Clustering of independently diffusing individuals by birth and death processes

Martin Meyer,[1,2] Shlomo Havlin,[1,3] and Armin Bunde[1]
[1]*Institut für Theoretische Physik, Justus-Liebig-Universität, Heinrich-Buff-Ring 16, 35392 Giessen, Germany*
[2]*Sonderforschungsbereich 173, Universität Hannover, Callinstrasse 3-3a, D-30176 Hannover, Germany*
[3]*Minerva Center and Departement of Physics, Bar-Ilan University, Ramat-Gan, Israel*
(Received 26 December 1995; revised manuscript received 21 May 1996)

We study analytically a model for the dynamics of populations and epidemics that includes birth and death processes in a system of independently diffusing individuals. When birth and death rates are close to each other, the individuals tend to cluster around their center of mass and the population (disease) territory migrates as a whole. The clustering phenomenon is caused by the natural asymmetry between birth and death processes, and does not require any communication among the individuals or between the individuals and the substrate. [S1063-651X(96)13010-5]

PACS number(s): 87.10.+e, 05.40.+j

I. INTRODUCTION

Diffusion models, where the features of birth (infection) and death (recovery) are included, are commonly used for studying the dynamics of populations, epidemics [1–5], and chemical reactions with creation and annihilation processes [6]. For these ''branching diffusion'' models, averaged quantities like the mean individual density $\langle n(\mathbf{r},t)\rangle$ at position $\mathbf{r}$ and time t can be easily obtained [1,2]. However, since $\langle n(\mathbf{r},t)\rangle$ gives no information about spatial correlations between two individuals in a single configuration, it cannot describe effects like clustering or segregation that are observed in field experiments [7–10].

Here we introduce an interesting branching diffusion model with reduced fluctuations, where we can calculate the pair correlation function $P_2(\mathbf{r},t)$ rigorously. We find two different types of phenomena, *migration* for short times and *spreading* for long times, with a crossover time $\sim 1/(\alpha_+ - \alpha_-)$ that approaches infinity when the birth and death rates α_+ and α_- become equal.

In the case of *migration* [Fig. 1(a)], the mean distance between the individuals saturates, i.e., the individuals cluster around their center of mass and the population territory (epidemics area) does not increase, but moves as a whole with time. This dynamical localization is based *only* on the natural asymmetry between birth (infection) and death (recovery) processes — birth and infection occur only next to an individual, while death and recovery occur everywhere — and

