

Article

Deep Gene Networks and Response to Stress

Sergey Vakulenko ^{1,*} and Dmitry Grigoriev ^{2,†}¹ Saint Petersburg Electrotechnical University “LETI”, 197376 Saint Petersburg, Russia² CNRS, Mathématiques, Université de Lille, 59655 Villeneuve d’Ascq, France; dima.grigoryev@gmail.com

* Correspondence: savakulenko@etu.ru or vakulenfr@gmail.com; Tel.: +7-921-584-2914

† These authors contributed equally to this work.

Abstract: We consider systems of differential equations with polynomial and rational nonlinearities and with a dependence on a discrete parameter. Such systems arise in biological and ecological applications, where the discrete parameter can be interpreted as a genetic code. The genetic code defines system responses to external perturbations. We suppose that these responses are defined by deep networks. We investigate the stability of attractors of our systems under sequences of perturbations (for example, stresses induced by environmental changes), and we introduce a new concept of biosystem stability via gene regulation. We show that if the gene regulation is absent, then biosystems sooner or later collapse under fluctuations. By a genetic regulation, one can provide attractor stability for large times. Therefore, in the framework of our model, we prove the Gromov–Carbone hypothesis that evolution by replication makes biosystems robust against random fluctuations. We apply these results to a model of cancer immune therapy.

Keywords: stability; evolution; response to stress; deep networks



Citation: Vakulenko, S.; Grigoriev, D. Deep Gene Networks and Response to Stress. *Mathematics* **2021**, *9*, 3028. <https://doi.org/10.3390/math9233028>

Academic Editors: Ovanes Petrosian and Witold Pedrycz

Received: 13 October 2021

Accepted: 23 November 2021

Published: 26 November 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In this paper, we consider systems of differential equations with polynomial and rational nonlinearities and with a dependence on a discrete parameter (hybrid systems). Such systems arise in biological and ecological applications, where the discrete parameter allows us to incorporate in the model a genetic regulation by a genetic code. We investigate the stability of attractors of these systems under stress. The attractors define system functioning regimes, while stresses are sudden changes of system parameters (for example, heat shocks). Our aim is to investigate the attractor stability under infinite sequences of shocks. It is inspired by the following idea of M. Gromov and A. Carbone: “Homeostasis of an individual cell cannot be stable for a long time as it would be destroyed by random fluctuations within and out of cell. There is no adequate mathematical formalism to express the intuitively clear idea of replicative stability of dynamical systems” ([1], p. 40). It is clear that replicative stability is important for the evolution of cancer cells, viruses, for example, such as COVID 19, and bacteria such as *E. coli*. There is no doubt that replication helps viruses and bacteria survive in changing environments (see, for example, Lensky’s experiment with *E. coli* [2]). However, there is a non-trivial question: how to estimate chances that such replicative evolution can continue eternally even under action of therapeutic agents? For COVID-19, it is an important question of will we have fourth or fifth pandemic waves, etc., or can the virus vanish? In this paper, however, we consider cancer cells as an example because our analyses are more applicable to this case.

Using a model, proposed in this paper, one can formulate the Gromov–Carbone idea in rigorous mathematical terms and, under some assumptions, to prove this hypothesis. Our model exploits ideas of artificial intelligence theory; namely, we suppose that the response to stress is defined by deep gene networks (DGN), which are similar to deep neural networks (DNNs). It allows us to use estimates obtained recently in DNN theory [3–5]. Note that deep networks were recently applied as interpretable models of the gene network regulations in [6].

To describe replicative stability, we incorporate into our model a regulation by the discrete parameter (for biosystems and ecosystems it is a gene regulation) and consider how the stochastic stability depends on regulation. Our results help us to explain, for example, why cancer cells exhibit a long survival time even under a strong therapy. Indeed, cancer cells give us an example of evolving biosystem affected by a sequence of stresses (for example, immune therapy or chemotherapy). It is worth noting that replicative stability for biochemical systems, introduced in this paper, is a new idea for the theory of stability (robustness) of these systems. Earlier, a number of works (see fundamental papers [7,8] and references therein) considered a more usual approach, which investigates reactions of the biochemical system to the small disturbances of kinetic rates or reagent concentrations. We also investigate the stability in the vicinity of small changes in some states (attractors). However, we take into account variations of the system genotype. We suppose that these variations compensate for a significant part of external perturbations that reinforces the system stability. The gene regulation system evolves in such a way that this compensation mechanism becomes more and more effective, which provides stability under fluctuations for large times.

Let us first outline the concept of stochastic stability. We consider organisms as biochemical machines, which can be modeled by a system of ordinary differential equations. A regime of our biosystem functioning is defined by an attractor \mathcal{A} . Let us denote by $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ the probability that the system state lies in a δ -neighborhood of the attractor \mathcal{A} within the time interval $[0, \tau]$. The parameter $\delta > 0$ defines the size of a homeostasis domain, where the system stays viable (here we follow viability theory, see, for example, [9]). In the framework of that viability approach, the Gromov–Carbone hypothesis reduces to two claims. The first assertion is that for biological systems with fixed parameters and a fixed genetic code,

$$\mathbb{P}_{\delta, \mathcal{A}, \tau} \rightarrow 0 \quad \text{as } \tau \rightarrow +\infty. \quad (1)$$

This means that all biosystem states, sooner or later, will be destroyed by fluctuations. The second assertion is that a sequence of modifications of the system genotype can lead to eternal stability, where the viability probability stays positive for all times:

$$\mathbb{P}_{\delta, \mathcal{A}, \tau} > \beta > 0 \quad \forall \tau > 0, \quad (2)$$

where $\beta > 0$. The constant β can be interpreted as follows: imagine a large population of cancer cells, bacteria or viruses. Let this population be affected by consecutive therapeutic attacks. The constant β describes the relative fraction of cells (bacteria, viruses) that will survive.

In other words, systems with fixed parameters are unstable under environment fluctuations but the evolution of the genetic code can stabilize them. This replicative stability is considered in [10] for some network models; see [11] for an overview. For some classes of hybrid dynamical systems involving a dependence on discrete genotypes $s(t)$, which can evolve in time, it is shown that inequality (2) implies

$$\sup_{t \in [0, T]} L(s(t)) \rightarrow +\infty \quad T \rightarrow \infty, \quad (3)$$

i.e., the genotype length L has a tendency to increase. Here we assume that the genotype s is a Boolean string (s_1, \dots, s_N) and the length L of s is N . However, it is well known that the genotype length does not correlate directly with the morphological complexity and that evolution does not always lead to a genotype length increase. For example, *Drosophila*, nematode and human have numbers of coding genes of the same order (although bacteria genomes are much shorter than the nematode genome). Nonetheless, it is well known that the gene regulation system has become more and more complex throughout evolution. In this paper, we use deep networks (DNs) to describe gene regulation and show that DN network growth makes biosystems stable for long time periods.

Let us outline the model in more detail. In contrast to some earlier works [12–14], which used either formidable computations [12,13], or sophisticated nonlinear asymptotics [14], we use deep networks. These networks consist of, in general, many layers. The depth $L = L_{net}$ of a network is the number of hidden layers, and the size $S = S_{net}$ is the total number of units. In practice, networks with depth $L_{net} = O(1)$ can be considered as shallow, while deep networks typically have many layers. Connections between units in these networks depend on coding sequences $s = (s_1, \dots, s_N)$, $s_i \in \{0, 1\}$ of Boolean genes, which could be switched off ($s_i = 0$) or switched on ($s_i = 1$). The inputs of the DN's are the stress parameter values and the system states. The DN produces an output, which depends on unit interconnections and thus on the genetic code, which determines these interconnections. The main idea behind our model of regulation is that the DN output and the stress impact cancel out each other.

To describe stress impacts, we use a model of subsequent environmental shocks. This idea is consistent with the Gould–Eldredge concept [15], where evolution is represented as a sequence of intermittent short bursts and longtime stasis periods, the evolution bursts appear in response to stresses induced by environment. Let $\tau_1, \tau_2, \dots, \tau_k, \dots$ be an infinite increasing sequence of time moments τ_j . These moments can be considered as moments of environment shocks (stresses), when the environment suddenly changes. At these moments, some system parameters can change. A well-known example is a heat shock connected with temperature, which affects kinetic rates in biochemical systems and produces mutations [16]. For cancer cells, these shocks can describe a therapy. Our aim is to estimate the probability that a local attractor \mathcal{A} is robust under all those environmental shocks within a large (maybe, infinite) time interval.

