Field Theory of a Reaction-Diffusion Model of Quasispecies Dynamics

Romualdo Pastor-Satorras
Ricard V. Solé

SFI WORKING PAPER: 2001-05-024

SANTA FE INSTITUTE

Field theory for a reaction-diffusion model of quasispecies dynamics

Romualdo Pastor-Satorras[1] and Ricard V. Solé[1,2,3]
[1]*Complex Systems Research Group, Departament of Physics, FEN Universitat Politècnica de Catalunya, Campus Nord B4, 08034 Barcelona, Spain*
[2]*NASA-associated Astrobiology Institute, Carretera del Ajalvir Km 4, Madrid*
[3]*Santa Fe Institute, 1399 Hyde Park Road, New Mexico 87501, USA*
(May 4, 2001)

Abstract

RNA viruses are known to replicate with extremely high mutation rates. These rates are actually close to the so-called error threshold. This threshold is in fact a critical point beyond which genetic information is lost through a second-order phase transition, which has been dubbed the "error catastrophe." Here we explore this phenomenon using a field theory approximation to the spatially extended Swetina-Schuster quasispecies model [J. Swetina and P. Schuster, Biophys. Chem. **16**, 329 (1982)], a single-sharp-peak landscape. In analogy with standard absorbing-state phase transitions, we develop a reaction-diffusion model whose discrete rules mimic the Swetina-Schuster model. The field theory representation of the reaction-diffusion system is constructed. The proposed field theory belongs to the same universality class than a conserved reaction-diffusion model previously proposed [F. van Wijland *et al.*, Physica A **251**, 179 (1998)]. From the field theory, we obtain the full set of exponents that characterize the critical behavior at the error threshold. Our results present the error catastrophe from a new point of view and suggest that spatial degrees of freedom can modify several mean field predictions previously considered, leading to the definition of characteristic exponents that could be experimentally measurable.

PACS numbers: 87.10.+e, 02.50.-r, 64.60.Ak

I. INTRODUCTION

RNA viruses offer a unique opportunity for exploring long term evolution under controlled conditions due to their high mutation rates [1,2]. Their evolutionary success is to a large extent due to their small-sized genomes, but specially to their enormous plasticity and adaptability to changing environments [3]. These viruses display the highest possible mutation rates and, as a consequence, their populations, the so-called *molecular quasispecies* [4], are extremely heterogeneous. The quasispecies structure has numerous implications for the biology of viruses. The most relevant of them is that mutant swarms are reservoirs of variants with potentially useful phenotypes in the face of environmental change. As extremely simplified entities at the border of a life-like state, they are specially apt to mathematical modeling [5,6].

The replication of these molecules implies two basic reactions [4,7]: (i) error-free copy, when a (molecular) species I_i replicates by using available monomers (A), i.e.:

$$(A) + I_i \xrightarrow{A_i Q_i} 2I_i \tag{1}$$

and (ii) mutation:

$$(A) + I_i \xrightarrow{\Psi_{ij}} I_i + I_j \quad (i \neq j) \tag{2}$$

The parameters A_i and Q_i are the replication rate and the quality factor, respectively. $Q_i \in [0, 1]$ is a measure of the correctness of the replication process, and it is maximum ($Q_i = 1$) if no mutations occur. Ψ_{ij} are the mutation rates which can lead to transitions between species $j \to i$.

The standard approach to the quasispecies dynamics (in the limit of very large populations) is based on the continuous Eigen model [4]. Here a set of molecules which can replicate and mutate is considered. The basic equations are:

$$\frac{dx_i}{dt} = (A_i Q_i - D_i)x_i + \sum_{j \neq i} \Psi_{ij} x_j + \Phi_i, \tag{3}$$

where x_i, with $i = 1, 2, ..., n$, accounts for the population size of each species, D_i stands for spontaneous degradation of molecules (assumed to be linear), and Φ_i is an outflow term which takes into account the removal of molecules from the system. If we introduce the constraint of constant population size (CP), $\sum_i x_i = \text{const.}$, the previous equations read:

$$\frac{dx_i}{dt} = (A_i Q_i - D_i - \bar{E})x_i + \sum_{j \neq i} \Psi_{ij} x_j, \tag{4}$$

where the mean value of the so-called excess productivity $E_i = A_i - Q_i$ is given by $\bar{E} = \sum_i (A_i - D_i)x_i / \sum_j x_j$.

