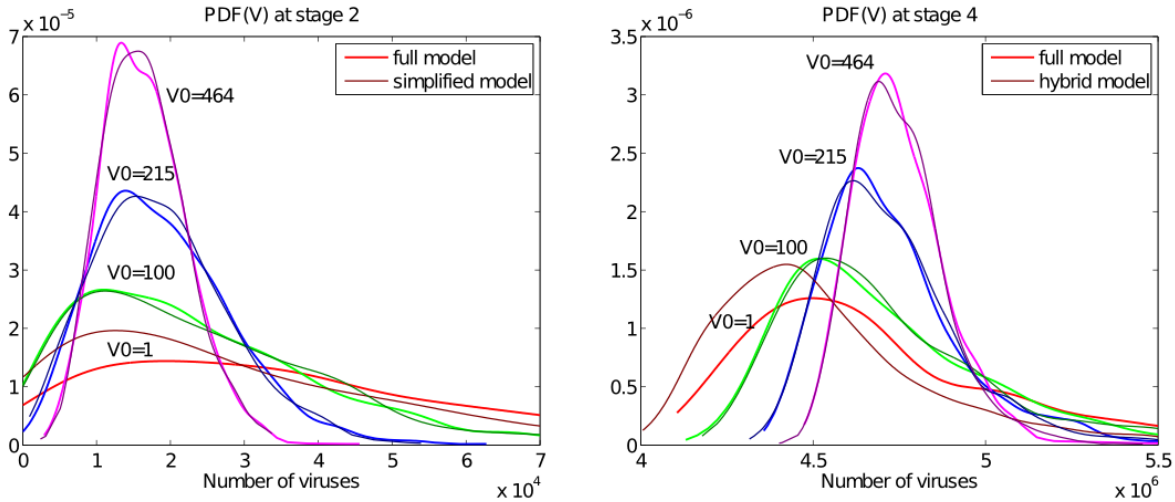

FIGURE 7. Smoothed normalized histograms of number of virions at stage 2 (left) and stage 4 (right) for different initial number of viruses V_0, calculated using full (3.1) and simplified (4.2) stochastic models. The histograms are calculated for ensemble of 10k infection development realizations. The curves with the darker color indicate the histograms for the full model, curves with the same but lighter color indicate the histograms for the hybrid model.

cells under the well-mixing assumption. To deal efficiently with the description of population dynamics in state space domain characterized by a mix of small and large species densities, we have developed a hybrid modelling method and identified the criteria for switching between stochastic to semi-deterministic and finally to deterministic types of modelling.

It has been shown that random effects result in a time-dependent distribution of the virus population size that can be described by a PDF. The PDF becomes more narrow with an increase in the dose of infection reflecting the transition to a deterministic mode of behaviour. The model under study has been used to quantify the dependence of the probability of the systemic infection development versus extinction of the virus. HIV infection is considered to result from the transmission of a single viral variant [18]. Understanding of the stochastic and selective forces that restrict transmission and may be targets for prevention strategies, requires further development of stochastic models that account for multiple bottlenecks, scales and compartments in order to quantify the probability of fateful decision between extinction and systemic spread of the virus upon transmission.

In are future work we are going to investigate stochastic models with delay between the cell contamination and free virion release [15]. Then we will have to model the dynamics by Markov processes with continuous state spaces instead of Markov chains with countable state spaces.

Acknowledgements. This study was supported by the Russian Science Foundation Grant number 15-11-00029 (to D.G. and G.B.). M.K. acknowledges a support within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

Compliance with ethical standards

Conflict of interests The authors declare that they have no conflict of interests.

References

[1] Alizon, S., Magnus C., 2012, "Modelling the course of an HIV infection: Insights from ecology and evolution," Viruses 4, pp. 1984–2013.

[2] Banks, H.T., Davidian, M., Hu, S., Kepler, G.M., Rosenberg, E.S., 2008, "Modelling HIV immune response and validation with clinical data," J. Biol. Dyn. 2, pp. 357–385.