Let us outline our results. The first result (Theorem 1) concerns a connection between gene regulation complexity and uncertainty of environment. The important measure of genetic regulation complexity is the number of equilibria of gene regulation network N_{reg} . This question on the number of the equilibria has received the attention of many works; in particular, it was studied for the famous NK model (pioneered in [17]). The NK model is a mathematical model describing a “tunably rugged” fitness landscape, depending on two parameters, N and K . The parameter N is the length of a Boolean genotype s length, and K defines the topology of fitness landscape via the number of genes involved in the trait control. The NK model has the main applications in evolutionary biology. First, it was supposed that N_{reg} is a relatively small, $N_{reg} \propto N^\kappa$, where an exponent $\kappa > 0$. In particular, for the Hopfield model with symmetrical interaction [18], this relation holds with $\kappa = 1$. However, later it was shown that in the NK model, N_{reg} can be exponentially large [19]. A similar result is obtained for another simple model, which describes a special class of Hopfield models with asymmetric interaction and a special interaction topology [20]. Experimental data [21,22] show that in gene interaction networks there are strongly connected nodes that can be named hubs, or centers. The hubs communications go via a number of weakly connected nodes. This centralized connectivity has been found in gene–gene and protein–protein interaction networks [21,22], and such networks can support many equilibrium states.

To explain why they have a number of equilibria, we define an entropy $H_{ext,\delta}$ of environment uncertainty, depending on the robustness level δ and gene regulation. This quantity is the Kolmogorov δ -entropy (see, for example, [23,24]) on the metric space of perturbations ξ with a special metric. We show that if an attractor describing biosystem functioning does not collapse under all perturbations (conserves its topological structure) then there exists a connection between N_{reg} and the entropy $H_{ext,\delta}$:

$$\log_2 N_{reg} \geq H_{ext,\delta}. \quad (4)$$

This means that entropy of regulation must be no less than the environment uncertainty.

This result, in particular, shows that, to kill cancer cells, we should use a combined therapy (for example, chemotherapy together with a heat therapy). Hyperthermia therapy is heat treatment for cancer that can be a powerful tool when used in combination with

chemotherapy or radiation. Chemotherapy resistance develops over time as the tumors adapt (due to replicative stability). Chemoresistance against several anticancer drugs could be reversed by the addition of heat therapy.

Our results show that before all, it is necessary to attack hubs of the gene regulation network. In fact, N_{reg} critically depends on the gene interaction topology. Even a small increase in the number of hubs in gene regulation networks leads to a sharp increase in N_{reg} [20]. These results are consistent with the conclusions of [25,26], where it was shown that network hubs buffer environmental variations and also make switches in regulatory networks. Our results are supported by some experimental data, at least qualitatively [27]. A number of hub genes are identified in cancer cell regulation. According to these results, drugs attacking the hub genes should be especially effective. A list of such candidate drugs targeting hub genes can be found in Table 3 from [27].

There is also a connection with basic results of hard combinatorial problem theory. Fundamental results [28] on Boolean functions and the K -SAT problem assert that to satisfy many constraints, we should use a number of Boolean variables. Our result (4) can be considered as an analog of these results. If we have a genotype s of the length N , then $\log_2 N_{reg} \leq N$; thus, for Boolean regulation models, (4) implies that the genome length N should be more than the entropy $H_{ext,\delta}$.

Furthermore, we show that it is possible to provide stochastic attractor robustness by an appropriate genotype choice at each environment shock $\tau_1 < \tau_2 < \dots$. If, in contrast, the genotype is fixed, then the stochastic stability tends to zero as $\tau \rightarrow +\infty$. For cancer these results are confirmed by numerical simulations, where we use a model from [29] extended to take into account the evolution of cancer cells. If the cancer cells do not evolve, immune therapy may be effective, otherwise cancer wins sooner or later.

To conclude this introduction, let us note that the estimates of stochastic stability, obtained in the paper (see Theorem 3), allows us to find the main factors, which determine the organism fate under the stress stability: the sparsity of reaction network, the stress sensitivity, the stress dimension (the number of perturbed external parameters affecting the rates of biochemical reactions) and the size of gene regulation network.

The paper is organized as follows. In the next section, we describe models with random perturbations and gene regulation. In Section 4, we state the model of environmental shocks and formulate some assumptions needed for the mathematical tractability of the problem. In Section 5, we formulate some definitions, and further, we find a necessary condition for the attractor stability.

In Section 7, we consider attractor robustness under a sequence of environment shocks. In this section, we apply DNN's that finally give Theorem 3, which states an estimate of replicative stability. The last section contains a discussion and conclusions. To simplify reading, the technical parts of demonstrations are relegated into Appendix A.

2. Materials and Methods

We use standard methods of mathematical analysis and theory of differential equations. Moreover, we apply the known results for approximations of Lipschitz smooth functions by deep neural networks [3–5]. For numerical simulations, and to produce Figure 1, we used Matlab and the standard Euler method.

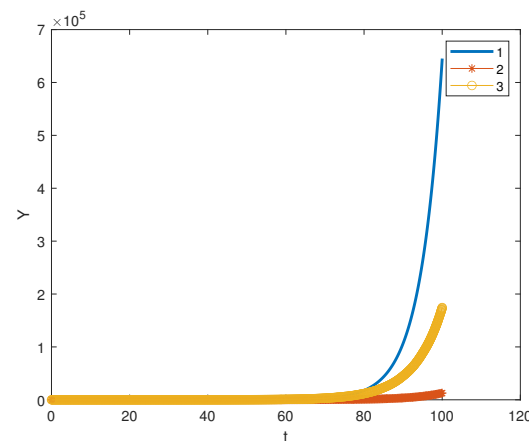


Figure 1. This plot shows dynamics of concentration of tumor cancer cells in three cases: no therapy (the curve 1), therapy in the case when tumor cells do not evolve (the curve 3) and therapy in the case when tumor cells mutate with the probability $p_{mut} = 0.001$ (the curve 2). The parameters are taken from [30]: $r_0 = 1.310^4$, $\rho = 0.1254$, $\alpha = 2.02 \cdot 10^{-7}$, $c_1 = 3.42 \cdot 10^{-10}$, $c_2 = 1.1 \cdot 10^{-7}$, $d_1 = 0.0412$, $b = 2.010^{-9}$, $r = 0.18$. We have a suite of 20 shocks, which at the time moments $t = 5, 10, \dots, 100$, $\Delta t = 5$. For the increasing rate, the tumor cell rate increment is $d_r = 0.01$. For the decreasing and increasing curve, the infusion dose is $\eta = 2 \cdot 10^8$.

3. Model with Random Parameters

We describe our models in two steps, first we formulate a model without gene regulations and then with regulation.

We consider systems of differential equations with a dependence on a random process $\zeta(t)$. They have the form

$$\frac{dv}{dt} = f(v, \zeta(t)) \quad t \geq 0, \tag{5}$$

where $v(t) = (v_1(t), \dots, v_n(t))^{tr} \in D$, where D is a compact domain in \mathbb{R}^n with a smooth boundary ∂D , $v_i(t)$ are biochemical species concentrations, $f(v, \zeta) = (f_1(v, \zeta), \dots, f_n(v, \zeta))^{tr}$, the reaction terms f_i are sufficiently smooth functions of v , for example, multivariate polynomials in v , and also smooth functions of $\zeta \in \mathbb{R}^p$. The functions $\zeta(t)$ are random processes of a trajectory, which are piecewise constant in t and take values in a compact subset \mathcal{P}_{ext} of \mathbb{R}^p , where the quantity p is a positive integer (dimension of stress parameter ζ). These processes $\zeta(t)$ describe multiplicative noise. The source of this noise is an environment of a biosystem. Other assumptions of $\zeta(t)$ will be described later.

We complement the system of stochastic Equation (5) by initial conditions

$$v(0) = v^{(0)}. \tag{6}$$

To define f we use the standard models of chemical kinetics and population dynamics. The simplest choice is a polynomial model, which can be derived by the law of mass action:

$$f_i(v, \zeta) = \sum_{\mathbf{a} \in \mathcal{R}_i} C_{i,\mathbf{a}}(\zeta) v_1^{a_1} v_2^{a_2} \dots v_n^{a_n}, \tag{7}$$

where $\mathbf{a} = (a_1, \dots, a_n)$ is a multi-index with integers $a_i \geq 0$, \mathcal{R}_i are finite subsets of $I_n = \{1, \dots, n\}$ and $C_{i,\mathbf{a}}$ are coefficients that determine kinetic rates. Note that the well-known Lotka–Volterra and generalized Lotka–Volterra systems (which have ecological and economical applications, see [31,32]) are also included in the class of systems defined by (7).