One of the most important results of Eigen's theory was the finding that a phase transition takes place when mutation rates are tuned. Specifically, let us assume for simplicity that each species is composed by a string of elements $\{S_1, ..., S_\nu\}$ of size ν [8]. We consider in total a population of N strings. A particular kind of sequences, composed by the elements $\{S_1^{(0)}, ..., S_\nu^{(0)}\}$, represent the correct genomic sequences of the species, and is called

the *master sequence*. Each time a string is chosen to be replicated, it does so with some sequence-dependent probability $r_i = P(\{S_i\})$. If replication occurs, each unit can mutate with probability μ, and reproduce exactly with probability $1-\mu$. Mutation introduces disorder into the system, and it can be shown that a well-defined mapping exists between replication dynamics and the two-dimensional Ising model in such a way that the temperature is given by $T \approx -|\log(\mu/(1-\mu))|$ [9–11]. For small μ the replication system reaches a steady state in which the probability of observing the master sequence is finite. On the other hand, for values of μ larger than a given threshold, the probability of observing the master sequence is vanishing small.

Eigen's theory predicts that genetic information becomes lost for mutation rates higher than the critical rate μ_c, due to a breakdown of heredity and the lack of selection—the so-called "error catastrophe." It has been shown that this phenomenon indeed occurs in RNA viruses, which replicate close to the error threshold [3]. Actually, experimental data reveal that real viruses have mutation rates $\mu \approx 1/\nu$, that is, inversely proportional to the size of the genomic content, consistently with the prediction, and this has led to the claim that increased mutation rates might be able to bring virus populations into extinction. Such a strategy has been recently shown to hold *in vitro* and is likely to be feasible *in vivo* [12].

The presence of a critical mutation rate allows to interpret the error catastrophe in the framework of standard *absorbing-state phase transitions* (APT) [13,14]. APT are a class of non-equilibrium transitions in which, by the variation of a control parameter, the system crosses from an active phase with everlasting activity, to an absorbing phase, in which the system remains trapped forever, with no possibility to escape. In the framework of species replication with mutation, the active phase is identified with the low mutation regime, while the absorbing phase corresponds to the high mutation regime. Most APT are phase transitions of second order. If we characterize the system by an appropriate order parameter ψ, which in this case corresponds to the density of master sequences in the system, by tuning the parameter μ we observe the typical behavior

$$\begin{aligned} \psi &= 0, && \text{for } \mu > \mu_c \\ \psi &\simeq (\mu_c - \mu)^{\beta}, && \text{for } \mu < \mu_c, \end{aligned}$$

close to the critical point μ_c. The previous expression serves to define the critical exponent β. The analogy with error catastrophe is in this sense clear: for mutation rates larger than the error threshold, the virus is inviable and it quickly dies. For small mutation rates the virus is able to survive and reaches viable populations whose size is an increasing function of $\mu_c - \mu$. Further extending the analogy with APT, we can consider the spatial and time dependence of the order parameter ψ, and define the correlation function $g(r,t) = \langle \psi(r',t')\psi(r'+r,t'+t)\rangle$, where the bracket denote averages over different realizations of the system. According to the dynamic scaling ansatz [15], we expect to observe close to the error threshold the behavior

$$g(r,t) = r^{-(d-2+\eta)} F\left(\frac{r}{\xi}, \frac{t}{\xi^z}\right) \tag{5}$$

which defines the correlation length ξ, related to the distance to the μ_c by

$$\xi \sim (\mu_c - \mu)^{\nu_\perp}. \tag{6}$$

Eqs. (5) and (6) define the new critical exponents η, $\nu_\perp$ and z, which determine the scaling of the correlation function with respect to changes in the mutation rate μ.

Based in the previous analogy, in this paper we propose to study the phenomenon of the error catastrophe from the point of view of an APT by analyzing a reaction diffusion model with captures the essence of the replication-plus-mutation mechanism of quasispecies dynamics. The model allows the construction of an associated field theory, representative of the same universality class, following a standard technique outlined in the work of Doi, Cardy, and others [16–19]. The field theory developed here is shown to correspond to a conserved reaction diffusion model previously proposed by van Wijland *et al* [20] (see also Ref. [21]), in which the critical exponents were obtained performing a one-loop renormalization group analysis.

An important aspect, seldom considered in previous studies, is the effect of spatial degrees of freedom in quasispecies models. An exception is Adami's work on artificial life systems, in which a set of replicating bit strings of code spread on a two-dimensional lattice [22]. Under appropriate conditions, it was shown that the population spontaneously evolves to the error threshold, although no characterization of the model behavior at this critical point was performed. Besides, only a few experimental studies have recently reported the presence of several patterns of virus distribution that cannot be explained in terms of spatially-implicit quasispecies dynamics [23,24].

