EXISTENCE AND DYNAMICS OF STRAINS IN A NONLOCAL REACTION-DIFFUSION MODEL OF VIRAL EVOLUTION*

NIKOLAI BESSONOV†, GENNADY BOCHAROV‡, ANDREAS MEYERHANS§, VLADIMIR POPOV¶, AND VITALY VOLPERT ‖

Abstract. In this work, we develop a mathematical framework for predicting and quantifying virus diversity evolution during infection of a host organism. It is specified as a virus density distribution with respect to genotype and time governed by a reaction-diffusion integro-differential equation taking virus mutations, replication, and elimination by immune cells and medical treatment into account. Conditions for the existence of virus strains that correspond to localized density distributions in the space of genotypes are determined. It is shown that common viral evolutionary traits like diversification and extinction are driven by non-local interactions via immune responses, target-cell competition and therapy. This provides with a mechanistic explanation for clinically relevant properties like immune escape and drug resistance selection, and allows to link virus genotypes to phenotypes.

Key words. virus density distribution, genotype, virus infection, immune response, resistance to treatment, nonlocal interaction

AMS subject classifications. 35K57

1. Introduction.

Virus infections of host organisms are initiated when viruses overcome protective body boundaries like skin and mucosal surfaces. The first step then is local virus amplification in target cells at the entry site, after which viruses can subsequently spread to other tissues or organs. Due to the explosive amplification capacity of viruses, e.g., a CD4 T cell infected by a single HIV particle can produce around 1000 progeny particles within 24 hours [32], the host requires a multitude of restricting defense mechanisms both at the cellular level (virus restriction factors) as well as at the organism level (immune system) ([7], chapter 1). These systems together with error-prone virus replication and therapeutic interventions drive the evolution of viruses within genotype space.

Genetic evolution of viruses can play a fundamental role in virus pathogenesis [50, 51, 12, 29, 14, 47]. However, the prediction of virus evolution remains a challenge for theoretical and empirical research [22, 46, 41, 27]. Today, the general approach to analyse virus evolution relies on the notions of fitness and fitness landscapes which map the genotypes to reproductive success [52]. While these are essential for predicting the selection of mutants, a quantitative a priori description of the fitness landscape remains difficult as it is determined by multiple factors and constraints [40]. A combination of experimental and mathematical modelling studies provided interesting

*Submitted to the editors DATE.

†Institute of Problems of Mechanical Engineering, Russian Academy of Sciences, 199178 Saint Petersburg, Russia (nickbessonov1@gmail.com).

‡Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow, Russian Federation; Moscow Center for Fundamental and Applied Mathematics at INM RAS; Sechenov First Moscow State Medical University, Moscow 119991, Russian Federation (g.bocharov@inm.ras.ru).

§ ICREA, Pg. Lluis Companys 23, 08010 Barcelona, Spain; Infection Biology Laboratory, Universitat Pompeu Fabra, 08003 Barcelona, Spain (andreas.meyerhans@upf.edu).

¶Peoples Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation (volodimir.a@gmail.com).

‖Institut Camille Jordan, UMR 5208 CNRS, University Lyon 1, 69622 Villeurbanne, France; INRIA Team Dracula, INRIA Lyon La Doua, 69603 Villeurbanne, France; Peoples Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation (volpert@math.univ-lyon1.fr).

means to specify fitness landscapes [40, 17, 57]. They consider as a structure variable for describing virus diversity (i) the binding affinity of viral antigens to T cells [46], (ii) single nucleotide variants of the HIV-1 genome [57], (iii) antibody binding affinity and (iv) RNA virus capsid folding stability [40]. The mathematical models were formulated with systems of integro-differential equations, stochastic ODEs and algebraic equations, respectively.

Since the pioneering works by Fisher [15], Kolmogorov, Petrovskii and Piskunov [25], and later Kimura [23], the reaction-diffusion equation and the diffusion-convection equation (backward Kolmogorov equations) were used to describe the mutation-driven evolution of gene frequencies in a population. The models were then tailored for the analysis of intra-host genetic evolution of viruses subjected to immune responses using a reaction-diffusion equation with a non-local competition growth term for the virus dynamics and an ODE for immune cells [44], or a system of reaction-convection equations for virus dynamics and immune cells [22].

The concept of quasispecies (clouds of related strains) represents a fruitful framework for the analysis of evolving virus populations [6, 13]. Mathematically, it can be expressed using either a discrete or a continuous description of the state space. Several models were formulated with high-dimensional ODE systems that describe the evolution of individual mutants [36]. The complementary approach which we follow here in our work is based on considering a density distribution function to represent the abundance of viruses in some defined genotype space $u(t, x), x \in \mathbf{X}$. It provides a less developed but powerful tool to study a range of fundamental issues concerning the dynamics of virus infections including (i) the presence of wild type virus versus mutants, (ii) the branching of genotypes, (iii) mutant extinction, and (iv) escape mechanisms. The following processes were taken into account: (1) random generation of mutations of viral genomes, (2) diversity of target cells in which viral replication takes place, (3) antigen cross-reactivity of immune cells, (4) bell-shaped responsiveness of antigen-specific immune cells, (5) competition between quasispecies, and (6) impact of antiviral drugs.

All six processes are described by formulating a general reaction-diffusion model with time delay and a number of non-local terms. The model is an abstraction from the details of genome sequences and virion structures. It enables the study of a broad range of dynamical behaviours of viral quasispecies in the respective genotype space including travelling waves, standing waves, mono- and bistable waves, irregular dynamics, pulses and branching waves. Whereas the travelling waves [42, 43, 41] and branching waves [46] have received some attention previously, the other dynamical modes remain to be systematically investigated. We elaborate the conditions underlying various evolutionary behaviours and relate them to the within-host viral evolution subjected to antiviral immune responses and antiviral therapy.

2. Mathematical model of adaptive evolution. Similar to biological species, virus strains can be characterized by their genotypes. Consider the 1D space of genotypes $\mathbf{X}$ in which the viral quasispecies population can be described as a density function of genotypes $x \in \mathbf{X}$, i.e. $u(t, x)$ at time $t \in [0, \infty)$. The population density distribution $u(x, t)$ represents a strain if it is localized in the sense that it has a maximum at some $x = x_0$, and it rapidly decays as $|x - x_0|$ grows. A typical dependence is given by the normal distribution described by the diffusion-advection equation (Supplementary materials, SM). Note that localized density distributions do not exist as stationary solutions of the diffusion equation. In this and in the next section, we will introduce the model and will investigate the existence and dynamics of virus strains.

2.1. Governing equation for the within-host quasispecies dynamics. We formulate a general mathematical model to investigate the evolutionary dynamics of the virus density distribution $u(t,x)$ depending on its genotype, being considered as a continuous variable, and time. As shown in Figure 1, the model takes into account mutations, quasispecies competition for target cells, immune-mediated elimination subject to genotype density-dependent regulation and antiviral treatment. The model is formulated in a form of reaction-diffusion equation with time lags and non-local regulation:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ruH - duC - \sigma(x)u. \tag{2.1}$$

This equation is considered on the interval $0 \le x \le L$ with some initial and boundary conditions. For the mathematical convenience, we consider this equation on the whole axis. In numerical simulations, it will be considered on a bounded interval with periodic boundary conditions, so that it can be extended on the whole axis by periodicity. In this case, the problem is defined for all real values of x. In the biological interpretation, we will keep in mind a bounded interval of genotypes. Diffusion coefficient D is a positive parameter related to the mutation rate. Derivation details of the diffusion equation for describing the mutation-driven evolution of viruses are presented in SM. Parameters r and d are positive.

The virus reproduction term $ruH = ru(1 - qJ(u))$ is proportional to the virus density u and to the concentration of available host cells H. In the model, virus replication in target cells is described by a conventional logistic term. However, it takes into account the genotype-dependent target cell tropism via the non-local interaction term $J(u)$. Three biologically relevant functional forms can be specified:

1. When there is a one-to-one correspondence between the virus genotype x and the type of infected cells, then $J(u) = u(x,t)/H_0$. Here H_0 is a carrying capacity and $u(x,t)/H_0$ is the proportion of infected host cells.
2. The above case is biologically restrictive since, in general, viruses with different genotypes can infect the same target cells. In this case, the logistic term is replaced by the term $ru(1-qJ(u))$, where $J(u) = \int_{-\infty}^{\infty} \phi(x-y)u(y,t)dy$, q is a positive parameter, and the kernel $\phi(x-y)$ shows how the infection of host cells depends on the distance between the genotypes. Such terms with non-local consumption of resources were previously considered in the framework of a nonlocal Fisher-KPP equation, in particular, for addressing evolution problems [10, 19, 53].
3. If the infection of target cells does not depend on the virus genotype, that is $\phi(x-y) \equiv$ const, then the reproduction term becomes $ru(1-qU(t))$, where $U(t) = \int_{-\infty}^{\infty} u(y,t)dy$. This is the case of global consumption of resources in population dynamics [54].

The third term of eqn (2.1) describes the immune-mediated elimination of viruses. The elimination term duC is proportional to the virus density and to the concentration of immune cells C. Since production of virus-specific immune cells by the organism is stimulated by viral antigens, their abundance can be considered as a function of virus concentration, $C(x,t) = f(u_\tau)$, $u_\tau = u(x,t-\tau)$. This approximation can be obtained from a more complete model presented by a system of two equations [9]. Time delay τ takes into account the duration of clonal expansion of immune cells stimulated by the antigen. The function $f(u)$ is growing for small u and decreasing for large u since virus can downregulate cell proliferation and upregulate their death via a number of mechanisms, including functional exhaustion, activation-induced apoptosis, etc.