In real biochemical applications, we are often dealing with binary reactions, where only two exponents $a_i \neq 0$, and sparse f_i , where, although the reagent number n may be large, for each i , the number m_i of non-zero coefficients $C_{i,\mathbf{a}}$ in each row is bounded,

$m_i \ll n$. In a more sophisticated case (which arises when in a slow-fast polynomial system we express fast variables via slow ones) we have rational v nonlinearities, for example,

$$f_i(v, \xi) = \sum_{\mathbf{a} \in \mathcal{R}_i} C_{i,\mathbf{a}}(\xi) \frac{P_{i,\mathbf{a}}(v)}{Q_{i,\mathbf{a}}(v)}, \tag{8}$$

where $P_{i,\mathbf{a}}(v)$ and $Q_{i,\mathbf{a}}(v)$ are polynomials. We suppose that $Q_{i,\mathbf{a}}(v)$ are separated from zero:

$$|Q_{i,\mathbf{a}}(v)| > \delta_0 > 0, \tag{9}$$

for all i, \mathbf{a} and $v \in D$. Such systems appear in biochemical kinetics, population dynamics and cancer modeling. As an example, let us consider the following model of cancer immunotherapy [29,30]:

$$\frac{dv_1}{dt} = r_0 + \frac{\rho v_1 v_2}{\alpha + v_1} - c_1 v_1 v_2 - d_1 v_1 + \xi(t), \tag{10}$$

$$\frac{dv_2}{dt} = r v_2 (1 - b v_2) - c_2 v_1 v_2, \tag{11}$$

where v_1 denotes the concentration of cells with antitumor activity in the tumor site (CTL cells), and v_2 is the concentration of tumor cells. This model thus concerns two cell populations: normal cells and cancer ones. Parameters $r_0, c_1, c_2, b, r, d_1, \alpha, \rho$ are positive coefficients, which can be explained as follows. The parameter r_0 is the normal rate of immune cell flow into the tumor site, c_1 is the immune cells death rate due to interaction with tumor cells, c_2 is the fraction of tumor cells killed by immune cells, $1/b$ is the tumor cells carrying capacity, r is the tumor cells growth rate, d_1 is a natural death rate of immune cells, α is the steepness coefficient of immune cell recruitment and ρ is the maximal immune cells recruitment rate [29]. For $\xi(t) \equiv 0$ this model and its generalizations are investigated in [30], the term $\xi(t)$ describes a therapy by infusion of CTL cells (it is proposed in [29]).

To provide the existence and uniqueness of solutions to the Cauchy problem (5) and (6) for all $t \in (-\infty, \infty)$, we assume the following. Let

$$f(v, \xi) \cdot n(v) \leq 0 \quad \forall v \in \partial D \quad \forall \xi \in \mathcal{P}_{ext} \tag{12}$$

where $n(v)$ is a unit normal vector directed outside D at the point $v \in \partial D$. Then solutions of the Cauchy problem (5) and (6) exist and are unique for all $t > 0$.

It is worth noting that polynomial and rational functions f_i may not satisfy conditions (12). Since we further investigate a local stability of local attractors, we can restrict ourselves to considering narrow neighborhoods of those attractors \mathcal{A} . Therefore, we can truncate nonlinearities in $f_i(v)$ setting for instance $\hat{f}_i(v) = f_i(v)\chi_{\mathcal{A}}(v)$, where $\chi_{\mathcal{A}}(v) \equiv 0$ outside of an open neighborhood W of \mathcal{A} .

3.1. Extended Model with Regulation

Different complicated models of gene regulation were proposed, for example, [33,34] for *Drosophila* morphogenesis, however, we apply another approach using DNN's and inspired by recent biological works (for example, [6]).

Let $s(t) = (s_1(t), \dots, s_N(t))$ be a time dependent gene expression vector. We have N Boolean genes s_i , which take values in the set $S_N = \{0, 1\}^N, s_i \in \{0, 1\}$. The i -th gene may be switch on (when $s_i = 1$), or switch off (when $s_i = 0$). The Boolean string s encodes interconnections in the networks of gene regulation (see Section 7.2). We incorporate genes s into system (5) using the following

Assumption GR Suppose that coefficients $C_{i,\mathbf{a}}$ depend on the genetic code s :

$$C_{i,\mathbf{a}} = C_{i,\mathbf{a}}(\xi, s) \tag{13}$$

and they are Lipschitz in ξ for each s .

This assumption means that regulation goes via a genetic control of kinetic rates. Moreover, we assume that the network finds a genotype s , which produces an optimal answer to stress, almost instantly. We obtain thus that s is involved in the vector field $f(v, \zeta)$ via kinetic coefficients $C_{i,a}$: $f = f(v, \zeta, s)$.

Let us introduce vector fields describing perturbations. Consider, for example, a heat shock. All kinetic rates depend on the temperature T : $C_{i,a} = C_{i,a}(T)$, sometimes in a complicated manner. However, typically we have a reference value $T = \bar{T}$ corresponding to an optimal functioning regime (for example, if T is the human body temperature, then \bar{T} is about 36 degrees). Let $T = \bar{T} + \Delta T$, where ΔT is a temperature increment corresponding to a heat shock. Then in the case (7) one obtains that the perturbations g_i of the f_i have the form

$$g_i(v, T, s) = \sum_{a \in \mathcal{R}_i} \tilde{C}_{i,a,s} v_1^{a_1} v_2^{a_2} \dots v_n^{a_n}, \tag{14}$$

where

$$\tilde{C}_{i,a,s} = C_{i,a}(\bar{T} + \Delta T, s) - C_{i,a}(\bar{T}, s).$$

In the general case, let $\bar{\zeta}$ be an optimal parameter value corresponding to a normal system functioning. Then we denote by

$$g(v, \zeta, s) = f(v, \zeta, s) - f(v, \bar{\zeta}) \tag{15}$$

the perturbation of f , which appears as a result of a parameter ζ variation. Let us denote the non-perturbed field $f(v, \bar{\zeta})$ by $\bar{f}(v)$. Then we have the two following systems: the non-perturbed one

$$\frac{dv}{dt} = \bar{f}(v), \tag{16}$$

and the perturbed system:

$$\frac{dv}{dt} = \bar{f}(v) + g(v, \zeta, s) \tag{17}$$

where, in general case, ζ and s can depend on t .

3.2. Attractors and Semiflows

Let us denote by $\mathcal{A}_{s,\zeta}$ a compact invariant locally attracting set (a local attractor) of semiflow $S_{s,\zeta}^t$ generated by the Cauchy problem for system (17) for given fixed s and ζ . The semiflow generated by the non-perturbed system (16) will be denoted by S^t and the corresponding attractors by \mathcal{A} . We consider mainly two cases: \mathcal{A} is a stable hyperbolic equilibrium v^{eq} or a stable hyperbolic limit cycle $V(t)$ with a period T_0 . (A reminder that hyperbolicity for an equilibrium means that the spectrum of the corresponding linearized operator does not intersect an imaginary axis, while for cycles it means that the corresponding Floquet multipliers do not lie on the unit circle of the complex plane. For the general theory of hyperbolic sets, see [35,36]). We are going to estimate the probability that a local attractor (or an invariant set) of our system is stable under those shocks. We define that probability in the next subsection.

4. Stochastic Stability under Shocks

4.1. Environmental Shocks

To describe stresses, we use the model of subsequent environmental shocks. Let $\tau_1, \tau_2, \dots, \tau_k, \dots$ be an infinite increasing sequence of time moments τ_j such that $\min_j(\tau_{j+1} - \tau_j) \gg \Delta\tau \gg \Delta t$, where Δt and $\Delta\tau$ are characteristic times of gene regulation functioning and biochemical model dynamics, respectively. At the moments τ_j we have environment shocks, when an interaction between the environment and the biosystem changes. We suppose that within interval $[\tau_j, \tau_{j+1})$ the parameter $\zeta = \zeta_j$, where at each step j we choose ζ_j randomly as follows in two steps. The first step is a random choice of the set $\mathcal{P}_{ext,j}$ of perturbations affecting the system at j -th interval $I_j = (\tau_j, \tau_{j+1}]$. We choose the set $\mathcal{P}_{ext,j}$ randomly from a countable set $E = \{\mathcal{P}_{ext,1}, \mathcal{P}_{ext,2}, \dots\}$. The index k in $\mathcal{P}_{ext,k}$ defines a kind

of interaction between the organism and its environment. Each set $\mathcal{P}_{ext,j}$ is equipped with a probabilistic measure μ_j .

The second step is a random choice of $\xi_j \in \mathcal{P}_{ext,j}$. The shocks ξ_j are independent, and at each step j and each k , we sample ξ_j according to the measure μ_j . Therefore, we obtain a sequence of independent random variables $\xi_j, j = 1, \dots, +\infty$, which describes subsequent random changes of our stress environment.