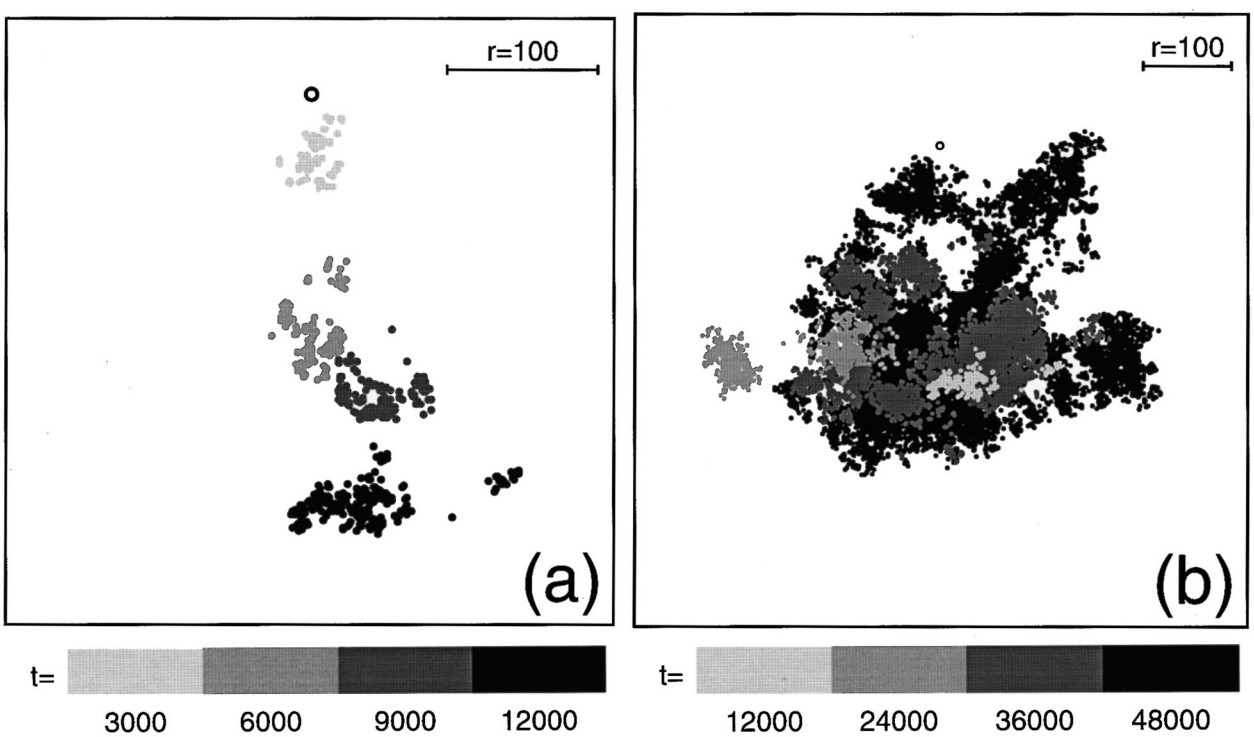

FIG. 1. Positions of 100 diffusing individuals in two dimensions for $2\,dD=1$, $\alpha_-=0.1$, and $\alpha_+=0.1001$ for several times (a) in the short time regime and (b) in the long time regime. Different colors refer to different times. The white ring marks the origin, where the individuals started at $t=0$.

 © 1996 The American Physical Society

does *not* require any interaction between the individuals or between the individuals and the substrate. In the case of *spreading* [Fig. 1(b)], the mean distance between the individuals increases all the time, but the center of the growing territory almost does not move.

II. MODEL

We consider N_0 individuals, initially located at the origin of a d-dimensional coordinate system. At $t=0$, the individuals start to diffuse independently, with a diffusion constant D. At certain instances of time t_k, one randomly chosen individual gives birth, i.e., is replaced by two individuals (at the same position $\mathbf{r}$), and another randomly chosen individual is removed (dies) with probability p_d. According to this rule, the population can never be extinct. At criticality ($p_d=1$), the number of individuals remains constant, and in this case the model essentially reduces the diffusion reproduction model studied by Zhang, Serva, and Polikarpov [11]. By definition, the death rate α_- and the birth rate α_+ are related to p_d by $\alpha_-=p_d\alpha_+$ [12].

It is natural to assume that the birth and death processes are Poissonian, i.e., the increments Δt between two successive time instances are exponentially distributed. Since for $N(t)$ individuals at time t, $\alpha_+N(t)$ is the number of birth events per unit time, it follows that the mean increment time is $\langle\Delta t\rangle=1/[\alpha_+N(t)]$. The mean number of individuals increases exponentially, $\langle N(t)\rangle=N_0\exp(\Delta\alpha t)$, where $\Delta\alpha\equiv\alpha_+-\alpha_-$ is the growth rate.

By construction, the fluctuations in $N(t)$ are only caused by the fluctuations of the time intervals Δt and (for $p_d<1$) of the number of dying individuals at each step. Thus the number $N(t)$ of individuals shows only weak statistical fluctuations around $\langle N(t)\rangle$ (that vanish in the ''critical'' case $\Delta\alpha=0$). In contrast, in the usual branching diffusion models [1], where birth and death events occur *independently* with rates α_+ and α_-, large fluctuations lead to a nonzero dying-out probability (which is one in the critical case). We show below that our reduced fluctuation model shows rich phenomena in the spatial dynamics of the community, and has the advantage of being solvable analytically at and above criticality.

To calculate $\langle n(\mathbf{r},t)\rangle$, we start with the probability density $P_i(\mathbf{r},t)$ that the ith individual is at position $\mathbf{r}$ at time t. Notice that we can trace the ancestory of any given individual to a unique ancestor for any given earlier time (see the genealogical tree). Thus we can define a generalized trajectory of individual i for all times $0\leqslant t'\leqslant t$ by identifying the position of i for times t' before its birth with the position of its unique ancestor at that time t'. Since the diffusion process is unaffected by the birth process, the statistical properties of this generalized trajectory are the same as for a single diffusing particle, and $P_i(\mathbf{r},t)$ is identical to the well known Gaussian probability density $G(\mathbf{r},t)$ for a *single* diffusing particle to move a distance $\mathbf{r}$ during time t,

$$P_i(\mathbf{r},t)=G(\mathbf{r},t)\equiv\left(\frac{1}{4\pi Dt}\right)^{d/2}\exp\left(-\frac{r^2}{4Dt}\right), \tag{1}$$

where D is the diffusion constant. The mean individual density $\langle n(\mathbf{r},t)\rangle$ is simply the product of $\langle N(t)\rangle$ and $G(\mathbf{r},t)$,