To take into account the cross-reactivity of immune cells and their specificity-dependent distribution with respect to the virus-elimination efficacy, the following non-local convolution terms are specified:

- The simplest case is a one-to-one correspondence between virus genotype (antigen) and immune cells (lymphocytes). This assumption reflects the basic antigen-specific stimulation and is described by the function $d \cdot u \cdot f(u_\tau)$.
- However, because of antigen cross-reactivity and bystander stimulation effects, a specific lymphocyte can be activated by the cytokines released due to antigens of another specificity. Thus, the genotype dependence will not be strictly localized. We assume that virus genotypes of some range $[x_l, x_h]$ lead to clonal expansion of a certain repertoire of lymphocyte specificities. Then, instead of the function $f(u_\tau)$, we consider a more general expression $f(S(u_\tau))$, where $S(u_\tau)(x,t) = \int_{-\infty}^{\infty} \psi(x-y)u(y,t-\tau)dy$. The function $\psi(x)$ shows how the initiation of clonal expansion of immune cells depends on the virus genotype.
- Immune cells can also eliminate viruses in some range of genotypes although with a different efficacy. In this case, the virus elimination term takes the form $u(x,t)\int_{-\infty}^{\infty} \theta(x-y)f(S(u_\tau)(y,t))dy$. The dependence of virus elimination on immune cells is determined by the function $\theta(x)$. In this work we will consider a particular case of this equation where $\theta(x)$ is approximated by a δ-function.

Without loss of generality we can set $\int_{-\infty}^{\infty} \phi(x)dx = \int_{-\infty}^{\infty} \psi(x)dx = 1$.

Finally, the last term of eqn (2.1) describes genotype-dependent virus mortality or inhibition of its replication. Virus death rate is determined by its degradation in the infected host which can occur via multiple processes, including the intracellular ones, such as virus internalization and cytosolic degradation, the degradation of the viral RNA by cellular defense system, and stability to various physical and chemical factors in the extracellular environment. In is known that virus strains differ in their pH sensitivity [45], thermal and mechanical stability of viral particle with its structure defined by the genome [31, 11]. Next, drug resistance is a fundamental property of many virus infections [1, 26, 36]. It is associated with a partial or complete reduction of the inhibitory effect of the drugs on virus replication. For some drugs, a single-point mutation confers high level of resistance whereas for other single drugs or combination of drugs acquisition of several mutations is needed to reduce the inhibition to zero [36]. Overall, the genome-dependent differences in the virus death rate and inhibition are taken into account in the model phenomenologically by specifying the properties of $\sigma(x)$.

Taking into account all these assumptions, we arrive at the following model to study the intrapatient evolution of viral quasi-species driven by target-cell competition, immune forcing and antiviral therapy:

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru\left(1 - q\int_{-\infty}^{\infty} \phi(x-y)u(y,t)dy\right) - u\int_{-\infty}^{\infty} \theta(x-y)f(S(u_\tau)(y,t))dy - \sigma(x)u, \tag{2.2}$$

where

$$S(u_\tau)(y,t) = \int_{-\infty}^{\infty} \psi(y-z)u(z,t-\tau)dz.$$

The phenomenological features of the nested mathematical models considered in this study and the represented biological processes are summarized in Figure 1. Equation (2.2) has a very rich dynamics. We use it to examine the role of nonlocal nonlinear interactions for the emergence of various modes of virus evolutionary dynamics subject to immune- and drug-mediated forcing within a host, i.e. the evolutionary paths characterized by dominance of wild type viruses versus mutants, branching of genotypes, extinction, and escape kinetics. The mathematical conditions of their existence are elaborated below.

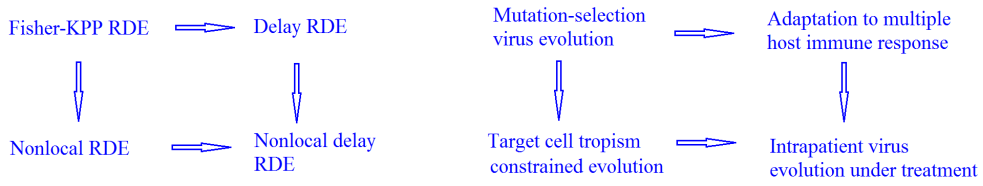

FIG. 1. *Nested mathematical models of different nature and complexity used in the study. Left: Fisher-KPP reaction-diffusion equation extended by considering non-local interactions and time-delay of the immune response. Right: Processes considered to determine the within-host mutation-selection driven viral evolution.*

Since equation (2.2) manifests extremely complex dynamics with many different patterns, in this work we will study a particular case where (a) the kernels $\psi(x)$ and $\theta(x)$ are the δ-functions such that the corresponding terms become local, (b) there is no time delay, $\tau = 0$, and we will focus our attention on the existence and dynamics of pulse solutions. In the following article, we will consider equation (2.2) under different assumptions more specific for the analysis of periodic patterns and waves.

3. Existence of virus strains. We will begin the analysis of the equation (see the assumptions (a), (b) above)

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru(1 - qU(t)) - uf(u) - \sigma(x)u \tag{3.1}$$

$U(t) = \int_{-\infty}^{\infty} u(x,t)dx$, with the investigation of its positive stationary solutions decaying at infinity called pulse solutions. In the context of virus distribution in the space of genotypes such solutions characterize virus strains. In this section, we will prove the existence of such solutions in the case of the interaction of virus reproduction with immune response or with the genotype dependent mortality.

3.1. Competition for target cells and immune response. Consider equation (3.1) on the whole axis assuming that $\sigma(x) \equiv 0$ and $f(u) \geq 0$ for $u \geq 0$. For simplicity of presentation, we set $D = r = q = 1$. We look for a positive stationary solution of this equation $u(x)$ with the limits $u(\pm\infty) = 0$ at infinity. Such solution is called a pulse solution. It satisfies the problem

$$u'' + u(1 - U) - uf(u) = 0, \quad u(\pm\infty) = 0, \tag{3.2}$$

where $U = \int_{-\infty}^{\infty} u(x)dx$. Set $k = 1 - U$. Then problem (3.2) writes:

$$u'' + u(k - f(u)) = 0, \quad u(\pm\infty) = 0. \tag{3.3}$$

Existence of solution of this problem can be studied analytically. Suppose that such solution exists and denote it by $u_k(x)$, where the subscript k shows its dependence on the parameter k. From the definition of k we obtain the following equation with respect to k:

$$1 - \int_{-\infty}^{\infty} u_k(x)dx = k. \tag{3.4}$$

Existence of its solution determines the existence of solution of problem (3.2). In order to get an analytical solution of this problem, let us consider a particular example $f(u) = b - u$. Then equation (3.3) becomes as follows: $u'' - pu + u^2 = 0$, where $p = b - k$. Set $u(x) = pw(\sqrt{p}x)$. Then $w(y)$ satisfies the equation $w'' - w + w^2 = 0$. It has a positive solution $w_0(y)$ such that $w_0(\pm\infty) = 0$. Hence, $u_k(x) = pw_0(\sqrt{p}x)$, and from equation (3.4) we obtain:

$$U_0\sqrt{b-k} = 1 - k, \tag{3.5}$$

where $U_0 = \int_{-\infty}^{\infty} w_0(y)dy$.

Theorem 3.1. *Consider a stationary solution of equation (3.1) decaying at infinity, that is, a solution of the problem (3.2) where $U = \int_{-\infty}^{\infty} u(x)dx$, $f(u) = b - u$, $b > 0$. Then there exist such positive values b_1, b_2, $b_1 < b_2$ that this problem has a positive solution for $b_1 < b < b_2$, and it does not have positive solution for $0 < b < b_1$ and $b > b_2$.*

The proof of the proposition follows from the analysis of equation (3.5). The method presented here can be generalized for the functions $f(u) = b - u^n$, $n > 1$.

Existence of a positive stationary virus density distribution provided by this theorem corresponds, from the biological point of view, to the existence of a virus strain. Existence of pulses for other functions $f(u)$ and their stability are confirmed by numerical simulations (Section 4). Thus, the mechanism leading to the existence of stable virus strains is determined by the interaction of the nonlocal term in the reproduction rate with the immune response. Such solutions do not exist without an immune response, that is, for $f(u) \equiv 0$. Decreasing functions $f(u)$ correspond to the decaying branch of the immune response where high virus concentration eliminates the immune cells. Conditions on the parameter b in Theorem 3.1 signify that virus strain persists for some intermediate intensity of immune response.

Let us note that if $\sigma(x) \equiv 0$, then the solution is invariant with respect to translation in space. From the biological point of view, this means that for the genotype-independent virus death, virus quasi-species can form around any value of average genotype determined by the initial condition (initial viral load).

3.2. Genotype-dependent mortality. In this section we study the existence of pulses (virus strains) in the case of genotype dependent mortality determined either by virus viability for different genotypes or by an antiviral treatment.

3.2.1. Piece-wise constant mortality rate. Consider the problem

$$u'' + u(1-U) - \sigma(x)u = 0\,, \quad u(\pm\infty) = 0 \tag{3.6}$$

on the whole axis, where $U = \int_{-\infty}^{\infty} u(x)dx$, $\sigma(x) = \sigma_0 > 1$ for $|x| \geq x_0$, and $\sigma(x) = 0$ for $|x| < x_0$, x_0 is some positive number. We look for a positive bounded solution of this equation. Clearly, it can exist only if $I(u) < 1$. Set

$$1 - U = k^2. \tag{3.7}$$

Then equation (3.6) can be written as follows:

$$u'' + k^2 u = 0, \;\; |x| < x_0, \quad u'' + k^2 u - \sigma_0 u = 0, \;\; |x| \geq x_0. \tag{3.8}$$

Then we get

$$u(x) = c_1 \cos(kx), \;\; |x| < x_0, \quad u(x) = c_2 e^{\pm\lambda x}, \;\; |x| \geq x_0,$$

where c_1 and c_2 are positive constants, $\lambda = \sqrt{\sigma_0 - k^2}$ ($k^2 < \sigma_0$). From the continuity of the solution and of its first derivative at $x = \pm x_0$ we obtain the following equalities:

$$c_1 \cos(kx_0) = c_2 e^{-\lambda x_0}\,, \quad c_1 k \sin(kx_0) = c_2 \lambda e^{-\lambda x_0}\,. \tag{3.9}$$

Dividing the second equation by the first one, we get the equation with respect to k:

$$\sqrt{\sigma_0 - k^2} = k \tan(kx_0). \tag{3.10}$$

Let us recall that we look for a solution $k < 1$ of this equation. Such solution exists if x_0 is greater than some critical value x_0^*, and it does not exist if $x_0 < x_0^*$. For x_0 large enough, there are multiple solutions satisfying this condition.