4.2. Attractor Stochastic Stability

To formulate the stochastic stability problem, we use the viability approach [9]. Let \mathcal{A} be an attractor under consideration. For $\delta > 0$ we introduce a small tubular neighborhood of that attractor by

$$W_\delta(\mathcal{A}) = \{v : \text{dist}(v, \mathcal{A}) < \delta\}, \tag{18}$$

where the distance between a state v and a subset $B \subset D$ is defined as

$$\text{dist}(v, B) = \inf_{w \in B} |v - w|, \tag{19}$$

where $|v|$ is the standard Euclidean norm in \mathbb{R}^n .

Let us introduce probability that the state of our system lies in $W_\delta(\mathcal{A})$ within the time interval $[0, \tau]$:

$$\mathbb{P}_{\delta, \mathcal{A}, \tau} = \text{Prob}\{v(t) \in W_\delta(\mathcal{A}) \quad \forall t \in [0, \tau]\}. \tag{20}$$

Naturally we suppose that the initial value $v^{(0)} \in W_\delta(\mathcal{A})$. The probability $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ depends on that initial value. Since our aim to obtain estimates of $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ uniform in $v^{(0)} \in \mathcal{A}$ we will omit dependence on $v^{(0)}$. The probability $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ can be considered as a measure of attractor \mathcal{A} stochastic stability.

5. Robustness with Gene Regulation

5.1. Definitions

Let us formulate certain definitions. We consider ergodic attractors \mathcal{A} . The statement of ergodic attractor's theory can be found in [37]. For such attractors and sets, the time average can be replaced by averages over the attractor taken by an invariant measure ρ defined on the attractor. Let us introduce the time averages along trajectories

$$\langle \phi \rangle_{\mathcal{A}} = \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \phi(v(t)) dt \tag{21}$$

where $\phi(v)$ is a smooth function, and $v(t)$ is a trajectory on the attractor. For ergodic attractors these averages exist and do not depend on the choice of the trajectory $v(t, v(0))$ on the attractor for almost all $v(0)$ with respect to the measure ρ . Of course rest point and limit cycle attractors are ergodic, but theory [37] also concerns the non-trivial situation, where \mathcal{A} may be chaotic.

Let us denote by Φ the set of all smooth functions ϕ defined on D . We denote by $Lip(\phi)$ the Lipschitz constant of ϕ :

$$|\phi(v) - \phi(\bar{v})| \leq Lip(\phi)|v - \bar{v}|.$$

Definition 1. We say that a local attractor \mathcal{A} of the semiflow S^t defined by Equation (17) with $\zeta = \bar{\zeta}$ is ϵ -robust with respect to the stress parameter ζ and the genotype s if

$$|\langle \phi \rangle_{\mathcal{A}} - \langle \phi \rangle_{\mathcal{A}_s, \bar{\zeta}}| < Lip(\phi)\epsilon, \quad \forall \phi \in \Phi, \tag{22}$$

i.e., the time averages over the perturbed and unperturbed attractors are ϵ -close.

Note that if the attractor is hyperbolic and $\epsilon > 0$ is small, then due to the persistence of hyperbolic sets [35,36], the perturbed attractor is topologically equivalent to the non-perturbed one.

Definition 2. A local attractor \mathcal{A} of the non-perturbed flow S^t defined by system (16) is ϵ -robust with gene regulation with respect to the stress environment \mathcal{P}_{ext} if for each $\zeta \in \mathcal{P}_{ext}$ there is an appropriate genotype $\bar{s}(\zeta)$ such that

$$|\langle \phi \rangle_{\mathcal{A}} - \langle \phi \rangle_{\mathcal{A}_{\bar{s}, \zeta}}| < Lip(\phi)\epsilon, \quad \forall \phi \in \Phi \tag{23}$$

where $\mathcal{A}_{\bar{s}, \zeta}$ is the local attractor $\mathcal{A}_{\bar{s}, \zeta}$ of the semiflow $S^t_{\bar{s}, \zeta}$.

Due to the definition of robustness under gene regulation for each field $\zeta \in \mathcal{P}_{ext}$ we have a genotype $\bar{s}(\zeta)$ such that trajectory averages over the perturbed attractor $\mathcal{A}_{\bar{s}, \zeta}$ are ϵ -close to ones for non-perturbed attractor \mathcal{A} of system (16). We have thus the map $S : \zeta \rightarrow s(\zeta) \in S^N$ from the set \mathcal{P}_{ext} into the set of all genotypes S^N .

5.2. Robustness Condition

Let us formulate a claim.

Proposition 1. Let a local attractor \mathcal{A} of a non-perturbed system (16) be ergodic and ϵ -robust with respect to $\zeta \in \mathcal{P}_{ext}$ and $s \in S^N$. Then

$$\langle g(v, \zeta, s) \rangle_{\mathcal{A}} < C_1\epsilon, \tag{24}$$

where the constant $C_1 > 0$ is uniform in $\epsilon > 0$.

Proof. We take the averages of the right and left hand sides of Equations (16) and (17) that gives

$$\langle \bar{f}_i \rangle_{\mathcal{A}} = 0, \quad \langle \bar{f}_i + g_i \rangle_{\mathcal{A}_{s, \zeta}} = 0 \tag{25}$$

for all i . By Definition 1 of robustness one has

$$\left| \langle \bar{f}_i + g_i \rangle_{\mathcal{A}} - \langle \bar{f}_i + g_i \rangle_{\mathcal{A}_{s, \zeta}} \right| < c_0\epsilon,$$

where $c_0 > 0$ are uniform in $\epsilon > 0$. Then (25) and the last estimate give (24). \square

6. Gene Regulation and Robustness

In this section, we describe a connection between robustness and the gene regulation.

Environment Uncertainty

In this subsection, we introduce a measure of external perturbation uncertainty. Let $\delta > 0$ be a number. Consider the set \mathcal{P}_{ext} of perturbations ζ . Let us introduce a special distance between perturbations depending on the attractor and g :

$$dist_{\mathcal{A}, g}(\zeta, \tilde{\zeta}) = \min_{s \in S^N} \sup_{i=1, \dots, n} |\langle g_i(\cdot, \zeta, s) \rangle_{\mathcal{A}} - \langle g_i(\cdot, \tilde{\zeta}, s) \rangle_{\mathcal{A}}|,$$

where $S^N = \{0, 1\}^N$ denotes the set of all genotypes.

By \mathcal{C}_δ we denote a δ -covering of \mathcal{P}_{ext} consisting of the sets $B_{i, \delta}, i = 1, 2, \dots, N_{cover}$:

$$\mathcal{P}_{ext} = \bigcup_{j=1}^{N_{cover}} B_j, \quad diam(B_j) \leq \delta \quad \forall j,$$

where the diameter $\text{diam}(B)$ of a bounded set B is defined by

$$\text{diam}(B) = \sup_{\tilde{\zeta}, \tilde{\xi} \in B} \text{dist}_{\mathcal{A},g}(\tilde{\zeta}, \tilde{\xi}).$$

The formula for the diameter can be rewritten as

$$\text{diam}(B) = \min_{s \in S^N} \max_{i, \tilde{\zeta}, \tilde{\xi} \in B} |(\tilde{C}_{i,a}(\zeta, s) - \tilde{C}_{i,a}(\tilde{\zeta}, s)) \langle v^a \rangle_{\mathcal{A}}|. \tag{26}$$

We define $H_{ext,\delta}$ as the Kolmogorov δ -entropy of \mathcal{P}_{ext} :

$$H_{ext,\delta} = \inf_{C_\delta} \log_2 N_{cover}, \tag{27}$$

where we take the infimum over all δ -coverings.

Remark 1. The introduced entropy depends on the parameter δ , and decreases in $\delta > 0$. Moreover, $H_{ext,\delta}$ depends on the unperturbed attractor \mathcal{A} (we omit that dependence in our notation).

Remark 2. The Kolmogorov δ -entropy has numerous applications for coding, image processing and neural networks, see [23,24].

It is difficult to compute $H_{ext,\delta}$; however, we can find a rough estimate of this entropy via the dimension $d_{str} = \text{dim} \zeta$ of the stress parameter. Suppose that the diameter of \mathcal{P}_{ext} is bounded: $\text{diam}(\mathcal{P}_{ext}) \leq R_0$, where R_0 is a positive constant, which can be interpreted as a maximal absolute value of the stress parameter variations. Then we obtain

$$H_{ext,\delta} \leq d_{str} \log_2(R_0/\delta). \tag{28}$$

If the set \mathcal{P}_{ext} is not compact then $H_{ext,\delta} = +\infty$.