The interest in this problem is twofold: On the one hand, virus populations show heterogeneity in space, thus introducing further complexity in quasispecies dynamics and creating new opportunities to viral evolution. On the other hand, it would be important to know if spatially extended, mean-field models, are appropriate descriptions of the real quasispecies dynamics in space. Although RNA fitness landscapes are known to be rugged [25,26], here we consider the simplest, single-sharp peak landscape [27]. This model has been used as a null model of quasispecies populations and a field theory of the reaction-diffusion rules can be developed.

The paper is structured as follows. In Sec. II we propose, in analogy with the Swetina-Schuster model [27,28] a streamlined reaction-diffusion model which captures the minimal elements in the reproduction/mutation mechanism of quasispecies dynamics. Sec. III reports a mean-field analysis, which allows to pinpoint the key parameters of the model. In Sec. IV we construct the field theory corresponding to the model. In analogy with the analysis performed by van Wijland *et al.* and Kree *et al* [20,21], we obtain the relevant critical exponents. Finally, we interpret our results and put forward several experimental applications of them in Sec. V.

II. REACTION-DIFFUSION MODEL

The key ingredient of our model consists in the assumption that one of the sequences $B \equiv I_m$ has a high replication rate, while *all* the others $A \equiv I_{j\neq m}$ have the same, lower replication rate. The first sequence is called the master sequence and this approximation defines the so-called Swetina-Schuster model [27]. Second, by assuming that the sequences are long enough (consistently with real RNA viruses, where $\nu > 10^4$), backward mutations from A to B can be neglected [29].

In order to propose the reaction steps defining the model, we consider a simplified version of the Swetina-Schuster model (see Ref. [28] and references therein). In this model, a population of N strings of size ν evolves by a mechanism of replication with errors. Each string is defined by a sequence $S_1, \ldots S_\nu$, with $S_i = \{0, 1\}$. At each time step, we select a string and replicate it, after removing another string chosen at random. The replication takes place with probability 1 for the master sequence, defined by $S_i = 1, \forall i$, and with probability $p < 1$ for the rest. The replication procedure replaces each element S_i of the string for $S'_i = S_i$ with probability $1 - \mu$, and for $S'_i = (S_i + 1) \bmod 2$ with probability μ. This version of the model shows in plain view the elementary steps of the error catastrophe: replication of the master sequence at a certain rate, mutation of the master sequence, and lack of backward mutation, for sufficiently large sequence length. Another important ingredient in the model to be remarked is the implicit constraint of constant number of sequences, realized in the random deletion step, and which is usually implemented in the quasispecies analytical models [4]. The model described in Ref. [28] seems to display all the features of the error catastrophe.

Using this simple framework, we can translate the dynamics of the Swetina-Schuster model at a microscopic level in terms of reactions among particles of type B and A, corresponding to the master sequence and the mutants. respectively. The simplified model in Ref. [28] implicitly introduces interactions among sequences, by means of the random deletion step. Thus, our reaction diffusion model considers all the possible binary reactions between particles of type B and A, that are compatible with the outcome of the rules used in [28]. The set of reactions that we consider is:

$$B + B \xrightarrow{\mu} B + A \tag{7a}$$

$$B + A \xrightarrow{1-\mu} B + B \tag{7b}$$

$$A + B \xrightarrow{\lambda} A + A \tag{7c}$$

The steps represent the replication/mutation of the first species, coupled with the random deletion of the second species. Thus, the reaction (7a) implements the replication with mutation of a master sequence B, which happens with an effective mutation rate μ, coupled to the random deletion of a sequence of type B; the reaction (7b) represent the exact replication of a master sequence, at rate $1 - \mu$, with the deletion of a A sequence; finally, the reaction (7c) stands for the exact replication of a sequence A, at rate λ, together with the deletion of a master sequence. All the remaining binary reactions with replication/mutation plus random deletion do not alter the total number of particles, and are thus not considered. The proposed set of reactions mimic the conserved nature of the model imposed by the random deletion of sequences in Ref. [28], in a more nature way than in the original quasispecies model, Eq. (3), in which one had to impose an external flow term Φ_i in order to ensure conservation.

The set of equations (7) constitutes the core of our model. Spatial effects are taken into account by allowing the different particles to diffuse with respective diffusivities D_A and D_B [30]. Given the interpretation of the different particles, it is natural to consider $D_A = D_B$, that is, both master sequence and mutants diffuse with the same speed. However, for the sake of completeness, we will develop the formalism with $D_A \neq D_B$, and make them equal only as a last step.