We can now determine the integral U:

$$U = \int_{-\infty}^{\infty} u(x)dx = \frac{2c_1}{k}\sin(kx_0) + \frac{2c_2}{\lambda}e^{-\lambda x_0}\,.$$

Taking into account the first relation in (3.9), we have:

$$U = 2c_1\left(\frac{1}{k}\sin(kx_0) + \frac{1}{\lambda}\cos(kx_0)\right).$$

The coefficient c_1 is determined from equation (3.7):

$$c_1 = (1-k^2)/(2h(k)), \quad h(k) = \frac{1}{k}\sin(kx_0) + \frac{1}{\sqrt{\sigma_0 - k^2}}\cos(kx_0),$$

and $c_2 = c_1 e^{\lambda x_0}\;\cos(kx_0)$.

Thus, problem (3.6) has a positive solution decaying at infinity if the width $2x_0$ of the admissible interval where the mortality rate is less than the reproduction rate is sufficiently large. Otherwise, such solution does not exist, and the virus population goes to extinction. A similar analysis for an arbitrary diffusion coefficient D shows that a virus exists if D is less than some critical value. If the mutation rate determined by the diffusion coefficient is sufficiently large, then a virus strain does not exist.

3.2.2. Continuous mortality rates. We study here problem (3.6) under the assumption that $\sigma(x)$ is a bounded non-negative sufficiently smooth function. In order to prove the existence of positive solutions of this problem, we will apply the topological degree method. We begin with a priori estimates of solutions.

Lemma 3.2. *Let $u(x)$ be a positive solution of problem (3.6). Then $U < 1$.*

The proof of the lemma follows directly from the maximum principle. Indeed, if $U \geq 1$, then $u(x)$ is a solution of the equation $u'' + q(x)u = 0$ with $q(x) \leq 0$ and $q(x) \not\equiv 0$. Therefore, $u(x)$ cannot have positive maximum or negative minimum. Hence $u(x) \equiv 0$.

Lemma 3.3. *Suppose that $\sigma(x) = \sigma_0 > 1$ for $|x| \geq x_1$ with some positive σ_0 and x_1. Then $u(x_1) < \sqrt{\sigma_0}/2$.*

Proof. For $x \geq x_1$ equation (3.6) writes: $u'' - au = 0$, where $a = \sigma_0 - (1 - U) < \sigma_0$, $a > 0$. Then

$$u(x) = u(x_1)e^{-\sqrt{a}(x-x_1)}\ , \quad \int_{x_1}^{\infty} u(x)dx = \frac{u(x_1)}{\sqrt{a}} > \frac{u(x_1)}{\sqrt{\sigma_0}}\ .$$

Hence,

$$1 > U > 2\int_{x_1}^{\infty} u(x)dx > \frac{2u(x_1)}{\sqrt{\sigma_0}}\ .$$

This inequality proves the lemma.

□

Lemma 3.4. *Suppose that $\sigma(x)$ is a continuous function and $\sup_x \sigma(x) \leq M$. Then a positive solution $u(x)$ admits an estimate which depends only on M.*

Proof. Solution $u(x)$ of problem (3.6) satisfies the boundary problem

$$v'' + b(x)v = 0, \quad v(\pm x_1) = u(\pm x_1)$$

on the interval $-x_1 \leq x \leq x_1$. Here $b(x) = 1 - I(u) - \sigma(x)$ is a bounded continuous function, $|b(x)| \leq M + 1 \equiv m$. According to the previous lemma, the boundary values of the solution are bounded. Therefore, it is sufficient to estimate a maximum of the solution inside the interval. Suppose that the function $v(x)$ has a global maximum at some point $x_0 \in [-x_1, x_1]$. Then

$$|v'(x)| = |\int_{x_0}^{x} v''(y)dy| \leq mv(x_0)|x - x_0|.$$

Hence

$$v(x) = v(x_0) + \int_{x_0}^{x} v'(y)dx \geq v(x_0) - \frac{1}{2}mv(x_0)(x - x_0)^2 = v(x_0)g(x),$$

where $g(x) = \frac{1}{2} - m(x - x_0)^2$. Denote by Ω the interval in $[-x_1, x_1]$, where this function is positive. Then $\int_\Omega g(x)dx \geq \kappa > 0$, where the constant depends only on M and possibly on x_1. Hence, $1 > \int_{-\infty}^{\infty} v(x)dx > \kappa v(x_0)$. This estimate prove the lemma.

□

We will use the topological degree theory to prove the existence of solutions. Lemma 3.3 above provides a priori estimates of solutions. Consider the operator

$$A_\theta(u) = u'' + u(1-U) - \sigma_\theta(x)u,$$

acting from the weighted Hölder space $C^{2+\alpha}_\mu(\mathbb{R})$ into the space $C^\alpha_\mu(\mathbb{R})$. Here $0 < \alpha < 1$, $\theta \in [0,1]$ is a parameter. The space $C^{k+\alpha}_\mu(\mathbb{R})$ is defined as ensemble of functions $u(x)$ such that $u(x)\mu(x) \in C^{2+\alpha}(\mathbb{R})$. The weight function $\mu(x)$ is increasing at infinity with a polynomial rate. We set $\mu(x) = 1 + x^2$. The introduction of polynomial weighted is used for the definition of topological degree for elliptic operators in unbounded domains [53]. Moreover, it does not change the essential spectrum, and the integral U is well defined.

We will suppose for simplicity that $\sigma_\theta(x)$ is an infinitely differentiable function with respect to x and θ. Other conditions will be specified later.

Denote by L_θ the operator obtained by the linearization of the operator $A_\theta(u)$ about $u = 0$: $L_\theta v = v'' + v - \sigma_\theta(x)v$. The spectrum of the operator L_θ consists of the essential spectrum and of eigenvalues. The eigenvalue with the maximal real part is called the principal eigenvalue. If it is greater than the upper limit of the essential spectrum, then it is simple, positive, and the corresponding eigenfunction is positive [5, 55].

Lemma 3.5. *Suppose that the principal eigenvalue of the operator L_θ is positive for $\theta_0 \le \theta \le \theta_1$ and for some θ_0, θ_1. Then there exists $\epsilon > 0$ such that $u_m = \sup_x u(x) \ge \epsilon$ for any positive solution of equation $A_\theta(u) = 0$, $\theta_0 \le \theta \le \theta_1$.*

Proof. Suppose that the assertion of the lemma does not hold and there is a sequence of solutions $u_k(x)$ for $\theta = \theta_k$ such that $u_{m_k} \to 0$. Without loss of generality we can assume that $\theta_k \to \theta_*$ for some $\theta_* \in [\theta_0, \theta_1]$. Then

$$0 = A_{\theta_k}(u_k) = A_{\theta_k}(0) + L_{\theta_k}u_k + o(\|u_k\|) = L_{\theta_k}u_k + o(\|u_k\|).$$

Set $v_k = u_k/\|u_k\|$. Then $L_{\theta_k}v_k = o(1)$. Since L_{θ_k} is proper with respect to v and θ [53], then the sequence v_k is compact and we can choose a convergent subsequence $v_k \to v_0$. Hence, $L_{\theta_*}v_0 = 0$. Since the functions $u_k(x)$ are positive, then $v_0(x) > 0$ for all x. Therefore, the operator L_{θ_*} has a zero eigenvalue with a positive eigenfunction. However, the only positive eigenfunction corresponds to the principal eigenvalue [56]. We obtain a contradiction with the assumption that the principal eigenvalue of the operator L_{θ^*} is positive.

□

Theorem 3.6. *Suppose that $\sigma(x) = \sigma_0 > 1$ for $|x| \ge x_1$ with some positive σ_0 and x_1,and the principal eigenvalue of the problem*

$$u'' + u - \sigma(x)u = \lambda u \tag{3.11}$$

that is, the eigenvalue with the maximal real part is positive. Then problem (3.6) has a positive solution converging to 0 at infinity.

Proof. Set $\sigma_\theta(x) = (1-\theta)\sigma(x) + \theta\sigma_0$. Since $\sigma_0 > 1$, then the operator L_1 (i.e., L_θ for $\theta = 1$) has the spectrum in the left half-plane. Let us note that the essential spectrum $S_e(L_\theta)$ of the operator L_θ does not depend on θ, and Re $S_e(L_\theta) \le -\delta < 0$ for some positive δ. Denote the principal eigenvalue of this operator, that is, the eigenvalue

with the maximal real part, by $\lambda_0(\theta)$. According to the assumption of the theorem $\lambda_0(0) > 0$. It is a monotonically decreasing function of $\theta \in [0,1]$ [56], and there exists such $\theta_0 \in [0,1]$ that

$$\lambda_0(\theta_0) = 0, \;\; \lambda_0(\theta) > 0 \;\text{ for }\; 0 < \theta \le \theta_0, \;\; \lambda_0(\theta) < 0 \;\text{ for }\; \theta_0 < \theta \le \theta_1.$$

Here θ_1 is some value in the interval $(\theta_0, 1]$. Since the eigenvalue can approach the essential spectrum, we cannot guarantee its existence for all $\theta \in [0,1]$.

Let us consider the equation $A_\theta(u) = 0$ in a small vicinity of the bifurcation point $\theta = \theta_0$. For this value of parameter, the trivial solution $u = 0$ loses its stability leading to the appearance of other solutions $u_\theta(x)$ and $\tilde{u}_\theta(x)$. One of them is positive and another one is negative since the principal eigenfunction $v_0(x)$ is positive [56]. Furthermore, the index of this solution, that is, the value of the degree with respect to a small ball containing this solution, equals 1. Indeed, from the homotopy invariance of the degree it follows that

$$\mathrm{ind}(0) + \mathrm{ind}(u_\theta) + \mathrm{ind}(\tilde{u}_\theta) = 1$$

for all $\theta > \theta_0$ and sufficiently close to θ_0. Since $\mathrm{ind}(0) = -1$ being equal $(-1)^\nu$, where $\nu = 1$ is the number of positive eigenvalues of the linearized operator, then $\mathrm{ind}(u_\theta) = \mathrm{ind}(\tilde{u}_\theta) = 1$.