There holds

Theorem 1. Let an ergodic attractor \mathcal{A} of a non-perturbed system (16) be ϵ -robust with gene regulation. Then for sufficiently small $\epsilon > 0$ one has

$$\log_2 N_{reg} \geq H_{ext,\delta} \tag{29}$$

where $\delta = C\epsilon$, where $C > 0$ is a constant uniform in ϵ as $\epsilon \rightarrow 0$.

For a proof, see Appendix A. The idea is simple: using Proposition 1 one can show that if the same genotype provides the stability for all perturbations from a set B , then $\text{diam}(B) < C_1\epsilon$.

By results of [20] we can find a connection between the number of centers (hubs) in the gene regulation network and the entropy $H_{ext,\delta}$. To find a rough estimate, we suppose that we have a starlike network with N_c centers and each center is connected to N_s satellites. Then the equilibrium number N_{reg} has the order $N_s^{N_c} = \exp(N_c \ln N_s)$ [20] and

$$\log_2 N_{reg} \approx \log_2 \left(\exp(N_c \ln N_s) \right) = c_0 N_c \ln N_s. \tag{30}$$

Equation (30) shows that the main factor determining the effectiveness of the gene response to stresses is the number of hubs involved in the corresponding gene regulation, and $N_c \ln N_s \geq \text{const} H_{ext,\delta}$ (we use here that \ln is a very slow increasing function of its argument). The estimate (28) then gives the following topological condition of the stability:

$$N_c > c_0^{-1} \log_2 R_0 (\ln N_s)^{-1} d_{str}, \tag{31}$$

i.e., roughly speaking, to support stability, the number of the gene network hubs should be proportional to the dimension of the stress parameter.

7. Stability under Many Shocks

In this section, we use the model of gene regulation from Section 4.1. We are going to prove the Gromov–Carbone hypothesis and show that the complexity growth of the gene regulation network can lead to a successful evolution, where attractors stay robust under all shocks. First, in the coming subsection, we study the case when the gene regulation does not work and the genotype s is fixed.

7.1. Instability for Systems with Fixed Parameters

Let us prove a claim. Remember that on the set \mathcal{P}_{ext} of possible perturbations ζ , a probabilistic measure μ is defined.

Theorem 2. *Let v^{eq} be a hyperbolic equilibrium for system (17) with $\zeta = \bar{\zeta}$. Suppose that for each s the range of the map $\zeta \rightarrow g(v^{eq}, \zeta, s)$ contains a closed ball B_R of radius $R > 0$ centered at 0, and the measure μ on \mathcal{P}_{ext} is continuous with respect to the Lebesgue measure on B_R .*

Then for sufficiently small $\delta > 0$

$$\mathbb{P}_{\delta, v^{eq}, \tau} < (1 - \delta_1)^\tau, \quad \tau > 0 \tag{32}$$

where $\delta_1 > 0$.

Proof. Using the theorem assumptions, we choose such a perturbation ζ such that

$$|g(v^{eq}, \zeta, s)| = \delta^{1/2}. \tag{33}$$

If $\delta > 0$ is small enough then by implicit function theorem and using that our equilibrium v^{eq} is hyperbolic, one can show that the equation

$$\tilde{f}(v) + g(v, \zeta, s) = 0$$

has a solution $v^{eq, \zeta}$ close to v^{eq} : $|v^{eq, \zeta} - v^{eq}| < c\delta^{1/2}$. This solution is a hyperbolic equilibrium attractor for system (17). If a trajectory $v(t)$ stays inside the δ -neighborhood of v^{eq} for all $t > 0$, this trajectory converges to $v^{eq, \zeta}$:

$$|v(t) - v^{eq, \zeta}| < C_0 \exp(-c_0 t), \quad C, c_0 > 0.$$

Therefore, $|v^{eq} - v^{eq, \zeta}| < 2\delta$. However, according to Proposition 1, one has

$$|\langle g(v^{eq}, \zeta, s) \rangle_{\mathcal{A}}| < c_1 \delta$$

which creates a contradiction with (33) for sufficiently small $\delta > 0$. Therefore, we have a non-zero probability p_0 to find a ζ such that all trajectories $v(t)$, which start in a δ -neighborhood W_δ of v^{eq} , leave that neighborhood within a sufficiently large interval $[0, T_0]$. Now we use the fact that the parameters ζ_k are i.i.d. random variables. In fact, consider the intervals $[0, T_0], [T_0, 2T_0], \dots, [nT_0, (n + 1)T_0]$. The probability to find ζ_k such that the state leaves W_δ within $[kT_0, (k + 1)T_0]$ is not less than p_0 . All these leaving events are independent, because ζ_k are independent. Therefore, the probability of leaving W_δ within $[0, nT_0]$ is not less than p_0^n , and as a result, we obtain (32). \square

This result confirms Gromov–Carbone’s hypothesis on the fundamental instability of biosystems under large parameter variations. However, if we suppose that the gene regulation system can evolve, it is possible that inequality (2) holds for a $\beta > 0$. Then, evolution can continue within a large time frame, with a positive probability. We consider that evolution in the coming subsection.

7.2. Evolution and DNN's

To investigate the evolution problem in more detail, we apply recent results of the deep networks theory [4,24,38]. Recently, people started to apply ReLU networks to model gene regulation [3,6,39]. The deep networks with ReLU activators allow us to overcome the curse of dimensionality [38].

Let the gene regulation be defined by deep networks F_{Net} with ReLU activator functions. Remember that the depth $L = L_{net}$ of a network is the number of hidden layers, and the size $S = S_{net}$ is the total number of units. Let us denote by N_{layer} the maximal number of units in the network layers.

The following estimate follows from results [4]. Let $G(u)$ be a Lipschitz function defined on a compact set $K \subset \mathbb{R}^m$. Then there is a network F_{net} with $L_{net} > 3$ layers and N_L units in each layer such that

$$\sup_{v \in K} |G(v) - F_{net}(v)| < 2\sqrt{2m}N_L^{-2/m}Lip(G), \tag{34}$$

where $Lip(G)$ is the Lipschitz constant of G .

Note that this approximation is made by a network with real-valued synaptic weights, for example, for a single hidden layer perceptron, we have

$$F_{Net}(v) = \sum_{l=1}^{N_L} a_l \sigma\left(\sum_{k=1}^n w_{lk}v_k + w_0\right), \tag{35}$$

where a_l, w_{lk} are real valued coefficients. We however need a network whose parameters depend on a binary string s . For ReLU sigmoids σ , this binary coding problem can be resolved as follows. Consider for simplicity the case (35) (a multilayered case can be considered in a similar way). We first find values of real valued coefficients a_l, w_{lk} , providing the needed approximation. Each w_{lk} we approximate by M bits within a precision $\epsilon_b \ll \epsilon_k$:

$$|w_{lk} - w_{lk}^b| < \epsilon_b,$$

where

$$w_{lj}^b = (1/2 - \bar{s}_j) \sum_{l=-M_1}^{M_2} s_{j,l} 2^{-l}, \quad \bar{s}_j, s_{j,l} \in \{0, 1\}.$$

The ReLU function is Lipschitz with the constant 1. Therefore,

$$\sup_{v \in D} |F_{net}(v) - F_{net,B}(v)| < C|w - w^b| < C_1\epsilon_b$$

for bounded inputs u, v and some constants $C, C_1 > 0$, and $F_{net,B}$ denotes the network of the same structure as F_{net} but with the weights w^b . The same procedure with the replacement of w_l to w_l^b can be performed for coefficients a_l .

Remark. If the network F_{net} has the size S_{net} , then we use $N_g = O(S_{net}) \log_2 \epsilon_b$ genes to construct the network $F_{net,B}$ encoded by a binary string $s \in \{0, 1\}^{N_g}$.

7.3. Result on Replicative Stability

We assume that assumption **GR** holds and consider the non-perturbed system (16). Let \mathcal{A} be a local attractor of the semiflow S^t generated by this system. The following theorem describes the probability of stochastic robustness of this attractor with gene regulation under a sequence of shocks $\xi_k, k = 1, \dots, M_s$. Together with (16) we consider the same systems with gene regulation terms defined by DNN's (which are considered in the previous subsection):

$$\frac{dv}{dt} = f(v, \xi_k, s^{(k)}) \quad t \in (\tau_k, \tau_{k+1}]. \tag{36}$$

Within each interval $I_k = (\tau_k, \tau_{k+1}]$, the system under an environmental shock is defined by the parameter $\zeta_k \in \mathcal{P}_{ext,k}$, $k = 1, 2, \dots$, where $\mathcal{P}_{ext,k}$ is the k -th set of external perturbations. The shocks ζ_k are independent and we choose the set $\mathcal{P}_{ext,k}$ randomly from a countable set $\{\mathcal{P}_{ext,1}, \mathcal{P}_{ext,2}, \dots\}$. The index k defines a kind of interaction between the organism and its environment. At the k -th step the corresponding gene regulation networks $F_{net,k}$ have the size S_k , the depth $L > 3$ and N_k units in each layer. The weights of connections in that network are encoded by a genotype $s^{(k)}$ as it is explained in the previous subsection. We suppose that this network finds the correct choice $s^{(k)} = \bar{s}(\zeta_k)$ instantly. These networks $F_{i,net,k}$ solve the following approximation problem:

Approximation problem AP For a given target function

$$g_i(v, \zeta_k, s) = \sum_{\mathbf{a} \in \mathcal{R}_i} \tilde{C}_{i,\mathbf{a}}(\zeta_k, s)$$

of v to find a deep network $F_{i,net,k}$ such that the error

$$err_k = \sup_{v \in D} |g_i(v, \zeta_k, s) - F_{i,net,k}| \tag{37}$$

is minimal.