III. MEAN-FIELD ANALYSIS

In order to gain some preliminary intuition on the behavior of the model, we analyze it by applying a standard mean-field analysis. Let us denote by ρ_B and ρ_A the density of species B and A, respectively. Since the reactions (7) conserve the number of particles, the total density $\rho_B + \rho_A$ is constant in time:

$$\rho_B + \rho_A = \rho. \tag{8}$$

The classic (mean-field) equations for the densities ρ_B and ρ_A are readily found to be:

$$\frac{\partial \rho_B}{\partial t} = (1 - \lambda - \mu)\rho_B \rho_A - \mu \rho_B^2 \tag{9a}$$

$$\frac{\partial \rho_A}{\partial t} = -(1 - \lambda - \mu)\rho_B \rho_A + \mu \rho_B^2. \tag{9b}$$

Combining this with the conservation condition (8) we obtain a single equation for the density ρ_B of master sequences:

$$\frac{\partial \rho_B}{\partial t} = (1 - \lambda - \mu)\rho \rho_B - (1 - \lambda)\rho_B^2. \tag{10}$$

This equation has two stable stationary states, depending on the value of μ:

$$\rho_B = 0, \qquad \text{for } \mu > 1 - \lambda \tag{11}$$

$$\rho_B = \frac{1 - \lambda - \mu}{1 - \lambda}\rho \qquad \text{for } \mu < 1 - \lambda. \tag{12}$$

At the mean-field level we observe the presence of a standard absorbing-state phase transition at a critical point $\mu_c = 1 - \lambda$. In the subcritical regime, $\mu > \mu_c$, the order parameter (in this case the density of master sequences) vanishes; in the supercritical region, $\mu < \mu_c$, the order parameter has a power-law dependence on μ:

$$\rho_B \simeq (\mu_c - \mu)^\beta, \tag{13}$$

which defines the critical exponent β in the mean-field approximation, $\beta_{\rm MF} = 1$. As we will see in the next section, the presence of fluctuations will change the value of β at the relevant, experimental, dimensions. A very interesting property of this conserved RD model is that the critical point is independent of the total particle density ρ, and is given as a function only of the reproduction rate λ. This situation should be compared with the conservative RD systems proposed so far, in which the total particle density plays the role of the tuning parameter, and must be tuned to a critical density ρ_c in order for the system to display critical behavior [20,31].

The mean-field solution also provides the expression for the probability $\mathcal{P}$ that there is at least one master sequence in the steady-state regime [32]:

$$\mathcal{P} = \Theta(\mu_c - \mu), \tag{14}$$

where Θ is the Heaviside function.

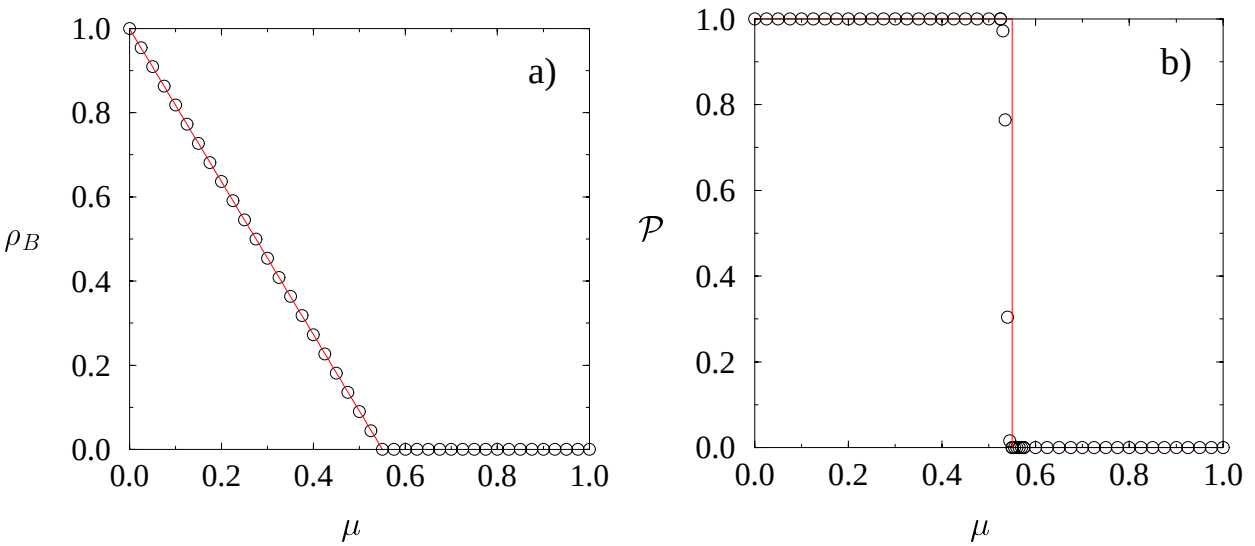

FIG. 1. Density of master sequences (a) and probability of finding at least one master sequence at the end of the run (b) for numerical mean-field simulations of the reaction-diffusion system (7). The full lines correspond to the analytic predictions. In particular, performing a linear regression in (a), we recover the expected critical point $\mu_c = 0.550 \pm 0.001$.