It follows from Lemma 3.4 and exponential decay of solutions at infinity that $\|u\|_{C^{2+\alpha}_\mu(\mathbb{R})} < M_0$ for some positive constant M_0 and for any positive solution u of the equation $A_\theta(u) = 0$. Next, from Lemma 3.5 we conclude that $\|u\|_{C^{2+\alpha}_\mu(\mathbb{R})} > \delta(\theta)$ for some positive $\delta(\theta)$, $\theta < \theta_0$. Consider the following domain

$$\Omega = \{u \in C^{2+\alpha}_\mu(\mathbb{R}),\; u(x) > 0, x \in \mathbb{R},\; \delta_0 < \|u\|_{C^{2+\alpha}_\mu(\mathbb{R})} < M_0\}$$

for some $\delta_0 > 0$ sufficiently small. Choose $\theta_2 < \theta_0$ such that $\delta(\theta) > \delta_0$ for $0 \le \theta \le \theta_2$. Since $A_\theta(u) \neq 0$ for $u \in \partial\Omega$, $0 \le \theta \le \theta_2$, then the value of the degree $\gamma(A_\theta, \Omega)$ does not depend on $\theta \in [0, \theta_2]$. Hence, $\gamma(A_0, \Omega) = \gamma(A_{\theta_2}, \Omega) = \mathrm{ind}(u_{\theta_2}) = 1$, and equation $A_0(u) = 0$ has a solution in Ω.

□

Let us note that a similar method to prove the existence of pulses can be used for the case for nonlocal consumption where the integral U is replaced by the integral $J(u)$. In order to illustrate the application of this theorem, consider the following example.

Example 3.7. Let $\beta(x) = 0$ for $|x| \le x_0$ and $\beta(x) = \sigma_0 > 1$ for $|x| > x_0$ (cf. Section 3.1). Consider an averaged function

$$\sigma(x) = \int_{x-\epsilon}^{x+\epsilon} \beta(y)dy$$

for some small $\epsilon > 0$. This function satisfies conditions of Theorem 3.6. The principal eigenvalue λ_0 of the operator $Lu = u'' + u - \sigma(x)u$ is real and simple [56]. Consider it as a function of x_0, $\lambda_0 = \lambda_0(x_0)$. Since the function $\sigma(x)$ monotonically decreases for all x as x_0 increases, then $\lambda_0(x_0)$ is a monotonically increasing function. If x_0 is sufficiently large, then $\lambda_0(x_0)$ becomes positive, and the theorem is applicable. Thus, similar to the piece-wise constant mortality rate considered in Section 3.1, a virus strain exists if the interval of admissible genotypes is sufficiently large.

We can estimate the critical value of x_0. If $x_0 > \pi/2 + \epsilon$, then the function $u_0(x) = \cos x$ for $|x| \le \pi/2$ and $u(x) = 0$ for $|x| > \pi/2$ is a lower function for the operator Lu, that is, the solution of the Cauchy problem $\frac{\partial u}{\partial t} = Lu$ with the initial condition $u(x) = u_0(x)$ is a growing function of t for any $x \in \mathbb{R}$. Therefore, $\lambda_0(x_0) > 0$ for $x_0 > \pi/2 + \epsilon$.

3.3. Existence of solutions of a local problem. Finally, we will prove the existence of pulses for the local problem where the integral U is replaced by u. Consider the equation

$$u'' + u(1-u) - \sigma(x)u = 0 \tag{3.12}$$

on the whole axis. Here $\sigma(x) = \sigma_0 > 1$ for $|x| \ge x_0$, and $\sigma(x) = 0$ for $|x| < x_0$, x_0 is some positive number. We look for a positive bounded solution of this equation with zero limits at infinity.

Equation

$$u'' + u(1-u) - \sigma_0 u = 0 \tag{3.13}$$

considered on the whole axis has a single (up to translation in space) positive solution $u_0(x)$ decaying to 0 at infinity. Hence, the same equation considered on the half-axis $x \ge x_0$ has a family of positive decaying solutions $v_h(x) = u_0(x + h)$ for all real h. The choice of one solution from this family of solutions is determined by the value $v_h(x_0)$. The functions $v_h(-x)$ provide solutions on the half-axis $x \le -x_0$.

Next, consider the equation

$$u'' + u(1-u) = 0 \tag{3.14}$$

on the interval $-x_0 < x < x_0$ with the boundary conditions $u(x_0) = u(-x_0) = a$ for some $a > 0$. Denote its solution by $u_a(x)$. If for some a and h,

$$u_a(x_0) = v_h(x_0), \quad u_a'(x_0) = v_h'(x_0), \tag{3.15}$$

then the function

$$u(x) = \begin{cases} u_a(x) & , \quad -x_0 < x < x_0 \\ v_h(|x|) & , \quad |x| \ge x_0 \end{cases}$$

is a solution of equation (3.12).

Lemma 3.8. There exists $x_* > 0$ such that for any $x_0 > x_*$ equation (3.12) has a positive solution decaying at infinity.

Proof. Consider the first-order system of equations

$$u' = p, \quad p' = -u(1-u) - \sigma_0 u \tag{3.16}$$

corresponding to equation (3.13), and the system

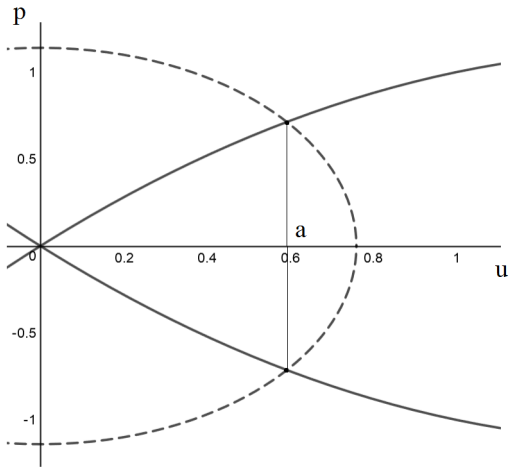

FIG. 2. *Trajectories of system (3.16) (solid lines) and of system (3.17) (dashed line).*

$$u' = p, \quad p' = -u(1-u) \tag{3.17}$$

corresponding to equation (3.14). Since $(0,0)$ is a saddle stationary point of system (3.16), then there are two trajectories approaching it from the half-plane $u > 0$ as $x \to \pm\infty$ (Figure 2, solid lines). The same stationary point is a center for system (3.17), and there is a family of limit cycle trajectories around it (dashed line). Denote by a the coordinate of their intersection, and consider two solid trajectories between $u = 0$ and $u = a$ and the arc of the dashed trajectory between the solid trajectories. These three trajectories provide solutions of equation (3.13) for $x > x_1$ and $x < -x_1$ and of equation (3.14) for $-x_1 < x < x_1$ with some x_1. Moreover, they satisfy conditions similar to (3.15) where x_0 is replaced by x_1. The value x_1 is determined by the length of the dashed arc, and in general x_1 is different from x_0.

We will now vary the value of a and consider the corresponding values $x_1 = x_1(a)$. For $a \to 0$, x_1 converges to some positive limiting value denoted by x_*. It can be explicitly found from the solution of the problem linearized about the stationary point $(0,0)$. Increasing a in such a way that the dashed trajectory converges to the stationary point $(1,0)$, we get $x_1(a) \to \infty$. Therefore, equation $x_1(a) = x_0$ has a solution for any $x_0 > x_*$.

□

We can use this lemma and the method of upper and lower functions in order to prove existence of solutions of equation (3.12) for more general functions $\sigma(x)$.

Theorem 3.9. Suppose that $\sigma(x) = 0$ for $|x| \le x_0$ and $\sigma(x) \ge 1$ for $|x| \ge x_1$, where $x_1 > x_0 > \pi/2$. Then equation (3.12) has a positive solution decaying at infinity.

Proof. Consider the arc of the periodic trajectories of system (3.17) in the half-plane $u > 0$ (cf. dashed trajectory in Figure 2). Denote the maximal value of u along such trajectory by u_m. Therefore, such trajectories exist for all $u_m \in (0,1)$. Each such trajectory corresponds to a solution of equation (3.14) on some interval $-\hat{x} \le x \le \hat{x}$, where $u(\pm\hat{x}) = 0, u(0) = u_m$. Furthermore, $\hat{x} \to \pi/2$ as $u_m \to 0$. Hence, for u_m sufficiently small, $\hat{x} < x_0$. Then the function

$$u_-(x) = \begin{cases} u_m(x) & , \quad -\hat{x} < x < \hat{x} \\ 0 & , \quad |x| \geq x_0 \end{cases}$$

is a lower function for equation (3.12).

Set $\sigma_0(x) = 0$ for $|x| < x_1$ and $\sigma_0(x) = 1$ for $|x| \geq x_1$. This function satisfies conditions of Lemma 3.8. Therefore, equation

$$u'' + u(1-u) - \sigma_0(x)u = 0$$

has a positive solution $u_+(x)$ decaying at infinity. Since $\sigma(x) \geq \sigma_0(x)$, then $u_+(x)$ is an upper function for equation (3.12). It remains to note that $u_-(x) < u_+(x)$, $x \in \mathbb{R}$ for u_m sufficiently small. The methods of upper and lower functions provides existence of solutions of equation (3.12).

□

Thus, pulses (strains) for the local problem exist for sufficiently wide admissible intervals. From this point of view, this result is similar to the existence result presented in Section 3.2. However, these two cases become essentially different in the case of two (or more) admissible intervals when the strains compete with each other. We will discuss this question in the next section. Since the problem is symmetric with respect to the center of the admissible interval, the maximum of the solution is reached at $x = 0$.