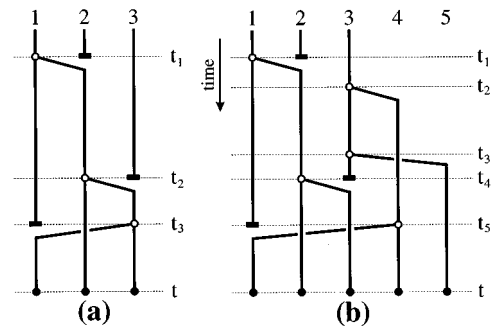

FIG. 2. The genealogical tree for $N_0=3$ for (a) the critical case, $\Delta\alpha=0$; and (b) a growing community, $\Delta\alpha>0$.

$$\langle n(\mathbf{r},t)\rangle=\langle N(t)\rangle G(\mathbf{r},t). \tag{2}$$

Equations (1) and (2) are common for branching diffusion models [1]. For obtaining the spatial correlations we need to analyze the genealogical tree of the individuals (see Fig. 2). If two individuals i and j at time t have a common ancestor, their line of ancestors split at some earlier time $t_s(i,j)>0$. We define $t_s(i,j)\equiv 0$, if i and j have no common ancestor. By definition, the generalized trajectories of i and j are identical for times smaller than $t_s(i,j)$. For times larger than $t_s(i,j)$, the trajectories are that of two independently diffusing particles starting at the common position of i and j at time $t_s(i,j)$. Thus the probability density for the distance $\mathbf{r}$ between the two individuals for a given $t_s\equiv t_s(i,j)$ is simply $G(\mathbf{r},2(t-t_s))$, where the factor of 2 is due to the fact that *both* individuals are moving.

The pair correlation function $P_2(\mathbf{r},t)$ is defined as the probability density that, after time t, the distance between two individuals is $\mathbf{r}$, and can be written as

$$P_2(\mathbf{r},t)=\int_0^t dt_s\psi_t(t_s)G(\mathbf{r},2(t-t_s)), \tag{3}$$

where $\psi_t(t_s)dt_s$ is the probability that the generalized trajectories of an arbitrarily chosen pair at time t were generated from a common ancestor in the time interval between t_s-dt_s and t_s. By definition, $\int_0^t\psi_t(t_s)dt_s=1$. As we show below, $\psi_t(t_s)$ can be calculated rigorously for arbitrary values of N_0, α_+ and α_-. For convenience, we begin with the critical case $p_d=1$ ($\Delta\alpha=0$).

III. CRITICAL CASE

For $p_d=1$, $N(t)=N_0$ remains constant. To calculate $\psi_t^{(c)}(t_s)$ (the superscript c denotes the critical case), we consider at time t that fraction of pairs that was generated from a common ancestor between t_s-dt_s and t_s, and divide it by that fraction of pairs that was generated from a common ancestor before t_s. Clearly, this ratio is independent of t and identical to the number of pairs generated at any time t_s, $N_0\alpha_+dt_s$, divided by the total number of pairs $N_0(N_0-1)/2$. Thus

$$\frac{\psi_t^{(c)}(t_s)dt_s}{\int_0^{t_s}\psi_t^{(c)}(t_s')dt_s'}=\frac{2N_0\alpha_+dt_s}{N_0/(N_0-1)}\equiv\frac{dt_s}{\tau_0}, \tag{4}$$

where $\tau_0 \equiv (N_0-1)/2\alpha_+$. The solution of this equation is

$$\psi_t^{(c)}(t_s) = \tau_0^{-1}\exp\left(-\frac{t-t_s}{\tau_0}\right) + \exp\left(-\frac{t}{\tau_0}\right)\delta(t_s). \tag{5}$$

The interpretation of Eq. (5) is clear: The splitting events are Poisson distributed with time constant τ_0, which for large times can be regarded as the mean value of $t-t_s$. The prefactor of the δ function is the probability that two individuals have no common ancestor.