In order to check the mean-field predictions we have performed single-site computer simulations of the model. Simulations start with N particles of type B, and proceed by iterating the four reactions described in Eq. (7). In practice this is done sequentially, selecting at random a reaction in (7) and applying it to a B particle with probability $k_i\rho_i$, with k_i the rate of the reaction and ρ_i is the fraction particles of type i (either A or B) that also participate in the reaction. In Fig. 1 we plot the outcome of numerical simulations with total number of particles $N = 10000$. Simulations are run until a time $T = 5000$, averaging over the last 500 time steps. Statistical averages are performed over 250 independent simulation. We fix the reaction rate $\lambda = 0.45$, which corresponds to a critical point $\mu_c = 0.55$. The plots in Fig. 1 show that the simulations recover perfectly the behavior predicted by the analytic mean-field calculations.

IV. FIELD THEORY

Our model consists in a set of particles of type B and A diffusing in a hypercubic lattice of mesh size h and size L, that interact probabilistically according to the rules (7) whenever they meet at the same lattice site. The dynamics of the model is defined through a master equation for the probability $P(\{n\},\{m\},t)$ of having a particle configuration $\{n\}$, $\{m\}$ of particles B and A, respectively, at time t. The configuration $\{n\} = \{n_1, n_2, \ldots, n_{L^d}\}$ represents the occupation number of each node in the lattice. The master equation for P is:

$$\frac{\partial}{\partial t}P(\{n\},\{m\},t) =$$
$$\frac{D_B}{h^2}\sum_{i,j}\left[(n_j+1)P(\ldots n_i-1, n_j+1, \ldots, \{m\},t) - n_i P(\{n\},\{m\},t)\right] \tag{15a}$$

$$+\frac{D_A}{h^2}\sum_{i,j}\left[(m_j+1)P(\{n\},\dots m_i-1,m_j+1,\dots,t)-m_iP(\{n\},\{m\},t)\right] \tag{15b}$$

$$+\mu\sum_i\left[(n_i+1)n_iP(\dots n_i+1,\dots m_i-1,\dots,t)-n_i(n_i-1)P(\{n\},\{m\},t)\right] \tag{15c}$$

$$+(1-\mu)\sum_i\left[(n_i-1)(m_i+1)P(\dots,n_i-1,\dots,m_i+1,\dots,t)-n_im_iP(\{n\},\{m\},t)\right] \tag{15d}$$

$$+\lambda\sum_i\left[(n_i+1)(m_i-1)P(\dots,n_i+1,\dots,m_i-1,\dots,t)-n_im_iP(\{n\},\{m\},t)\right], \tag{15e}$$

where D_B and D_A are the diffusion coefficients for B and A particles, i is summed over all the lattice sites, and j over the nearest neighbors of the site i. The first two terms in the rhs of Eq. (15) implement diffusion through a random hopping of particles between nearest neighbor sites. The initial condition $P(\{n\},\{m\},0)$ is given by a Poisson distribution, with an average density per site equal for both types of particles.

The next step consists in recasting the master equation into a "second quantized" form, following the procedure described by Doi [16–19]. We introduce two sets of annihilation and creation operators at each lattice site, $\hat{b}_i$ and $\hat{b}_i^\dagger$ for B particles, and $\hat{a}_i$ and $\hat{a}_i^\dagger$ for A particles, which fulfill the standard commutation rules

$$\left[\hat{a}_i,\hat{a}_j^\dagger\right]=\left[\hat{b}_i,\hat{b}_j^\dagger\right]=\delta_{ij}. \tag{16}$$

With this commutations rules, the operators have a bosonic character, that is natural given the multiple occupancy of sites allowed in the model. With the help of the vacuum state $|0\rangle$, defined by $\hat{a}_i|0\rangle=\hat{b}_i|0\rangle=0$, we construct an orthonormal basis of states $|n,m\rangle$, defined by

$$|n,m\rangle=\prod_i(\hat{b}_i^\dagger)^{n_i}(\hat{a}_i^\dagger)^{m_i}|0\rangle\,, \tag{17}$$

and work in the Fock space spanned by this basis. In terms of this Fock space, the state of the system at time t is represented by the vector state $|P(t)\rangle$, defined as

$$|P(t)\rangle=\sum_{\{n\},\{m\}}P(\{n\},\{m\},t)\,|n,m\rangle\,. \tag{18}$$

In terms of this vector state, the master equation Eq. (15) can be rewritten as a Schrödinger equation in imaginary time

$$\frac{\partial}{\partial t}|P(t)\rangle=-\hat{H}\,|P(t)\rangle\,, \tag{19}$$

with a Hamiltonian, or time-evolution operator, $\hat{H}$ defined by:

$$\hat{H}=\sum_{\langle ij\rangle}\left[\frac{D_A}{h^2}(\hat{a}_i^\dagger-\hat{a}_j^\dagger)(\hat{a}_i-\hat{a}_j)+\frac{D_B}{h^2}(\hat{b}_i^\dagger-\hat{b}_j^\dagger)(\hat{b}_i-\hat{b}_j)\right] \tag{20}$$

$$+\mu\sum_i(\hat{b}_i^\dagger-\hat{a}_i^\dagger)\hat{b}_i^\dagger\hat{b}_i^2+(1-\mu)\sum_i(\hat{a}_i^\dagger-\hat{b}_i^\dagger)\hat{b}_i^\dagger\hat{b}_i\hat{a}_i+\lambda\sum_i(\hat{b}_i^\dagger-\hat{a}_i^\dagger)\hat{a}_i^\dagger\hat{b}_i\hat{a}_i. \tag{21}$$

Eq. (19) can be formally solved in terms of the operator $\hat{H}$ yielding

$$|P(t)\rangle = \exp(-\hat{H}t)\,|P(0)\rangle\,. \tag{22}$$

From this solution, it is possible to derive all the statistical properties of the RD system, applying a projection technique [16–19]. For practical purposes, it is convenient to map this second-quantized form into a field theory, using a coherent state representation. Performing a time-slicing of the evolution operator in Eq. (22), via the Trotter formula, we can express the vector state $|P(t)\rangle$ as a path integral, weighted with the exponential of an action S, over a set of classical fields a^*, a, b^*, and b, which are related with the two types of particles. After taking the continuum limit ($h \to 0$), the vector state can be written as the path integral over space and time dependent fields

$$|P(t)\rangle = \int \mathcal{D}a\;\mathcal{D}a^*\;\mathcal{D}b\;\mathcal{D}b^*\;\exp(-S[a,a^*,b,b^*])\,|P(0)\rangle\,, \tag{23}$$

where the action S has the form [33]

$$\begin{aligned} S[a,a^*,b,b^*] = \int d^dx \int dt \Big\{ & a^*[\partial_t - D_A\nabla^2]a + b^*[\partial_t - D_B\nabla^2]b \\ & + \mu(b^*-a^*)b^*b^2 + (1-\mu)(a^*-b^*)b^*ab + \lambda(b^*-a^*)a^*ab\Big\}\,. \end{aligned}$$

Within this formalism, we can compute the average value of any observable $F(\{n\},\{m\})$ performing the path integral

$$\langle F(t)\rangle = \mathcal{C}\int \mathcal{D}a\;\mathcal{D}a^*\;\mathcal{D}b\;\mathcal{D}b^*\;F(a,b)\exp(-S[a,a^*,b,b^*]), \tag{24}$$

where $\mathcal{C}$ is an appropriate normalization constant.

The final step in the derivation of the field theory consists in performing the shift

$$a^* = 1 + \bar{a}, \qquad b^* = 1 + \bar{b}, \tag{25}$$

and the change of variables

$$\begin{array}{ll} \psi = b, & \phi = a + b - \rho, \\ \bar{\psi} = \bar{b} - \bar{a}, & \bar{\phi} = \bar{a}. \end{array} \tag{26}$$

The final action describing the RD system is

$$\begin{aligned} & S[\psi,\bar{\psi},\phi,\bar{\phi}] = \\ & \quad \int d^dx \int dt \Big\{\bar{\psi}[\partial_t\psi - D_B\nabla^2\psi - r\psi + g_1\psi^2 - g_2\psi\phi] + \bar{\phi}[\partial_t\phi - D_A\nabla^2\phi + \gamma\nabla^2\psi] \\ & + \bar{\psi}^2[-g_3\psi - v_1\psi\phi + v_2\psi^2] + \bar{\psi}\bar{\phi}[-g_4\psi - v_3\psi\phi + v_4\psi^2]\Big\} \end{aligned} \tag{27}$$

where we have defined the coupling constants

$$\begin{aligned} r = g_4 &= (1-\lambda-\mu)\rho, \\ g_1 = v_4 &= 1-\lambda, \\ g_2 = v_3 &= 1-\lambda-\mu, \\ g_3 = \rho v_1 &= (1-\mu)\rho, \\ v_2 &= 1, \\ \gamma &= D_A - D_B. \end{aligned}$$

A naïve power counting shows that the critical dimension of this field theory is $d_c = 4$, and that the coupling constants v_i are irrelevant, and can be in principle discarded. With this final form, it is easy to recognize that the action (27) represents the same field theory analyzed by van Wijland *et al.* for the conserved reaction diffusion model

$$\begin{aligned} B + A &\longrightarrow 2B, \\ B &\longrightarrow A. \end{aligned}$$