4. Dynamics of strains. Existence of pulses is proved above for some particular cases. In this section, we will use numerical simulations to show their existence and stability for more general functions $f(u)$ and $\sigma(x)$. Namely, we will consider non-monotone functions $f(u)$ corresponding to the biologically realistic properties of immune response and we will consider the mortality rate $\sigma(x)$ with two admissible intervals in order to study the competition of virus strains. We present here numerical simulations of the equation

$$\frac{\partial u}{\partial t} = D\,\frac{\partial^2 u}{\partial x^2} + ru(1 - qJ(u)) - uf(u) - \sigma(x)u \tag{4.1}$$

on the interval $0 < x < L$ with the periodic boundary conditions. Extending the solution on the whole axis by periodicity, we define the integral $J(u) = \int_{-\infty}^{\infty} \phi(x - y)u(y,t)dy$ and consider a piece-wise constant kernel, $\phi(x) = 1/(2N), |x| \leq N, \phi(x) = 0, |x| > N$. This approach is convenient since it decreases the influence of the boundary. Otherwise, the integral $J(u)$ differs near the boundary, and perturbs the solution. If the peak of the solution is sufficiently far from the boundary, and the support of the kernel $\phi(x)$ is small enough, then the influence of the boundary is negligible. Continuation of the problem by periodicity implies that the function $\sigma(x)$ is also periodic, and the support of the kernel $\phi(x)$ is limited by L. For large values of N, the integral $J(u)$ approximates U and for small values of N, $J(u)$ converges to u. Thus, varying this parameter, we can study all three cases, local, nonlocal, and global.

Existence and stability of virus strains. We consider the function $f(u) = (k_1 u + k_2)e^{-k_3 u}$ with some positive values of constants $k_i, i = 1, 2, 3$ specified below, and set $\sigma(x) \equiv 0$. The growing branch of this function corresponds to the clonal expansion of immune cells in response to the antigen, its decaying branch describes the exhaustion of immune cells for high virus concentrations. Let us recall that the local reaction-diffusion equation can have pulse solution only in the bistable case, and it is unstable.

The nonlocal reaction-diffusion equation can have stable stationary solutions in the form of a single pulse, and also multiple pulses. The latter are not stationary solutions, they slowly move from each other with a decaying speed. Similar properties were already observed in [53, 54] for another nonlocal equation, and we will not discuss them here in detail.

Contrary to the previously studied cases where stable pulses exist only for the bistable nonlinearity [53, 54], equation (2.2) admits their existence in both cases, bistable and monostable, due to the decreasing function $f(u)$. In the bistable case, there are two pulses, one of them is stable, another one is unstable. The latter separates the trivial solution and the stable pulse. Therefore, in order for the solution of equation (2.1) to approach the stable pulse, the initial condition should be sufficiently large. This is different in the monostable case. For any small non-trivial initial condition, the solution converges to the stable pulse. This difference is important for the understanding of the emergence of virus strains.

Existence and stability of pulses depends on the parameter N characterizing the width of the kernel. If N is small enough, then the nonlocal equation is similar to the local one, and pulses are unstable. For N sufficiently large, they become stable.

Competition of virus strains. We now consider existence and properties of pulses due to the genotype dependent mortality rate $\sigma(x)$ and without immune response, $f(u) \equiv 0$. In order to study the interaction of different virus strains, we set

$$\sigma(x) = \begin{cases} 0 & , \quad x \in I_1, x \in I_2 \\ 1 & , \quad x \in [0,1], x \notin I_1, I_2 \end{cases} .$$

Hence, we consider two admissible intervals where the birth rate exceeds the mortality rate. Each admissible interval corresponds to a virus strain. The initial condition (Figure 3, right, $t = 0$) is localized at the center of the interval. Numerical simulations presented in this section are carried out in time interval $0 < t \le T$ with $T = 100$ (dimensionless time units). By stationary solutions we understand here the solutions which do not change in this time scale. Numerical simulations in much longer time scale ($T \sim 10^5$) show that these solutions can be slowly evolving.

Solution of equation (4.1) converges to the stationary solution whose structure depends on the value of the diffusion coefficient (Figure 3, left column). For D large enough, the maxima of solutions in the admissible intervals are equal to each other. For small D, solution in one of the intervals is much less than in the other interval, or it practically vanishes. Such behavior of solutions is determined by the competition between the two strains. The strength of this competition depends on the width N of the support of the kernel $\phi(x)$.

Figure 4 shows the results of the simulations for different values of N. For N sufficiently small, we obtain a solution similar to the case of the local equation where $J(u) = u$. In this case, the peaks of solutions in the admissible intervals are equal to each other, that is, the two virus strains have a similar population density. For larger values of N, though the peaks are still equal to each other, the solution is more restraint to the admissible intervals, and its value outside the intervals decreases (cf. upper left and middle left images). For N large enough, there is a transition to the stationary solution with one large and one small peak. The critical value of N is determined by the distance between the admissible intervals. If N is less than this distance, then there is no competition between the strains.

Epitope-specific immunodominant CTL responses. Suppose that the immune response depends on the virus genotype due to the immunodominance of antigen-specific

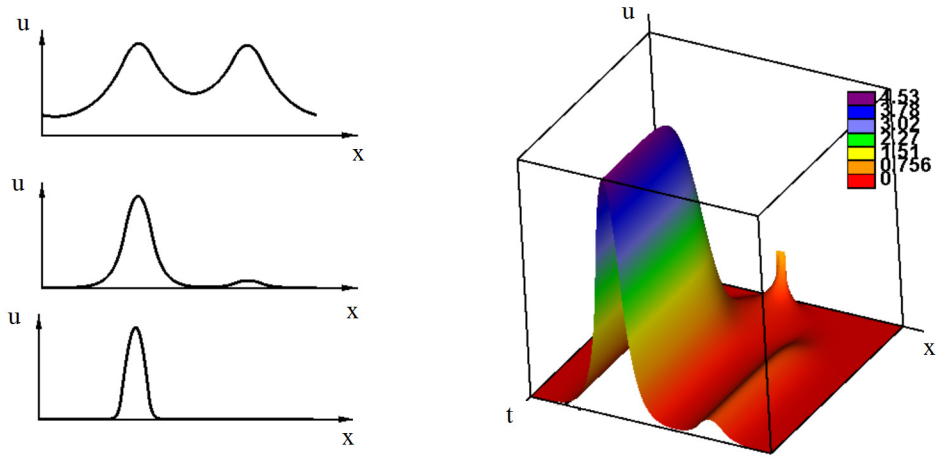

FIG. 3. *Stationary solutions of equation (4.1) with* $D = 0.005$ *(upper left),* $D = 0.001$ *(middle left),* $D = 0.0001$ *(lower left). The right image shows the solution* $u(x,t)$ *as a function of two variables for* $D = 0.001$*. Other values of parameters are as follows:* $r = 1, q = 1, L = 1, N = 0.5$*,* $I_1 = [0.3, 0.4], I_2 = [0.7, 0.8]$.

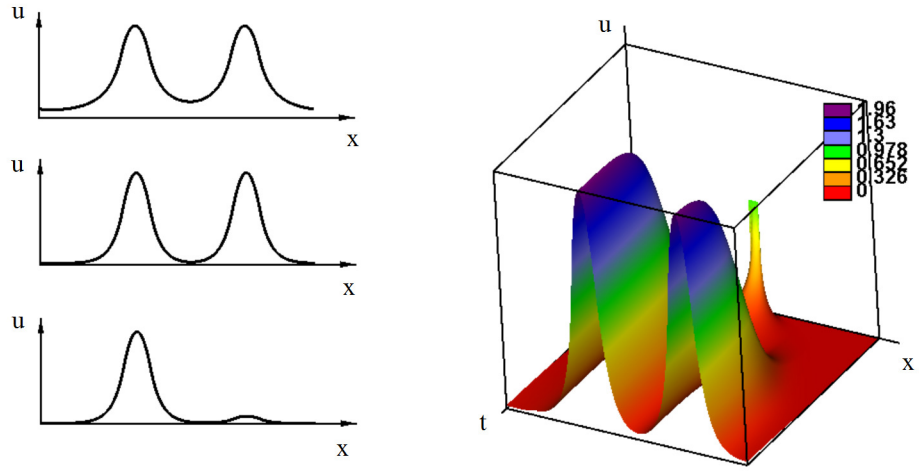

FIG. 4. *Stationary solutions of equation (4.1) with* $N = 0.01$ *(upper left),* $N = 0.2$ *(middle left),* $N = 1$ *(lower left). The right image shows the solution* $u(x,t)$ *as a function of two variables for* $N = 0.2$*. Other values of parameters are as follows:* $r = 1, q = 1, L = 1, D = 0.001$*,* $I_1 = [0.3, 0.4], I_2 = [0.7, 0.8]$.

immune cells resulting in a skewed response to only few specific epitopes [18]. Indeed, the magnitudes and frequencies of virus-specific CTL responses can spread unevenly across the viral genome, within and between most of the viral proteins [35]. In addition, mutations leading to changes of amino-acids in TCR epitopes result in changes of TCR 2D affinity [24]. This can also be a result of an immunotherapy that increases

the effect of the immune response in an antigen-dependent way. We consider equation (3.1) with $\sigma(x) \equiv 0$ and $f(u) = (k_1(x)u + k_2)e^{-k_3 u}$. Figure 5 shows two examples of numerical simulations. In the first case (left image), there is a single pulse solution corresponding to a virus strain. If immune response is genotype independent, then it is a stationary solution. In the case of a variable function $k_1(x)$, the pulse moves towards its minimum where the elimination of virus by immune cells is weaker. In the second example (right image), there are two pulses moving in the direction of the decrease of the function $k_1(x)$. After some time, one of the pulses splits into three other pulses (strains). They push the second pulse aside and stabilize in the left part of the interval where the immune response is weaker. Hence, antigen-dependent immune responses can lead to the emergence of new strains and to their evolution in the genotype space.