The strong spatial correlations between the individuals caused by the natural asymmetry between birth and death processes show up most clearly in the mean-squared pair distance $\langle r_2^2(t)\rangle = \int d^d r \mathbf{r}^2 P_2(\mathbf{r},t)$, which can be easily calculated by using Eq. (3) and changing the order of integration,

$$\langle r_2^2(t)\rangle = 4dD\int_0^t dt_s \psi_t^{(c)}(t_s)(t-t_s) = \rho_0^2\left[1-\exp\left(-\frac{t}{\tau_0}\right)\right], \tag{6}$$

with a localization radius $\rho_0 \equiv \sqrt{4dD\tau_0}$. Equation (6) shows that $\langle r_2^2(t)\rangle$ is linear in t for short times $t \ll \tau_0$, while for large times $t \gg \tau_0$, $\langle r_2^2(t)\rangle$ saturates and becomes identical to ρ_0^2. Accordingly, after an initial spreading period, the individuals stay together in one community, that typically covers a region of radius ρ_0.

We can study this feature further by calculating the pair correlation function itself, which can be obtained by inserting Eq. (5) into Eq. (3),

$$P_2^{(c)}(\mathbf{r},t) = \left(\frac{1}{8\pi D}\right)^{d/2}\left[t^{-d/2}\exp\left(-\frac{t}{\tau_0}-\frac{r^2}{8Dt}\right) + \tau_0^{-1}\int_0^t dt_s t_s^{-d/2}\exp\left(-\frac{t_s}{\tau_0}-\frac{r^2}{8Dt_s}\right)\right]. \tag{7}$$

For short times, $t \ll \tau_0$, $P_2^{(c)}(\mathbf{r},t)$ is dominated by the first term in the brackets, and we recover the pair correlation function of independently diffusing individuals, without birth and death processes $P_2^{(c)}(\mathbf{r},t) = G(\mathbf{r},2t)$. For sufficiently large times, the first term vanishes, and the second term becomes independent of time since the upper limit of the integration can be extended to infinity. Hence, for large t, $P_2^{(c)}(\mathbf{r},t)$ becomes stationary, and $P_2^{(c)}(\mathbf{r},t) = P_2^{(c)}(\mathbf{r},\infty)$, which reflects the fact that the interindividual distance saturates and there is no spreading. The straightforward calculation of the integral leads to a modified Bessel function, which for large r can be approximated by an exponential times a power of r, yielding

$$P_2^{(c)}(\mathbf{r},\infty) \sim r^{-(d-1)/2}\exp\left(-\sqrt{2d}\frac{r}{\rho_0}\right). \tag{8}$$

Thus the shape of $P_2^{(c)}(\mathbf{r},t)$ changes from a Gaussian (typical for diffusion) to a simple time-independent exponential (typical for localized behavior) at large times.

Next we consider the motion of the center of mass of the community, $\langle r_{\rm cm}^2(t)\rangle \equiv \langle (N^{-1}\Sigma_i \mathbf{r}_i)^2\rangle$, which can be expressed by $\langle r_2^2(t)\rangle \equiv (N(N-1))^{-1}\Sigma_{i,j}\langle(\mathbf{r}_i-\mathbf{r}_j)^2\rangle$ and the mean-squared displacement $\langle r^2(t)\rangle \equiv N^{-1}\Sigma_i\langle \mathbf{r}_i^2\rangle = 2dDt$,

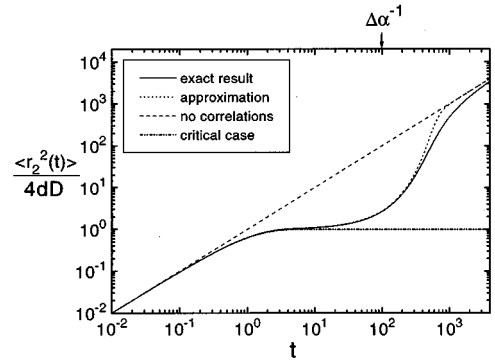

FIG. 3. The mean-squared pair distance in units of $4dD$ as function of time for $\tau_0=1$, $\Delta\alpha=0.01$ (—), and the approximate result $(\cdots)$ for the same parameters. For comparison, the results for $\tau_0=\infty$ (---) and $\tau_0=1$ and $\Delta\alpha=0$ $(\text{-}\cdot\text{-}\cdot\text{-}\cdot)$ are also shown.

$$\langle r_{\rm cm}^2(t)\rangle = \langle r^2(t)\rangle - \frac{N_0-1}{2N_0}\langle r_2^2(t)\rangle. \tag{9}$$