In their study, the authors worked out the renormalization group analysis for this system, for both cases: $D_A < D_B$ and $D_A = D_B$., providing the critical exponents up to a one-loop expansion. The case $D_A = D_B$, which also corresponds to the universality class of a model of population dynamics with pollution described by Kree *et al.* [21], is the relevant one in the problem under consideration. Quoting the results of Refs. [20,21], we have the critical exponents:

$$\beta = 1 - \frac{\varepsilon}{32} \tag{28a}$$

$$\nu_\perp = \frac{1}{2 - \varepsilon/2} \tag{28b}$$

$$\eta = -\frac{\varepsilon}{8} \tag{28c}$$

$$z = 2, \tag{28d}$$

where $\varepsilon = 4 - d$ gives the dimensionality of the system. The results for $\nu_\perp$ and z are exact, derived field-theoretically by analysing the symmetries of the action (27), and are thus valid for all dimensions. The values of β and η, on the other hand, and expansions around $\varepsilon = 0$, and thus they are expected to hold only for small values of ε.

In view of the results (28), the relevant exponents at the fisical dimension $d = 3$ are

$$\begin{aligned} \nu_\perp &= 3/2 \qquad & z &= 2, \\ \beta &\simeq 0.969 \qquad & \eta &\simeq -0.125. \end{aligned} \tag{29}$$

The values of $\nu_\perp$ and z are exact, while those for β and η represent an approximation given by the replacement of the small parameter ϵ for 1. In dimension $d = 2$ or less, the exponents (28) have to be taken with a grain of salt, due to the terms v_i in Eq. (27), which might become relevant at low dimensions [20].

V. DISCUSSION

The dynamical theory of molecular evolution developed by Eigen and Schuster reveals the presence of an intrinsic, sharp limit to molecular information carriers. This threshold is a generic feature of replicator systems involving reproduction and mutation. The Eigen-Schuster theory predicts that, under the effect of evolutionary pressures selecting for high variability, such replicators will evolve towards the error threshold. This is the case of RNA viruses and experimental evidence clearly supports this theoretical prediction.

Previous theoretical models have analysed the stochastic dynamics of quasispecies under different approaches. But all of them considered spatially-implicit models (mean-field-like),

paying no attention to local effects derived from incomplete mixing. Here we have explored this problem using the simplest quasispecies model, described by a single-sharp-peak replication landscape. The aim of our study was to see how the statistical behavior of a spatially extended molecular replication system would differ from the mean-field predictions.

We have considered a simplified reaction-diffusion model where two types of "particles" (the master sequence and the mutant sequences) diffuse, replicate, and mutate on a given spatial domain. Applying the standard approach to absorbing-state phase transitions, a field theory has been developed and it has been shown to be the same reported by Wijand *et al.* [20] and Kree *et al.* [21].

The main message from our study is that relevant differences between mean-field models and real dynamics are expected to be observed even in the simplest scenario considered here. Larger deviations should be expected in more realistic models incorporating a better description of molecular replicators and their dynamics. In particular, the sharpness of the transition (as defined by the β exponent) is not very different at different dimensions. This suggests that no measurable differences should be expected to be observed in experimental systems. The correlation exponent ν does change appreciably, from $\nu(d = 1) \approx 2/3$ to $\nu(d = 3) \approx 2$ thus leading a faster decorrelation at realistic dimensions. Such an increase in ν will enhance the coexistence of different strains (quasispecies) in a given spatial domain [34] and thus the probabilities of success for the virus.

Several caveats of this approach are worth mentioning. It is known that real RNA viruses have actually multipeaked landscapes [25,26]. We will expand our analysis to such situation in a future work (although the associated field theory is expected to be much harder to develop and analyze). However, in many situations the quasispecies is observed to be confined in a fitness peak, so that our previous analysis essentially holds [25,26] . Also, the real RNA virus dynamics takes place through a virus-cell interaction not considered here, while several previous theoretical models used in order to understand well-defined experimental results (where a cell population was present) have been shown to be successful in providing a full understanding of the evolutionary dynamics of RNA populations [35,36].

ACKNOWLEDGMENTS

The authors thank Santiago Elena, Peter Stadler, and Alessandro Vespignani for valuable comments and suggestions. This work has been supported by a grant PB97-0693 and by the Santa Fe Institute (RVS).

REFERENCES

[1] B. D. Preston and J. P. Dougherty, Trends. Microbiol. **4**, 16 (1996).

[2] E. Domingo and J. J. Holland, Annu. Rev. Microbiol. **51**, 151 (1997).