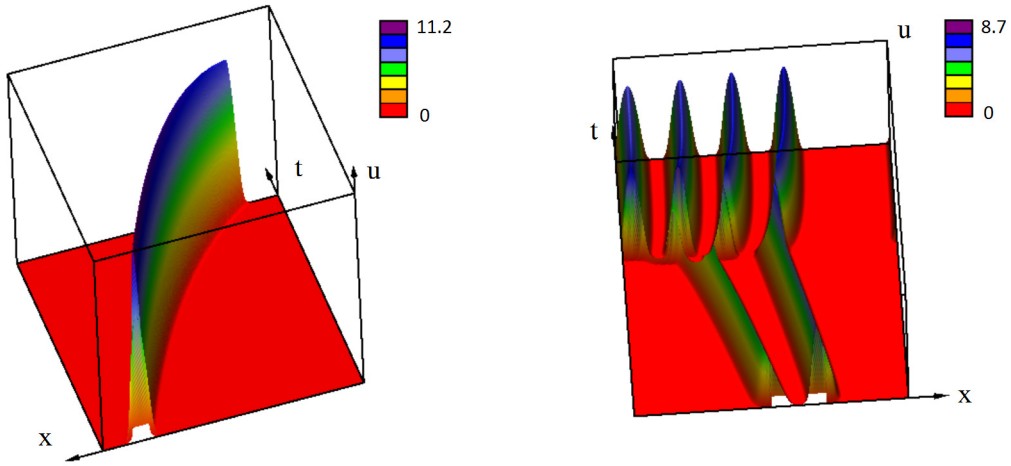

FIG. 5. *Virus dynamics in the case of a genotype-dependent immune response. The solution $u(x,t)$ is represented as a function of two variables. Left: a single virus strain evolves in the direction of a weaker immune response. Right: two virus strains evolve together while one of them splits into three other strains. The values of parameters are as follows: $r = 1, q = 0.4, L = 1, D = 0.00005, N = 0.1$; $f(u) = (k_1(x)u + k_2)e^{-k_3 u}$, $k_1(x) = k_1^0 x$; $k_1^0 = 10, k_2 = 1.5, k_3 = 1$ (left, $k_1^0 = 0$ during small initial period of time); $k_1^0 = 5, k_2 = 1.5, k_3 = 2$ (right).*

Elimination of a susceptible virus strain can lead to the emergence of a resistant strain. We begin with the case without an immune response ($f(u) = 0$) and consider equation (2.1) where $H = (1 - qJ(u))$. This non-local equation describes the competition of different genotypes for host cells. Suppose that the function $\sigma(x)$ has two admissible intervals where the virus reproduction rate is larger than its mortality. If the support of the kernel $\phi(x)$ in the integral $J(u)$ is sufficiently narrow, then the two virus strains coexist. If the support of the kernel is wide enough, then the strains compete with each other, and only one of them survives. The choice of the surviving strain is determined by their relative susceptibility or, if it is the same, by the initial condition.

Suppose that treatment acts on the first strain but not on the second. This assumption corresponds to the observation that susceptible strains are better fit. Then treatment eliminates the first strain. Since there is no more competition, the second strain, being resistant to treatment, emerges and persists. Before treatment, the function $\sigma(x)$ equals 0 in two admissible intervals I_1 and I_2, and $\sigma(x) = 1$ elsewhere.

An antiviral drug acts on the first strain. During treatment, the function $\sigma(x)$ becomes also equal 1 in the interval I_1 mimicking drug action.

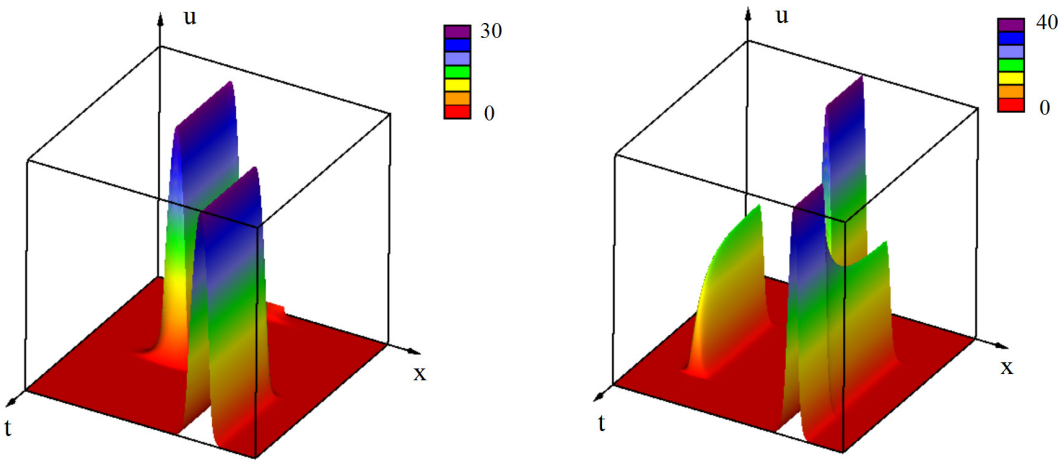

FIG. 6. *Emergence of a resistant virus strain due to antiviral treatment. Left: elimination of the first strain leads to the emergence of the second strain. Solution $u(x,t)$ of equation (3.1) is shown for the values of parameters $r = 1, q = 0.4, \tau = 0, L = 1, D = 0.0001, N = 0.5$, and admissible interval $I_1 = [0.3, 0.4], I_2 = [0.7, 0.8]$. Before treatment, $\sigma(x) = 0$ in both admissible intervals. Treatment is modeled by imposing $\sigma(x) = 1$ in the first interval. Right: three admissible intervals $I_1 = [0.2, 0.25], I_2 = [0.475, 0.525], I_3 = [0.75, 0.8]$, $N = 1$ with the presence of immune response (equation (2.2)), $f(u) = k_1 e^{-k_3 u}$, $k_1 = k_3 = 1$, and asymmetric initial condition. After the application of treatment, the central strain disappears leading to the emergence of two other strains. Their competition and the action of immune response result in the elimination of one of the two strains.*

The mechanism of the emergence of resistant strains is illustrated in Figure 6 (left). In the beginning, only the first strain is present. Applying treatment, we eliminate it. Its concentration decreases, while the concentration of the second strain increases. After some time, there is only the second strain. The second strain appears discontinuously, and thus is not as a result of a gradual evolution of the first strain.

The second example of the antiviral treatment is shown for the complete model (2.2) in which the immune response is taken into account (Figure 6, right). There are three admissible intervals, and in the beginning, there is only one virus strain corresponding to the central interval. After the application of antiviral treatment, this strain disappears, and two other strains emerge. Being in competition with each other, one of them disappears, while another one increases its concentration. Coexistence or disappearance of the strains are determined by the action of the immune response.

5. Discussion.

Model. In this study we have developed a general mathematical model for the analysis of the evolutionary dynamics of viruses in genotype space. The model, formulated with nonlocal delay reaction-diffusion equations, considers the viral evolution within an infected host under the impact of mutations, target-cell competition, cross-reactive immune responses and antiviral treatment. We provide a single mathematical framework that allows the assessment and prediction of the impact of a range of processes on the genetic changes during virus infections. This work contributes to a rapidly developing field of applications of nonlocal reaction-diffusion equations in the evolution theory [19, 37], ecology [4], cancer modelling [28], etc., and of their mathe-

matical analysis and nonlinear dynamics [2, 3, 20, 34, 49] (see more detailed literature review in [53]). Our study extends modelling of adaptive and innate immune response in various physiological situations, and the analysis of spatiotemporal dynamics of the within-host infections [21, 16].

Dynamics of virus strains. Similar to biological species, virus strains are characterized by their genotypes. From the mathematical point of view, they can be viewed as some localized density distributions around some average (most frequent) genotypes. Existence of such persistent distributions is not a priori given, and it should be obtained as stable stationary solutions of appropriate models. We propose in this work different mechanisms leading to the existence of such solutions:

- The diffusion-advection equation presented in Supplementary Materials possesses a solution in the form of normal distribution. It does not correspond to a single strain but to an ensemble of strains with the same number of mutations.
- In the case of a genotype-dependent virus mortality rate, a virus strain considered as a localized positive solution exists if the admissible interval is sufficiently large or if the diffusion coefficient (mutation rate) is sufficiently low. Such solutions exist with and without an immune response, and with the local or global competition for host cells. The latter can also describe the competition of different strains, and it provides a possible mechanisms for the resistance to treatment.
- Taking into account an immune response (without a genotype-dependent mortality rate) we obtain stable stationary solutions (strains) due to its interaction with the nonlocal or global terms in the virus production rate (competition for host cells).

The last mechanism is dominated by an immune response while the previous one by a genotype-dependent mortality. Both of them eliminate less fit viruses. An experiment in cell culture (without immune cells) can possibly give an additional information about these mechanisms and determine whether a virus strain persists or its genotype blur with time (depending on virus type and culture conditions).

Biological insights. Our model describes the extinction of a virus population when the random diffusion rate in the genotype space, which is proportional to the mutation probability, exceeds a certain threshold known as error catastrophe. It reproduces the emergence of viral mutants resistant to genotype-specific drugs and the escape from epitope-specific immune responses. Being consistent with these fundamental regularities of viral evolution, the model also predicts novel types of evolutionary dynamics such as travelling waves [15, 44, 22, 41], pulses, pulse bifurcations and periodic waves determined by non-local genotype-dependent interactions within a quasispecies. Elegant experiments with poliovirus infections of mice indicated that the size of the genomic diversity, or genotype space as in our conceptual model, influences virus adaptability [50, 51]. Our model now gives a potential mechanistic explanation for this feature, namely the adaptation in response to genotype-dependent non-local interactions of various origins.