[3] Bocharov, G., Chereshnev, V., Gainova, I., Bazhan, S., Bachmetyev, B., Argilaguet, J., Martinez, J., Meyerhans, A., 2012, "Human Immunodeficiency Virus Infection: From Biological Observations to Mechanistic Mathematical Modelling," Math. Model. Nat. Phenom. 7, pp. 78–104.

[4] Bocharov, G., Telatnikov, I., Chereshnev, V., Martinez, J., Meyerhans, M., 2015, "Mathematical modelling of the within-host HIV quasispecies dynamics in response to antiviral treatment," Russian Journal of Numerical Analysis and Mathematical Modelling. 30(3), pp 157–170.

[5] Boudec, J.Y.L., McDonald, D., Mundinger, J., 2007, "A generic mean field convergence result for systems of interacting objects," in Fourth International Conference on the Quantitative Evaluation of Systems (QEST 2007) p. 3. doi:10.1109/QEST.2007.8.

[6] Bouchnita, A., Bocharov, G., Meyerhans, A., Volpert, V., 2017, "Hybrid approach to model the spatial regulation of T cell responses," BMC Immunology, (in press).

[7] Bouchnita, A., Bocharov, G., Meyerhans, A., Volpert, V., 2017, "Towards a multiscale model of acute HIV infection," Computation, 5(1), p. 6; doi:10.3390/computation5010006

[8] Canini, L., Perelson, A.S., 2014, "Viral kinetic modeling: state of the art," J Pharmacokinet Pharmacodyn. 41(5), pp. 431–43.

[9] Castro, M., Lythe, G., Molina-ParÃs, C., Ribeiro, R.M., 2016, "Mathematics in modern immunology," Interface Focus, 6(2), 20150093

[10] Darling, R.W.R., Norris, J.R., 2008, "Differential equation approximations for Markov chains", Probability Surveys, 5, pp. 37–79, DOI: 10.1214/07-PS121

[11] Eftimie, R., Gillard, J.J., Cantrell, D.A., 2016, "Mathematical Models for Immunology: Current State of the Art and Future Research Directions," Bull. Math. Biol. 78(10), pp 2091-2134.

[12] S.N. Eithier, T.G. Kurtz, 1986, Markov Processes. Characterization and Convergence, Wiley.

[13] Gillespie, D.T., 1976, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", Journal of Computational Physics, 22(4), pp. 403–434.

[14] Grossman, Z., Meier-Schellersheim, M., Paul, W.E., Picker, L.J., 2006, "Pathogenesis of HIV infection: What the virus spares is as important as what it destroys," Nat. Med. 12, pp 289–295.

[15] Hattaf, K., Yousfi, N, 2016, "A class of delayed viral infection models with general incidence rate and adaptive immune response", Int. J. Dynam. Control 4, pp 254–265.

[16] Hattaf, K., Yousfi, N, 2016, "A generalized virus dynamics model with cell-to-cell transmission and cure rate", Advances in Difference Equations 174, DOI 10.1186/s13662-016-0906-3.

[17] Huang, G., Ma, W., Takeuchi, Y., 2009, "Global properties for virus dynamics model with Beddington–DeAngelis functional response," Applied Mathematics Letters, 22, pp. 1690–1693; doi:10.1016/j.aml.2009.06.004

[18] Joseph, S.B., Swanstrom, R., Kashuba, A.D., Cohen, M.S., 2015, "Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. Nat Rev Microbiol," 13(7), pp. 414–425.

[19] Korobeinikov, A., 2004, "Global properties of basic virus dynamics models," Bulletin of Mathematical Biology, 66, pp. 879–883; doi:10.1016/j.bulm.2004.02.001

[20] De Leenheer, P. and Smith, H.L., 2003, "Virus dynamics: a global analysis," SIAM J. Appl. Math., 63(4), pp. 1313–1327;

[21] Luzyanina, T., Bocharov, G., 2014, "Stochastic modelling of the impact of random forcing on persistent hepatitis B virus infection," Mathematics and Computers in Simulation. 96, pp. 54-65.