Further, to formulate the next theorem, let us introduce important parameters permitting to estimate the Lipschitz constants of the target functions g_i . The first parameter is the sparsity of the reaction network m_i , which is the number of elements in \mathcal{R}_i . We introduce the sparsity of the whole biochemical system by $Sparse(f) = \max_i m_i$. The second parameter is a sensitivity of kinetic rates under the perturbation of k -th type defined as follows. Let

$$Y_{sens,i,\mathbf{a},k} = \sup_{\zeta \in \mathcal{P}_{ext,k}} |\nabla_{\zeta} \ln \tilde{C}_{i,\mathbf{a}}|.$$

Then the k -th stress sensitivity Y_k is

$$Y_k = \max_{i,\mathbf{a} \in \mathcal{R}_i} Y_{sens,i,\mathbf{a},k}.$$

The third parameter is the productivity of the network, or in other words, the maximal (over all components and all reactions) reaction output

$$P_{\mathcal{A}} = \max_{i,\mathbf{a} \in \mathcal{R}_i, v \in \mathcal{A}} |C_{i,\mathbf{a}} v^{\mathbf{a}}|.$$

Then the Lipschitz constant $Lip_{\zeta}(g)$ of perturbation g as a function of ζ can be estimated via the introduced quantities:

$$Lip_{\zeta}(g) \leq Sparse(f) Y_k P_{\mathcal{A}}. \tag{38}$$

An estimate obtained in the next theorem is rough; nonetheless, it allows us to identify the main factors, which determine the organism fate under the stress stability: the sparsity of reaction network, the stress sensitivity, and the size of the gene regulation network.

Theorem 3. Suppose that \mathcal{A} is either a stable hyperbolic equilibrium or a stable hyperbolic limit cycle for the semiflow S^t defined by non-perturbed system (16), and we have stresses ζ_k at the moments $\tau_1, \dots, \tau_{M_s} < \tau$.

Then there is a sequence of networks $F_{i,net,k}$, $k = 1, 2, \dots, M_s$, $i = 1, 2, \dots, n$ such that for sufficiently small positive $\delta < \delta_0(f, \mathcal{A})$ the stochastic stability $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ satisfies

$$\mathbb{P}_{\delta, \mathcal{A}, \tau} > \Psi = \prod_{k=1}^K P_k, \tag{39}$$

where

$$P_k = \Pr \left[\text{Sparse}(f) Y_k P_{\mathcal{A}} < C_0 \delta N_k^{2/n} \right],$$

and C_0 is a constant uniform in δ as $\delta \rightarrow 0$.

Let us make some comments.

(i) The main idea is to provide stability as follows: at each time we choose $F_{i,net,k}$ resolving the approximation problem **AP**.

(ii) If the attractor \mathcal{A} is chaotic, the attractor contains unstable trajectories and a priori estimates from the proof fail. The standard ideas connected with hyperbolicity do not work because our perturbation is not autonomous in time and moreover it is not sufficiently smooth ($F_{net,k}$ are Lipschitz functions but they are not C^1).

(iii) We suppose that the number of coding genes in the layer N_{layer} may depend on the evolution step $k : N_{layer} = N_k$, where N_k is the number of genes in each layer at the k -th evolution step. The dependence on k is very important: it determines the evolution of the regulation network, and according to assertion 2, without that evolution, it is impossible to reach the robustness for all times.

(iv) From the biological point of view, assumptions of the theorem are not quite realistic because we suppose here that the gene regulation system instantly finds an optimal genotype $s^{(k)}$, but it is hard to study a completely realistic model. If the time lag between action of environmental stress and the system adaptive answer is sufficiently small with respect to a characteristic time of biochemical evolution, then one can expect that our results are still valid (although the mathematical analysis becomes more sophisticated). However, if that time lag is not small, our results are not applicable and the biochemical system can lose its adaptivity.

(v) For infinite sequences of perturbations ζ_k , the stability for all times is possible if the series $\sum_k (-\ln P_k)$ converges.

(vi) Due to results of Section 7.2 (see the remark at the end of that subsection), the number N_g of genes needed to encode a deep gene regulations network can be estimated as $N_k \log_2 N_k$, i.e., it is more than the network size up to a logarithmic factor. This is roughly consistent with experimental data. For example, for the *E. coli* metabolic network includes 744 reactions catalyzed by 607 enzymes and the *E. coli* genome contains 4288 protein coding genes [40].

Note that Theorems 1 and 3 lead to the following interesting corollary: if $H_{ext,\delta} = +\infty$, then, to provide eternal robustness (2), the size N_k of the gene regulation network cannot be bounded as $k \rightarrow \infty$: $\sup_k N_k = +\infty$.

Proof. Note that $F_{net,k}(v)$ are non-smooth (although Lipschitz bounded) functions of v ; therefore, we cannot use ideas connected with ϵ -shadowing, hyperbolicity, etc. Instead of this, we use rough straight forward estimates. Let us consider the Cauchy problems

$$\frac{d\bar{v}}{dt} = \bar{f}(\bar{v}), \quad \bar{v}(0) = \bar{v}^0,$$

$$\frac{dv}{dt} = \bar{f}(v) + g(v, \zeta_k, s^{(k)}), \quad v(\tau_k) = \bar{v}^0.$$

where $v^0 \in D$ and $t \in I_k = (\tau_k, \tau_{k+1}]$. Let $w(t) = v - \bar{v}$. Then for $w(t)$ one has

$$\frac{dw}{dt} = D\bar{f}(\bar{v})w + h(w) + g(v, \zeta_k, s^{(k)}), \quad w(0) = 0, \tag{40}$$

where $|h(w)| < Cw^2$ and $D\bar{f}(\bar{v})$ denote the derivative of \bar{f} with respect to v .

Lemma 1. Let the errors of approximation problem **AP** satisfy

$$err_k < \epsilon_0 \tag{41}$$

for each k . Then for sufficiently small $\epsilon_0 > 0$ one has

$$\sup_{t>0} |w(t)| < C_0 \epsilon_0, \quad (42)$$

where C_0 is uniform in $\epsilon_0 > 0$.

For the proof, which is quite standard, see Appendix A.

Using the lemma, at k -th evolution step we solve approximation problem **AP**. We use the estimate (34). This estimate shows that the probability to satisfy condition (41) for each k equals Ψ , which according to Lemma 1, gives the estimate of $\mathbb{P}_{\delta, \mathcal{A}, \tau}$ from below. This proves the theorem. \square

8. Numerical Simulations

To illustrate our analytical results on evolution, systems (10) and (11) are considered. To describe the evolution of cancer cells, we extend the model in [29]. Note that among all model parameters, r is a key parameter that defines the growth of the cancer cell. It is natural thus to assume that mutations increasing this parameter play the main role, and these mutations support replicative stability. We introduce a small mutation probability $p_{mut} > 0$. We solve systems (10) and (11) by the Euler method with a time step dt , and at each time step positive mutations can occur, increasing the growth parameter r . At each time step, the growth parameter either increases by a small quantity, or that parameter stays the same. The growth event occurs with the probability p_{mut} . We suppose that at these mutation time moments, we change the growth parameter r by an increment d_r , where $d_r > 0$ be a positive parameter. Therefore, r subsequently takes the values $r, r + d_r, r + 2d_r, \dots$

Following [29], we suppose that $\zeta(t)$ has the form of a sequence of periodical impulses:

$$\zeta(t) = \eta \sum_{k=1}^{\infty} \delta(t - k\Delta t), \quad (43)$$

where $\eta > 0$ is a parameter (the therapy intensity), and δ stands for the Dirac delta function. Such a choice corresponds to jumps in concentrations $v_1(t)$ at time moments $\Delta t, 2\Delta t, \dots$

We consider the three cases:

- (i) No therapy and no evolution of cancer cells, $\eta = 0$;
- (ii) Therapy and no evolution of cancer cells, $\eta > 0$ and $d_r = 0$;
- (iii) Therapy and evolution of cancer cells, $\eta > 0$ and $d_r > 0$.