The phenomena of strain-dependent virus cell tropism, escape of acute CTL responses and the selection of drug-resistant mutants are well documented for infections such as HIV, HCV and Influenza virus [33, 14]. Mathematical models have been extensively used to address various aspects of viral evolution mostly using ODEs or Fisher-KPP-type of equations. The non-local delay RDE model presented in our study serves as a further generalization of the above models. Using powerful analytical and numerical tools for its analyses, novel insights into the emergence of regular dynamic patterns pertinent to escape, diversification and extinction of viral strains in the infected host due to non-local genotype-dependent interactions with host cells

and the immune system are obtained. The model can be further applied to study specific patterns of viral adaptation towards antivirals. For example, the dependence of mutations on the genotype known as epistatic interactions can be described by considering the diffusion operator to be dependent on x, i.e. $D = D(x)$. This enables the analysis of complex mutational patterns, e.g., mutation pathways of the HIV-1 reverse transcriptase under AZT therapy [30].

In view of the demand for designing combination therapies to cure chronic viral infections like HIV and HBV, the formulated model and the developed analytical framework could be effectively used to quantify virus fitness, i.e. the genotype to phenotype mapping as an emergent property of non-local interactions between viruses and host factors parameterized in the genotype space. This may provide a better understanding of the adaptation of mutating viruses to changes of the host environment.

Treatment and resistance. Antiviral treatment can be modelled by means of a genotype dependent-mortality rate. The initial mortality rate $\sigma_0(x)$ can be replaced at some moment of time by another function $\sigma_1(x)$ where one of the admissible intervals is removed. This means that the corresponding virus strain is eliminated by the treatment. The response of other strains can be different depending on the competition between them for host cells and on the immune response. In particular, new strains can emerge in other admissible intervals. They could not develop before the application of treatment because of the competition with the dominating (e.g., wild-type) strain. Since the treatment does not act on these other admissible intervals, we can obtain resistant strains as a result of treatment.

Assuming that treatment can stimulate a genotype-dependent immune response (immunotherapy), we observe the evolution of virus strains in the genotype space towards less susceptible (more resistant) strains. This dynamics can be quite complex including gradual modification of the existing strains or the emergence of new strains.

These mechanisms of the emergence of resistance to treatment are different in comparison with the mechanism considered in [1, 38, 39]. The latter is based on the assumption of the existence of two endemic equilibria. One of them is dominated by a drug-susceptible virus and another one by a drug-resistant mutant. Before treatment, the former is stable, while the latter is unstable. Antiviral treatment changes their stability. The first mechanism suggested in our work, interpreted in terms of ODE systems, corresponds to the system of competition of species with two stable stationary points and convergence to one of them determined by the initial condition. The nonlocal reaction-diffusion equation describes a similar mechanism taking into account the virus distribution with respect to the genotype variable.

The nonlocal reaction-diffusion model considers the width of the admissible intervals (e.g., I_1 and I_2) where the mortality rate is smaller than the birth rate. In the context of drug-resistance analysis, these components of the model can be considered as quantitative representation of the genetic barrier to drug resistance. The latter is related to the number of mutations which are needed for the mutants to become insensitive to certain drugs, i.e. to reduce their mortality rate σ to zero. Therefore, the model provides a tool for examining the impact of various levels of the genetic barrier on the evolutionary dynamics of the viral population.

Limitations and perspectives. The qualitative model presented in this work does not take into account complex intracellular mechanisms of virus multiplication, involvement of different types of immune cells, various aspects of molecular and cellular regulation. Nonetheless, these simplifications allow us to reveal some generic biological mechanisms which can be more difficult to identify in more complete models.

Once these hypothetical mechanisms are delineated, more detailed models and biological experiments can be used to study them. We expect that this work will open interesting perspectives in mathematical modelling of viral infections. Furthermore, some of these mechanisms can also hold for the evolutionary dynamics of cancer cells.

Acknowledgements. The research was funded by the Russian Science Foundation (Grant no. 18-11-00171) to N.B., G.B. A.M. and V.V. V.P. and V.V. were partially supported by the RUDN University Strategic Academic Leadership Program. G.B. was partly supported by Moscow Center for Fundamental and Applied Mathematics (agreement with the Ministry of Education and Science of the Russian Federation No. 075-15-2019-1624). A.M. was also supported by a grant from the Spanish Ministry of Economy, Industry and Competitiveness and FEDER grant no. SAF2016-75505-R (AEI/MINEICO/FEDER, UE) and the Maria de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0370).

REFERENCES

[1] H.K. Alexander, S. Bonhoeffer. Pre-existence and emergence of drug resistance in a generalized model of intra-host viral dynamics. Epidemics, 4 (2012) 187202.
[2] M. Alfaro, J. Coville, G. Raoul. Bistable travelling waves for nonlocal reaction diffusion equations. Discret. Contin. Dyn. Syst. 34 (2014) pp. 1775–1791.
[3] M. Alfaro, A. Ducrot, T. Giletti. Travelling waves for a non-monotone bistable equation with delay: existence and oscillations. Proc. London Math. Soc., 116 (2018), no. 3, 729-759.
[4] M. Banerjee, V. Volpert. Spatio-temporal pattern formation in RosenzweigMacarthur model: effect of nonlocal interactions. Ecological complexity, 30 (2017), 2-10.
[5] H. Berestycki, L. Nirenberg, S.R.S. Varadhan, The Principal Eigenvalue and Maximum Principle for Second-Order Elliptic Operators in General Domains, CPAM, 1994, 4-92.
[6] Biebricher CK, Eigen M. What is a quasispecies? Curr Top Microbiol Immunol. 2006;299:1-31.
[7] G. Bocharov, V. Volpert, B. Ludewig, A. Meyerhans. Mathematical Immunology of Virus Infections. Springer, 2018.
[8] G. Bocharov, A. Meyerhans, N. Bessonov, S. Trofimchuk, V. Volpert. Spatiotemporal dynamics of virus infection spreading in tissues. PlosOne, DOI:10.1371/journal.pone.0168576 December 20, 2016.
[9] G. Bocharov, A. Meyerhans, N. Bessonov, S. Trofimchuk, V. Volpert. Modelling the dynamics of virus infection and immune response in space and time, International Journal of Parallel, Emergent and Distributed Systems. (2017)
[10] N.F. Britton. Spatial structures and periodic travelling waves in an integro-differential reaction-diffusion population model. SIAM J. Appl. Math., 6 (1990), 1663-1688.
[11] Buzon, P, Maity, S, Roos, WH. Physical virology: From virus selfassembly to particle mechanics. WIREs Nanomed Nanobiotechnol. 2020; 12:e1613.
[12] Coffin J, Swanstrom R. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. Cold Spring Harb Perspect Med. 2013 Jan 1;3(1):a012526.
[13] Esteban Domingo, Julie Sheldon, Celia Perales. Viral Quasispecies. Evolution Microbiol Mol Biol Rev. 2012 Jun; 76(2): 159216
[14] Natalia Echeverra, Gonzalo Moratorio, Juan Cristina, Pilar Moreno. Hepatitis C virus genetic variability and evolution. World J Hepatol. 2015 Apr 28; 7(6): 831845.
[15] Fisher R.A. The Wave of Advance of Advantageous Genes. Annals of Eugenics. (1937) 7 (4): 353369. doi:10.1111/j.1469-1809.1937.tb02153.x.
[16] A. Friedman, W. Hao. A mathematical model of atherosclerosis with reverse cholesterol transport and associated risk factors. Bull Math Biol. 77 (2015), 758781.
[17] Frost SD, Wrin T, Smith DM, Kosakovsky Pond SL, Liu Y, Paxinos E, Chappey C, Galovich J, Beauchaine J, Petropoulos CJ, Little SJ, Richman DD. Antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection. Proc Natl Acad Sci U S A. 2005 Dec 20;102(51):18514-9
[18] Gallimore A, Hengartner H, Zinkernagel R. Hierarchies of antigen-specific cytotoxic T-cell responses. Immunol Rev. 1998 Aug;164:29-36.
[19] S. Genieys, V. Volpert, P. Auger. Adaptive dynamics: modelling Darwin's divergence principle. Comptes Rendus Biologies, 329 (11), 876-879 (2006).
[20] S.A. Gourley, M.A.J. Chaplain, F.A. Davidson. Spatio-temporal pattern formation in a nonlocal