[22] Liu, J., Keele, B.F., Li, H., Keating, S., Norris, P.J., Carville, A., Mansfield, K.G., Tomaras, G.D., Haynes, B.F., Kolodkin-Gal, D., Letvin, N.L., Hahn, B.H., Shaw, G.M., Barouch, D.H., 2010, "Low-dose mucosal simian immunodeficiency virus infection restricts early replication kinetics and transmitted virus variants in rhesus monkeys," J. Virology, 84(19), pp. 10406–10412.

[23] Marino, S., Kirschner, D.E., 2016, "A multi-compartment hybrid computational model predicts key roles for dendritic cells in tuberculosis infection," Computation, 4, 39; doi:10.3390/computation4040039

[24] Noecker, C., Schaefer, K., Zaccheo, K., Yang, Y., Day, J., Ganusov, V.V., 2015, "Simple mathematical models do not accurately predict early SIV dynamics," Viruses 7, pp. 1189–1217; doi:10.3390/v7031189

[25] Nowak, M.A., May, R.M., 2000, Virus Dynamics. Mathematical Principles of Immunology and Virology, 250 pages. Chapter 3.

[26] O'Neill, M.E., 2015, "PCG: A family of simple fast space-efficient statistically good algorithms for random number generation," ACM Trans. Math. Softw. (submitted), 1-46.

[27] Pearson, J.E., Krapivsky, P., Perelson, A.S., 2011, "Stochastic theory of early viral infection: continuous versus burst production of virions," PLoS Comput Biol. 7(2), e1001058.

[28] Perelson, A.S., Nelson, P.W., 1999, "Mathematical analysis of HIV-1 dynamics in vivo," SIAM Rev., 41, pp. 3–44.

[29] Perelson, A.S., Ribeiro, R.M., 2013, "Modeling the within-host dynamics of HIV infection," BMC Biol., 11:96; DOI: 10.1186/1741-7007-11-96

[30] Prokopiou, S.A., Barbarroux, L., Bernard, S., Mafille, J., Leverrier, Y., Arpin, C., Marvel, J., Gandrillon, O., Crauste, F., 2014, "Multiscale Modeling of the Early CD8 T-Cell Immune Response in Lymph Nodes: An Integrative Study," Computation 2, pp. 159–181.

[31] Rebuli, N.P., Bean, N.G., Ross, J.V., 2016, "Hybrid Markov Chain models of S-I-R disease dynamics," Journal of Mathematical Biology, DOI: 10.1007/s00285-016-1085-2
[32] C. Safta, K. Sargsyan, B. Debusschere, H.N. Najm (2015) Hybrid discrete/continuum algorithms for stochastic reaction networks. *Journal of Computational Physics*, 281 (2015) 177–198
[33] Sazonov, I., Kelbert, M., Gravenor, M.B., 2008, "The speed of epidemic waves in a one-dimensional lattice of SIR models," Mathematical Modelling of Natural Phenomena, 3(4), pp. 28–47.
[34] Sazonov, I., Kelbert, M., Gravenor, M.B., 2011, "A two-stage model for the SIR outbreak: Accounting for the discrete and stochastic nature of the epidemic at the initial contamination stage," Mathematical Biosciences, 234, pp. 108–117.
[35] Sazonov, I., Kelbert, M., Gravenor, M.B., 2016, "Random migration processes between two stochastic epidemic centers," Mathematical Biosciences, 274, pp. 45–57.
[36] Schinazi, R.B., 2005, "A spatial stochastic model for virus dynamics," Journal of Statistical Physics, 128(3), pp. 771–779
[37] Shaw, G.M., Hunter, E., 2012, "HIV transmission," Cold Spring Harb Perspect Med. 2(11); doi:10.1101/cshperspect.a006965..
[38] Tange, O., 2011, "GNU Parallel - The Command-Line Power Tool," The USENIX Magazine, pp. 42–47.
[39] Tuckwella, H.C., Toubianab, L., 2007, "Dynamical modeling of viral spread in spatially distributed populations: stochastic origins of oscillations and density dependence," BioSystems, 90, pp. 546–559
[40] Van Kampen, N.G., 2007, "Stochastic Processes in Physics and Chemistry" (Chapter IX—The Langevin Approach, Pages 219-243) Elsevier, Amsterdam, Third edition: 2007, 463 pages, ISBN: 978-0-444-52965-7
[41] Xu, Y., Allena, L.J.S., Perelson A.S., 2007, "Stochastic model of an influenza epidemic with drug resistance," J. Theoretical Biology, 248, pp. 179–193
[42] Xu, Sh., Lu, W., Zhan, Zh., 2012, "A stochastic model of multivirus dynamics," IEEE Transactions on Dependable and Secure Computing, 9(1), pp. 30–45.
[43] Yuan, Z., Ma, F., Demers, A.J., Wang, D., Xu, J., Lewis, M.G., Li, Q., 2017, "Characterization of founder viruses in very early SIV rectal transmission," Virology. 502, pp. 97–105.