[3] E. Domingo and J. J. Holland, in *The Evolutionary Biology of Viruses*, edited by S. Morse (Raven, New York, 1994), pp. 161–184.

[4] M. Eigen, J. McCaskill, and P. Schuster, Advances in Chemical Physics **75**, 149 (1989).

[5] D. Alves and J. F. Fontanari, Phys. Rev. E **57**, 7008 (1998).

[6] M. Nowak and P. Schuster, J. Theor. Biol. **137**, 375 (1989).

[7] G. W. Rowe, *Theoretical models in biology* (Oxford University Press, Oxford, 1994).

[8] The string can be visualized as idealized genomic sequence defining the species, with a length of ν genes.

[9] I. Leuthäusser, J. Chem. Phys. **84**, 1884 (1986).

[10] I. Leuthäusser, J. Stat. Phys. **48**, 343 (1987).

[11] P. Tarazona, Phys Rev. A **45**, 6038 (1992).

[12] L. A. Loeb, J. M. Essigmann, F. Kazazi, J. Zhang, and K. D. Rose, Proc. Nat. Acad. Sci. **96**, 1492 (1999).

[13] J. Marro and R. Dickman, *Nonequilibrium phase transitions in lattice models* (Cambridge University Press, Cambridge, 1999).

[14] G. Grinstein and M. A. Muñoz, in *Fourth Granada Lectures on Computational Physics*, Vol. 493 of *Lecture Notes in Physics*, edited by P. L. Garrido and J. Marro (Springer-Verlag, Berlin, 1997).

[15] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, 1996).

[16] M. Doi, J. Phys. A: Math. Gen. **9**, 1465 (1976).

[17] M. Doi, J. Phys. A: Math. Gen. **9**, 1479 (1976).

[18] L. Peliti, J. Phys. I **46**, 1469 (1985).

[19] B. P. Lee and J. Cardy, J. Stat. Phys. **80**, 971 (1995).

[20] F. van Wijland, K. Oerding, and H. J. Hilhorst, Physica A **251**, 179 (1998).

[21] R. Kree, B. Schaub, and B. Schmittmann, Phys. Rev. A **39**, 2214 (1989).

[22] C. Adami, *Introduction to Artificial Life* (Springer, New York, 1999)

[23] See for example: J. K. Wong, C. C. Ignacio, F. Torriani, D. Havlir, N. J. S. Fitch, and D. D. Richman, J. Virol. **71**, 2059 (1997); A. B. van't Wout, J. Virol. **72**, 488 (1998); S. Navas, J. Martín, J. A. Quiroga, I. Castillo, and V. Carreno, J. Virol. **72**, 1640 (1998).

[24] Spatially extended replicating systems have been studied by using $Q\beta$ RNA viruses on a one-dimensional setup (see J. McCaskill and G. J. Bauer, Proc. Nat. Acad. Sci. **90**, 4191 (1993). Under controlled conditions (an the resence of an appropriate RNA polymerase) waves of mutant viruses can be observed and their general properties measured.

[25] C. L. Burch and L. Chao, Genetics **151**, 921 (1999).

[26] C. L. Burch and L. Chao, Nature **406**, 625 (2000).

[27] J. Swetina and P. Schuster, Biophysical Chemistry **16**, 329 (1982).

[28] R. V. Solé, S. C. Manrubia, M. Benton, S. Kauffman, and P. Bak, Trends. Ecol. Evol. **14**, 156 (1999).

[29] The numerical analysis of the Swetina-Schuster model shows that this approximation works very well for $\nu > 50$ [27]. Sequences larger than this value provide a sharp phase transition. Since the smallest RNA molecules have a size of $\nu \approx 80$ (not to mention

RNA viruses, whose genomes have sizes $\nu > 10^3$) the current approximation seems well justified.
[30] Such a diffusion-like approximation has been applied in continuous partial differential equation models of virus spreading in growing plaques [see L. You and J. Yin, J. Theor. Biol. **200**, 365 (1999)].
[31] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **62**, R5875 (2000).
[32] Note that this probability is not related with the surviving probability in spreading experiments usually defined in absorbing-state phase transitions [13].
[33] Since we are only interested in the properties of the stationary state of the system, we have omitted in the expression of the action boundary terms that are relevant for the study of the relaxational dynamics; see for example Ref. [20].
[34] R. V. Solé and J. Bascompte, *Complexity and self-organization in evolutionary ecology*, (Princeton U. Press, Princeton, to appear).
[35] L. Tsimring, H. Levine, and D. A. Kessler, Phys. Rev. Lett. **76**, 4440 (1996)
[36] R. V. Solé, R. Ferrer, I. González-García, J. Quer and E. Domingo, J. Theor. Biol. **198**, 47 (1999).