reaction-diffusion equation. Dynamical systems, 16 (2001), no. 2, 173192.
[21] W. Hao, L.S. Schlesinger, A. Friedman. Modeling granulomas in response to infection in the lung. PLoS One. 11 (2016), no. 3, e0148738.
[22] Haraguchi Y, Sasaki A. Evolutionary pattern of intra-host pathogen antigenic drift: effect of cross-reactivity in immune response. Philos Trans R Soc Lond B Biol Sci. 1997 Jan 29;352(1349):11-20.
[23] Motoo Kimura. Diffusion Models in Population Genetics Journal of Applied Probability, Vol. 1, No. 2. (Dec., 1964), pp. 177-232.
[24] Kolawole EM, Andargachew R, Liu B, Jacobs JR, Evavold BD. 2D Kinetic Analysis of TCR and CD8 Coreceptor for LCMV GP33 Epitopes. Front Immunol. 2018 Oct 15;9:2348.
[25] A. Kolmogorov, I. Petrovskii, N. Piskunov. A study of the diffusion equation with increase in the amount of substance," and its application to a biological problem. In V. M. Tikhomirov, editor, Selected Works of A. N. Kolmogorov I, pages 248270. Kluwer 1991, ISBN 90-277-2796-1. Translated by V. M. Volosov from Bull. Moscow Univ., Math. Mech. 1, 125, 1937.
[26] G. Lawyer, A. Altmann, A. Thielen, M. Zazzi, A. Sönnerborg, T. Lengauer. HIV-1 mutational pathways under multidrug therapy. AIDS Research and Therapy 2011, 8:26.
[27] Lind PA, Libby E, Herzog J, Rainey PB. Predicting mutational routes to new adaptive phenotypes. Elife. 2019 Jan 8;8. pii: e38822.
[28] A. Lorz, T. Lorenzi, J. Clairambault, A. Escargueil, B. Perthame. Modeling the effects of space structure and combination therapies on phenotypic heterogeneity and drug resistance in solid tumors. Bulletin of mathematical biology 77 (2015) (1), 1-22.
[29] Yi-Fan Lu, David B. Goldstein, Misha Angrist, and Gianpiero Cavalleri. Personalized Medicine and Human Genetic Diversity. Cold Spring Harb Perspect Med September 2014;4:a008581
[30] Martinez JP, Bocharov G, Ignatovich A, Reiter J, Dittmar MT, Wain-Hobson S, Meyerhans A. Fitness ranking of individual mutants drives patterns of epistatic interactions in HIV-1. PLoS One. 2011 Mar 31;6(3):e18375.
[31] Mauricio G. Mateu, Virus engineering: functionalization and stabilization, Protein Engineering, Design and Selection, Volume 24, Issue 1-2, January-February 2011, 5363.
[32] Mohammadi P, Desfarges S, Bartha I, Joos B, Zangger N, Muoz M, Gnthard HF, Beerenwinkel N, Telenti A, Ciuffi A. 24 hours in the life of HIV-1 in a T cell line. PLoS Pathog. 2013, 9(1):e1003161.
[33] Murin CD, Wilson IA, Ward AB. Antibody responses to viral infections: a structural perspective across three different enveloped viruses.Nat Microbiol. 2019 Mar 18.
[34] G. Nadin, L. Rossi, L. Ryzhik, B. Perthame. Wave-like solutions for nonlocal reaction-diffusion equations: a toy model. Math. Model. Nat.Phenom., 8 (2013), no. 3, 3341.
[35] Novitsky V, Cao H, Rybak N, Gilbert P, McLane MF, Gaolekwe S, Peter T, Thior I, Ndung'u T, Marlink R, Lee TH, Essex M. Magnitude and frequency of cytotoxic T-lymphocyte responses: identification of immunodominant regions of human immunodeficiency virus type 1 subtype C. J Virol. 2002 Oct;76(20):10155-68.
[36] Martin Nowak and Robert M. May. Virus Dynamics. Mathematical Principles of Immunology and Virology. 2000. Oxford University Press.
[37] B. Perthame, S. Genieys. Concentration in the nonlocal Fisher equation: the Hamilton-Jacobi limit. Math. Model. Nat. Phenom., 4 (2007), 135151.
[38] L. Rong, Z. Feng, A.S. Perelson. Emergence of HIV-1 drug resistance during antiretroviral treatment. Bulletin of Mathematical Biology, (2007) 69: 2027-2060.
[39] L. Rong, R.M. Ribeiro, A.S. Perelson. Modeling quasispecies and drug resistance in Hepatitis C patients treated with a protease inhibitor. Bulletin of Mathematical Biology, (2012) 74:1789-1817.
[40] Rotem A, Serohijos AWR, Chang CB, Wolfe JT, Fischer AE, Mehoke TS, Zhang H, Tao Y, Lloyd Ung W, Choi JM, Rodrigues JV, Kolawole AO, Koehler SA, Wu S, Thielen PM, Cui N, Demirev PA, Giacobbi NS, Julian TR, Schwab K, Lin JS, Smith TJ, Pipas JM, Wobus CE, Feldman AB, Weitz DA, Shakhnovich EI. Evolution on the Biophysical Fitness Landscape of an RNA Virus. Mol Biol Evol. 2018 Oct 1;35(10):2390-2400.
[41] Rouzine IM, Rozhnova G. Antigenic evolution of viruses in host populations. PLoS Pathog. 2018 Sep 12;14(9):e1007291
[42] Rouzine IM, Rodrigo A, Coffin JM. Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. Microbiol Mol Biol Rev. 2001 Mar;65(1):151-85
[43] Rouzine IM, Wakeley J, Coffin JM. The solitary wave of asexual evolution.roc Natl Acad Sci U S A. 2003 Jan 21;100(2):587-92.
[44] Sasaki A. Evolution of antigen drift/switching: continuously evading pathogens. J Theor Biol. 1994 Jun 7;168(3):291-308.

[45] Schelker M, Mair CM, Jolmes F, Welke R-W, Klipp E, Herrmann A, et al. (2016) Viral RNA Degradation and Diffusion Act as a Bottleneck for the Influenza A Virus Infection Efficiency. PLoS Comput Biol 12(10): e1005075.
[46] Schlesinger KJ, Stromberg SP, Carlson JM. Coevolutionary immune system dynamics driving pathogen speciation. PLoS One. 2014 Jul 23;9(7):e102821.
[47] Timofeeva TA, Asatryan MN, Altstein AD, Narodisky BS, Gintsburg AL, Kaverin NV. Predicting the Evolutionary Variability of the Influenza A Virus. Acta Naturae. 2017;9(3):48-54.
[48] T. M. Touaoula, M. N. Frioui, N. Bessonov, V. Volpert. Dynamics of solutions of a reaction-diffusion equation with delayed inhibition. DCDS-S, 13 (2020), no. 9, 2425-2442.
[49] S. Trofimchuk, V. Volpert. Traveling waves for a bistable reaction-diffusion equation with delay. SIAM Journal on Mathematical Analysis 50 (2018), (1), 1175-1199.
[50] Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature. 2006 Jan 19;439(7074):344-8
[51] Vignuzzi M, Wendt E, Andino R.Engineering attenuated virus vaccines by controlling replication fidelity. Nat Med. 2008 Feb;14(2):154-61
[52] de Visser JA, Krug J. Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 2014 Jul;15(7):480-90.
[53] V. Volpert. Elliptic partial differential equations. Volume 2. Reaction-diffusion equations. Birkhäuser, 2014.
[54] V. Volpert. Pulses and waves for a bistable nonlocal reaction-diffusion equation. Applied Mathematics Letters, 44 (2015), 21-25.
[55] A.I. Volpert, V.A. Volpert. Applications of the rotation theory of vector fields to the study of wave solutions of parabolic equations. Trans. Moscow Math. Soc., 1990, 52, pp. 59-108.
[56] V. Volpert, A. Volpert. Spectrum of elliptic operators and stability of travelling waves. Asymptotic Analysis, 2000, **23**, pp. 111-134.
[57] Zanini F, Puller V, Brodin J, Albert J, Neher RA. In vivo mutation rates and the landscape of fitness costs of HIV-1. Virus Evol. 2017 Mar 2;3(1):vex003.

6. Supplementary Materials. Virus distribution in the space of genotypes. Consider a class of genomes consisting of n loci, each of them admitting two states, normal state denoted by A and mutated state denoted by B. Then it can be considered as a sequence of n letters A and B. We will use notation $A_{n-k}B_k$ for all sequences with k letters B and $n-k$ letters A. Characterizing an individual virus in some population by its genome, that is, by a single sequence, we introduce the number of individuals u_k having genome $A_{n-k}B_k$, $k = 0, ..., n$.

A single mutation changing one letter A to B converts a sequence from $A_{n-k}B_k$ to a sequence from $A_{n-k-1}B_{k+1}$. This mutation can occur for any of $(n-k)$ letters A of this sequence. Assuming that the probability μ of this mutation does not depend on the position of the letter A (independent mutations), we conclude that the total probability of such event is $\mu(n-k)$. Hence, the flux from u_k to u_{k+1} equals $\mu(n-k)u_k$. Similarly, we determine three other fluxes:

$$u_k \to u_{k-1} : \mu k u_k, \quad u_{k+1} \to u_k : \mu(k+1)u_{k+1}, \quad u_{k-1} \to u_k : \mu(n-k+1)u_{k-1}.$$

We can now write the equation for the number of individuals u_k:

$$\frac{1}{\mu}\frac{du_k}{dt} = -(n-k)u_k - ku_k + (k+1)u_{k+1} + (n-k+1)u_{k-1}.$$

Setting $L = n\delta, x_{k\pm1} = (k \pm 1)\delta$, $D = \mu n \delta^2$, where δ is a positive number, we write the last equation as follows:

$$\frac{du_k}{dt} = D\frac{u_{k+1} - 2u_k + u_{k-1}}{\delta^2} - \mu L\frac{u_{k+1} - u_k}{\delta} + 2\mu\frac{x_{k+1}u_{k+1} - x_{k-1}u_{k-1}}{2\delta}.$$

This equation is a finite difference approximation of the equation

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + \mu\frac{\partial((2x-L)u)}{\partial x}\,, \tag{6.1}$$

where $u_k(t) = u(x_k, t)$, x_k corresponds to the sequence $A_{n-k}B_k$, $k = 0, ..., n$. Equation (6.1) describes the population density distribution $u(x,t)$ as a function of genotype x considered as a continuous variable and of time t. It is considered on the interval $0 \le x \le L$ with some initial and boundary conditions. Let us note that the convective term in this equation changes sign in the middle of the interval $x = L/2$. Parameters D and μ can be considered as independent.

In the case of equal boundary conditions, $u(0,t) = u(L,t) = u_0$, equation (6.1) has a stationary solution $u_s(x)$: $u_s(x) = u_0 e^{ax(L-x)} = be^{-a(x-L/2)^2}$, where $a = \mu/D$, $b = u_0 e^{aL^2/4}$. It is a normal distribution with the maximum at the center of the interval. It corresponds to the distribution of random sequences. If the boundary conditions are not equal to each other, then this distribution will be shifted with respect to the center of the interval. Let us note the variable x corresponds here to some set of genotypes with the same number of mutations and not to a single genotype. Therefore, $u(x,t)$ describes not a single strain but some set of strains. Such localized solutions of equation (6.1) exist due to the convective term, and they do not exist for the diffusion equation describing chain mutations.

Data on HIV mutations show that mutations can be dependent in the sense that mutation at some locus k_i can affect the probability of mutation at some other locus k_j [26]. Denote by k_i, $i = 1, 2, ...$ a sequence of consecutive mutations. Assuming that

they are reversible and have the same probability μ, we obtain the following equation for the densities u_{k_i}:

$$\frac{du_{k_i}}{dt} = \mu(u_{k_{i-1}} - 2u_{k_i} + u_{k_{i+1}}).$$

Instead of equation (6.1), we obtain in this case the diffusion equation. Such model of virus mutation were previously considered (see the main part of the paper for the references). Some mutations can affect the probability of several other mutations leading to the branching chain mutations where the equation for the virus density population should be considered on a system of interconnected intervals of genotypes. This question will be considered in the subsequent works.