A. Proof of Theorem 3.1

The proof of Theorem 3.1 follows closely that of Theorem 2.1, Chapter 11 in [12]. Consider a Markov process $\tilde{\mathbf{X}}_n(t) \in \mathbb{Z}^3$ with the transition rates $q^{(n)}_{m,m+m'} = n\nu_{m'}\big(\frac{1}{n}m\big)$. Here vectors $m, m' \in \mathbb{Z}^3$ take six possible values corresponding the columns of matrix (3.14). Functions $\nu_{m'}$ are defined in Table (3.1). Then process $\tilde{\mathbf{X}}_n(t)$ admits the following representation

$$\tilde{\mathbf{X}}_n(t) = \tilde{\mathbf{X}}_n(0) + \sum_{m'} m' \,\mathrm{Po}_{m'}\Big(n\int_0^t \frac{1}{n}\tilde{\mathbf{X}}_n(s)\,\mathrm{d}s\Big) \tag{A.1}$$

where $\mathrm{Po}_{m'}$ are six independent Poisson processes of rate 1. Setting $F(m) = \sum_{m'} m'\nu_{m'}(m)$ we obtain for the process $\mathbf{X}_n(t) = \frac{1}{n}\tilde{\mathbf{X}}_n(t)$ the following representation

$$\mathbf{X}_n(t) = \mathbf{X}_n(0) + \frac{1}{n}\sum_{m'} m' \,\widetilde{\mathrm{Po}}_{m'}\Big(n\int_0^t \nu_{m'}(\mathbf{X}_n(s))\,\mathrm{d}s\Big) + \int_0^t F(\mathbf{X}_n(s))\,\mathrm{d}s \tag{A.2}$$

where $\widetilde{\mathrm{Po}}_{m'}(t) = \mathrm{Po}_{m'}(t) - t$ is a compensated Poisson process. Using the Euclidean norm $|.|$ define

$$\epsilon_n(t) = \sup_{u\le t}\Big|\mathbf{X}_m(u) - \mathbf{X}_n(0) - \int_0^u F\big(\mathbf{X}_n(s)\big)\,\mathrm{d}s\Big|. \tag{A.3}$$

Now write ODEs (2.1) in the integral from

$$\mathbf{x}(t) = \mathbf{x}(0) + \int_0^t F(\mathbf{x}(s))\,\mathrm{d}s$$

and assume without loss of generality the Lipshitz condition in a neighbourhood of $\{\mathbf{x}(s), s \le t\}$ for a fixed $t \ge 0$: $|F(m_1) - F(m_2)| \le C|m_1 - m_2|$. Then (A.2) implies

$$|\mathbf{X}_n(t) - \mathbf{x}(t)| \le |\mathbf{X}_n(0) - \mathbf{x}(0)| + \epsilon_n(t) + C\int_0^t |\mathbf{X}_n(s) - \mathbf{x}(s)|\,\mathrm{d}s. \tag{A.4}$$

Hence by Gronwall's inequality

$$|\mathbf{X}_n(t) - \mathbf{x}(t)| \le |\mathbf{X}_n(0) - \mathbf{x}(0)| + \epsilon_n(t)e^{Ct}. \tag{A.5}$$

It remains to note that $|\epsilon_n(t)| \to 0$ in probability by the Law of Large Numbers. □