Results are shown in Figures 1 and 2.

We see that despite therapy, the cancer cell population starts to grow sooner or later, and the increasing curve nicely describes such a growth.

One can conclude that it is necessary to take into account cancer cell adaptivity and evolution, which allows us to describe cancer dormancy, an important effect in cancer treatment [41]. Many patients relapse due to dormant cancer cells. These rare and elusive cells can hide in specialized niches in distant organs before being reactivated to cause disease relapse [41]. The cancer dormancy is a very complex problem, and we postpone this issue for future articles.

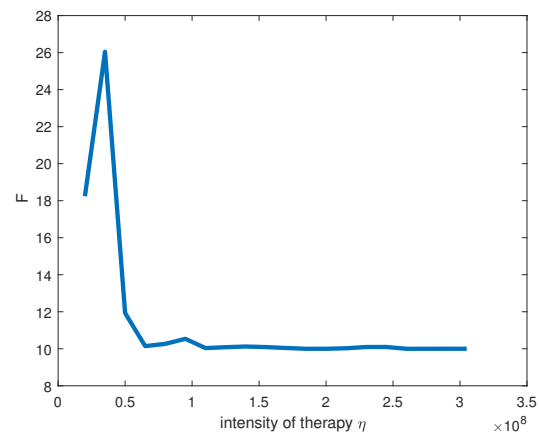


Figure 2. This plot shows how the replicative stability affects the abundance of cancer cells. The quantity F on the graph is the fraction, where the numerator is the abundance of the cancer cells for the replicative case and the denominator is the same abundance in the case when the mutations are absent. We see that F is almost constant for a large therapy dosage η . The parameters are selected in the same way as in the previous figure.

9. Discussion

In this paper, the idea of replicative stability proposed by M. Gromov and A. Carbone is investigated for systems that are standard models for metabolic kinetics and population dynamics. To this end, systems of ODE's are considered, which involve terms describing environmental shocks controlled by random parameters and gene regulation terms defined by networks. As a model of gene regulation, we used deep neural networks (DNN) here with ReLU activators. Since the deep networks overcome the curse of dimensionality, the problem of existence of genetic response to environmental stress can be resolved effectively.

In a framework of such approach, a connection is found between the uncertainty (entropy) of external stresses induced by a random environment and the topology of regulation networks. To compensate for the stress impact, the gene regulation networks should have a sufficient number of the hubs. Here the main result is an estimate of the hub number in regulation networks via the stress parameter dimension. Roughly speaking, to support stability, the number of the regulation network hubs should be not less than the dimension of the stress impact (this dimension can be simply understood as the number of different threats to the stability of the system). We think that this result is quite general and it is applicable to many systems. In fact, biological, ecological and economic systems are clearly becoming more complex in the process of evolution. However, it is not so easy to formalize these intuitive ideas on the complexity growth in precise mathematical terms. In this paper, using Kolmogorov's entropy, deep networks, and the Gromov–Carbone idea of replicative stability, we shed light on this hard problem. The main idea is that more complex systems support homeostasis in a more effective manner. The fluctuations of system internal parameters could be essentially less for systems having a regulation. For example, the DNA replication in such systems can be performed in a more reliable way and therefore more complicated systems stabilize their genetic code better, which give them a selective advantage.

From the biological point of view, the main result of this paper is that replicative stability provides a stability of populations within a large time period and even maybe an eternal stability (a sufficiently unpleasant conclusion, COVID-19 forever!). We show that it is possible due to gene regulation network evolution. We also present models that allow us to estimate the fraction of cancer cells (viruses, bacteria) that still survive after many stages of therapy.

Author Contributions: Conceptualization, D.G. and S.V.; writing—original draft preparation, S.V.; writing—review and editing, D.G. and S.V. All authors have read and agreed to the published version of the manuscript.

Funding: The first author was supported by the Ministry of Science and Higher Education of the Russian Federation by the Agreement N 075-15-2020-933 dated 13 November 2020 on the provision of a grant in the form of subsidies from the federal budget for the implementation of state support for the establishment and development of the world-class scientific center Pavlov center “Integrative physiology for medicine, high-tech healthcare, and stress- resilience technologies”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Proof of Lemma 1

Proof. Let us consider the Cauchy problem defined by the linear equation

$$du/dt = D\bar{f}(\bar{v}(t))u(t), \quad u(\bar{t}) = w.$$

The solutions of this equation define the semigroup $\mathcal{L}(t, \bar{t})$ as follows:

$$u(t) = \mathcal{L}(t, \bar{t})w.$$

Then we can rewrite (40) as an integral equation:

$$w(t) = \int_0^t \mathcal{L}(t, t')(h(w(t')) + g_k(v(t'), \bar{\zeta}(s(t')), s(t'))))dt'. \tag{A1}$$

Let us consider first the case when \mathcal{A} is a hyperbolic equilibria \bar{v} . Then the semigroup $\mathcal{L}(t, t') = \exp(\mathbf{L}(t - t'))$ induced by linear operator $\mathbf{L} : u \rightarrow Df(\bar{v})u$. It satisfies

$$|\mathcal{L}(t, t')u| \leq c_0 \exp(-\kappa(t - t'))|u|, \quad t > t', \tag{A2}$$

where c_0, κ are positive constants. Equation (A1) and that last estimate imply the integral inequality

$$|w(t)| \leq \bar{C} \int_0^t (|w(t')|^2 + \epsilon_0) \exp(-\kappa(t - t'))dt', \tag{A3}$$

where \bar{C} is a positive constant uniform in ϵ_0 . Let $X(t) = |w(t)|$ and let us denote by $\Phi(X(\cdot))$ the integral operator defined by the right hand side of the last inequality:

$$\Phi(X(\cdot))(t) = \int_0^t (X(t')^2 + \epsilon_0) \exp(-\kappa(t - t'))dt'.$$

We consider the operator Φ on the space of bounded non-negative functions $X(t)$ defined on $[0, +\infty)$ with the norm $\sup_{t \geq 0} X(t)$ if ϵ_0 is small enough then for an appropriate $C_0 > 0$ the operator Φ maps the subset $\{X(t) < C_0\epsilon_0\}$ into itself. This implies (42) for the case of an equilibrium.

In order to obtain (42) in the case of the limit cycle $\bar{v}(t)$, we slightly modify our arguments because then the semigroup $|\mathcal{L}(t, \tau)|$ is not exponentially decreasing in $|t - \tau|$. We proceed as follows. Let $\bar{v}(t)$ be a cycle. At the cycle, we represent the solution as $v(t) = \bar{v}(t + \phi(t)) + w(t)$, where ϕ is defined by the condition that distance $|v(t) - \bar{v}(t + \phi(t))|$ between the cycle and the solution $v(t)$ is minimal. For $w(t), \phi(t)$ we obtain two differential equations, and when we restrict on the subspace of w such that $(w, d\bar{v}/dt) = 0$ for all t the semigroup $|\mathcal{L}(t, \tau)|$ satisfies (A2). \square

Appendix A.2. Proof of Theorem 1

Appendix A.2.1. Auxiliary Lemma

Let us prove a preliminary Lemma.

Lemma A1. Let

$$\inf_{s \in S^N} |\langle g(\cdot, \zeta, s) - g(\cdot, \tilde{\zeta}, s) \rangle_{\mathcal{A}}| > \delta > 0 \quad (\text{A4})$$

where $\delta > C\varepsilon$ and $C > 0$ is an appropriate constant uniform in ε as $\varepsilon \rightarrow 0$. Then

$$\bar{s}(\zeta) \neq \bar{s}(\tilde{\zeta}). \quad (\text{A5})$$

Proof. Suppose that (A5) is not valid, i.e., $\bar{s}(\zeta) = \bar{s}(\tilde{\zeta}) = \bar{s}$. Due to Proposition 1, one obtains

$$|\langle g(\cdot, \zeta, \bar{s}) \rangle_{\mathcal{A}}| < C_1\varepsilon, \quad (\text{A6})$$

$$|\langle g(\cdot, \tilde{\zeta}, \bar{s}) \rangle_{\mathcal{A}}| < C_1\varepsilon. \quad (\text{A7})$$

These estimates imply

$$|\langle g(\cdot, \tilde{\zeta}, \bar{s}) - g(\cdot, \zeta, \bar{s}) \rangle_{\mathcal{A}}| < 2C_1\varepsilon.$$

Let $C > 2C_1$. Then one has a contradiction with (A4) that completes the proof. \square

Appendix A.2.2. End of Demonstration

To conclude the proof of Theorem 1, we just use the definition of $H_{ext,\delta}$. The proof is complete.

References

- Gromov, M.; Carbone, A. *Mathematical Slices of Molecular Biology*; Preprint IHES/M/01/03 (2001); Societe Mathematique de France: Paris, France, 2001.
- Lenski, R.E.; Wisler, M.J.; Ribeck, N.; Blount, Z.D.; Nahum, J.R.; Morris, J.J.; Zaman, L.; Turner, C.B.; Wade, B.D.; Maddamsetti, R.; et al. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. *Proc. R. Soc. Biol. Sci.* **2015**, *282*, 20152292. [[CrossRef](#)] [[PubMed](#)]
- Kishan, K.; Rui, L.; Cui, F.; Yu, Q.; Haake, A.R. GNE: A deep learning framework for gene network inference by aggregating biological information. *BMC Syst. Biol.* **2019**, *13*, 38.
- Shen, Z.; Yang, H.; Zhang, S. Nonlinear Approximation via Compositions. *Neural Netw.* **2019**, *119*, 74–84. [[CrossRef](#)] [[PubMed](#)]
- Weinan, E.; Wang, Q. Exponential convergence of the deep neural network approximation for analytic functions. *China Sci. Math.* **2018**, *61*, 1733–1740.
- Liu, Y.; Barr, K.; Reinitz, J. Fully interpretable deep learning model of transcriptional control. *Bionformatics* **2020**, *36*, i499–i507. [[CrossRef](#)] [[PubMed](#)]
- Mochizuki, A.; Bernold Fiedler, B. Sensitivity of chemical reaction networks: A structural approach. 1. Examples and the carbon metabolic network. *J. Theor. Biol.* **2015**, *21*, 189–202. [[CrossRef](#)]
- Brehm, B.; Fiedler, B. Sensitivity of chemical reaction networks: A structural approach. *arXiv* **2017**, arXiv:1606.00279v3.
- Aubin, J.P. *Dynamic Economic Theory: A Viability Approach*; Springer: Berlin/Heidelberg, Germany, 1997.
- Vakulenko, S.; Grigoriev, D. Stable growth of complex systems. In Proceedings of the Fifth Workshop on Simulation, St. Petersburg, Russia, 26 June–2 July 2005; pp. 705–709.
- Vakulenko, S. *Complexity and Evolution of Dissipative Systems*; de Gruyter: Berlin, Germany, 2014.
- Jiang, P.; Ludwig, M.L.; Kreitman, M.; Reinitz, J. Natural variation of the expression pattern of the segmentation gene even-skipped in *Drosophila melanogaster*. *Dev. Biol.* **2015**, *405*, 173–181. [[CrossRef](#)]
- Surkova, S.; Spirov, A.V.; Gursky, V.V.; Janssens, H.; Kim, A.R.; Radulescu, O.; Vanario-Alonso, C.E.; Sharp, D.H.; Samsonova, M.; Reinitz, J. Canalization of Gene Expression in the *Drosophila* Blastoderm by Gap Gene Cross Regulation. *PLoS Biol.* **2009**, *7*, e1000049. [[CrossRef](#)]
- Vakulenko, S.; Radulescu, O.; Manu, Reinitz, J. Size Regulation in the Segmentation of *Drosophila*: Interacting Interfaces between Localized Domains of Gene Expression Ensure Robust Spatial Patterning. *Phys. Rev. Lett.* **2009**, *103*, 168102–168106. [[CrossRef](#)]
- Gould, S.J. *The Structure of Evolutionary Theory*; The Belknap Press of Harvard University Press: Cambridge, MA, USA, 2002.
- Rutherford, S.L.; Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **1998**, *396*, 336–342. [[CrossRef](#)]

17. Kauffman, S.A. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **1969**, *22*, 437–467. [[CrossRef](#)]
18. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [[CrossRef](#)]
19. Samuelsson, B.; Troein, C. Superpolynomial Growth in the Number of Attractors in Kauffman Networks. *Phys. Rev. Lett.* **2003**, *90*, 098701. [[CrossRef](#)]
20. Vakulenko, S.; Morozov, I.; Radulescu, O. Maximal switchability of centralized networks. *Nonlinearity* **2016**, *29*, 2327. [[CrossRef](#)]
21. Jeong, H.; Tombor, B.; Albert, R.; Oltvai, Z.N.; Barabási, A.L. The large-scale organization of metabolic networks. *Nature* **2000**, *407*, 651–654. [[CrossRef](#)] [[PubMed](#)]
22. Jeong, H.; Mason, S.P.; Barabási, A.L.; Oltvai, Z.N. Lethality and centrality in protein networks. *Nature* **2001**, *411*, 41–42. [[CrossRef](#)] [[PubMed](#)]
23. Cohen, A.; Dahmen, W.; Daubechies, I.; DeVore, R. Tree Approximation and Optimal Encoding. *Appl. Comput. Harmon. Anal.* **2001**, *11*, 199–226. [[CrossRef](#)]
24. Grohs, P.; Perekrestenko, D.; Elbrächter, D.; Bolcskei, H. Deep Neural Network Approximation Theory. *arXiv* **2019**, arXiv:1901.02220v4.
25. Levy, S.; Siegal, M. Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PLoS Biol.* **2008**, *6*, e264. [[CrossRef](#)]
26. Seo, C.H.; Kim, J.R.; Kim, M.S.; Cho, K.H. Hub genes with positive feedbacks function as master switches in developmental gene regulatory networks. *Bioinformatics* **2009**, *25*, 1898–1904. [[CrossRef](#)]
27. Yang, W.; Zhao, X.; Han, Y.; Duan, L.; Lu, X.; Wang, X.; Zhang, Y.; Zhou, W.; Liu, J.; Zhang, H.; et al. Identification of hub genes and therapeutic drugs in esophageal squamous cell carcinoma based on integrated bioinformatics strategy. *Cancer Cell Int.* **2019**, *19*, 142. [[CrossRef](#)] [[PubMed](#)]
28. Friedgut, E. Sharp thresholds of graph properties, and the k -sat problem. *J. Am. Math. Soc.* **1999**, *12*, 1017–1055. [[CrossRef](#)]
29. Pang, L.; Chen, L.; Zhao, Z. Mathematical Modelling and Analysis of the Tumor Treatment Regimens with Pulsed Immunotherapy and Chemotherapy. *Comput. Math. Methods Med.* **2016**, *2016*, 6260474. [[CrossRef](#)]
30. Kuznetsov, V.; Makalkin, I.; Taylor, M.; Perelson, A. Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis. *Bull. Math. Biol.* **1994**, *56*, 295–321. [[CrossRef](#)]
31. Hofbauer, J.; Sigmund, K. *Evolutionary Games and Population Dynamics*; Cambridge University Press: Cambridge, UK, 1998.
32. Takeuchi, Y. *Global Dynamical Properties of Lotka-Volterra Systems*; World Scientific Pub Co., Inc.: Singapore, 1996.
33. Mjolsness, E.; Sharp, D.H.; Reinitz, J. A connectionist model of development. *J. Theor. Biol.* **1991**, *152*, 429–453. [[CrossRef](#)]
34. Reinitz, J.; Sharp, D.H. Mechanism of formation of eve stripes. *Mech. Dev.* **1995**, *49*, 133–158. [[CrossRef](#)]
35. Katok, A.; Hasselblatt, B. *Introduction to the Modern Theory of Dynamical Systems*; Cambridge University Press: Cambridge, UK, 1997.
36. Ruelle, D. *Elements of Differentiable Dynamics and Bifurcation Theory*; Elsevier: Amsterdam, The Netherlands, 2014.
37. Eckmann, J.P.; Ruelle, D. Ergodic theory of strange attractors. *Rev. Mod. Phys.* **1985**, *57*, 617–656. [[CrossRef](#)]
38. Montanelli, H.; Yang, H.; Du, Q. Deep ReLU networks overcome the curse of dimensionality for generalized bandlimited functions. *arXiv* **2020**, arXiv:1903.00735.
39. Martinez, C.A.; Barr, K.A.; Kim, A.; Reinitz, J. A synthetic biology approach to the development of transcriptional regulatory models and custom enhancer design. *Methods* **2013**, *62*, 91–98. [[CrossRef](#)]
40. Ouzounis, C.A.; Peter, D.; Karp, P.D. Global Properties of the Metabolic Map of *Escherichia coli*. *Genome Res.* **2000**, *10*, 568–576. [[CrossRef](#)] [[PubMed](#)]
41. Phan, T.G.; Croucher, P.I. The dormant cancer cell life cycle. *Nat. Rev. Cancer* **2020**, *20*, 398–411. [[CrossRef](#)] [[PubMed](#)]