Equation (9) follows directly from the definitions of the three quantities. By substituting Eq. (6) into Eq. (9), and expanding the exponential function up to the second order, one can identify three different time regimes in $\langle r_{\rm cm}^2(t)\rangle$:

$$\langle r_{\rm cm}^2(t)\rangle = \begin{cases} 2dD\dfrac{t}{N_0} & \text{for } t \ll \dfrac{2\tau_0}{N_0-1} \\ 2dD\dfrac{N_0-1}{2N_0\tau_0}t^2 & \text{for } \dfrac{2\tau_0}{N_0-1} \ll t \ll \tau_0 \\ 2dDt & \text{for } t \gg \tau_0. \end{cases} \tag{10}$$

Only for very short times, $t \ll 2\tau_0/(N_0-1)$, the center of mass moves as in the case of independently diffusing particles. In the intermediate time regime $2/(N-1) \ll t/\tau_0 \ll 1$, $\langle r_{\rm cm}^2(t)\rangle \propto t^2$ shows ballistic behavior, while for long times $\langle r_{\rm cm}^2(t)\rangle \approx \langle r^2(t)\rangle = 2dDt$, independent of N_0.

Thus we have found that in the critical case ($\Delta\alpha=0$) the community shows clustering and migration behavior after an initial spreading period for $t<\tau_0$: The mean-squared pair distance becomes constant, $\langle r_2^2(t)\rangle = \rho_0^2$, and hence the individuals are localized around their center of mass within a ''ball'' of radius ρ_0 [see Fig. 1(a)]. In this state of dynamical localization, the shape of the pair correlation function is a simple exponential, which is typical for localized behavior, and the community as a whole moves like a single diffusing entity.

IV. GENERAL CASE

Next we consider the noncritical case $\Delta\alpha>0$, where the number of individuals, $N(t)$, increases with time. In this case, the time scale τ_0 becomes time dependent, $\tau_0 \to \tau(t') \equiv (N(t')-1)/2\alpha_+ \simeq \tau_0 e^{\Delta\alpha t'}$, and Eq. (4) is generalized to

$$\frac{\psi_t(t_s)dt_s}{\int_0^{t_s}\psi_t(t_s')dt_s'} = \frac{dt_s}{\tau_0 e^{\Delta\alpha t_s}}. \tag{4a}$$

From the solution of Eq. (4a),

$$\psi_t(t_s)=\left(\frac{1}{\tau(t_s)}+\delta(t_s)\right)\exp\left(\frac{1}{\Delta\alpha\tau(t)}-\frac{1}{\Delta\alpha\tau(t_s)}\right),$$

a straightforward calculation yields

$$\langle r_2^2(t)\rangle=\frac{4dD}{\Delta\alpha}\exp\left(\frac{1}{\tau(t)\Delta\alpha}\right)\quad E_1\left(\frac{1}{\tau(t)\Delta\alpha}\right),$$

where $E_1(x)\equiv\int_x^\infty\exp(-\xi)/\xi d\xi$ is the exponential integral function. Figure 3 shows $\langle r_2^2(t)\rangle$ in units of $4dD$ for $\tau_0=1$ and $\Delta\alpha=0.01$. For comparison we also show the curves for independently diffusing individuals without death and birth processes ($\tau_0=\infty$) and for the critical case ($\tau_0=1$, $\Delta\alpha=0$). For short times $t\ll\Delta\alpha^{-1}$, the curve for $\Delta\alpha=0.01$ follows the critical curve, and shows a regime of dynamical localization, since the number of individuals is nearly constant. For very long times, the curve for $\tau_0=\infty$ is approached, which indicates that the correlations induced by the asymmetry between birth and death processes become irrelevant. In the intermediate time regime, the mean-squared pair distance grows exponentially.

A surprisingly good approximation for this overall behavior can be found by using the equation for the critical case, Eq. (6), and replacing the time independent scales τ_0 and ρ_0 by the time dependent ones $\tau(t)=\tau_0\exp(\Delta\alpha t)$ and $\rho(t)\equiv\sqrt{4dD\tau(t)}$. This approximation is shown as a dotted line in Fig. 3.

V. SUMMARY

We have studied a fluctuation reduced branching diffusion model, for which we could calculate rigorously the correlations between the individuals. We have found that the dynamical behavior of the community depends crucially on the growth rate $\Delta\alpha$. For times t smaller than the inverse growth rate $\Delta\alpha^{-1}$, the correlations lead to clustering of the individuals, and the community migrates as a whole. This clustering phenomenon is caused by the natural asymmetry between birth and death processes, and does not require interactions or communications among the individuals or between the individuals and the substrate. For $t\gg\Delta\alpha^{-1}$, the correlations become irrelevant, and the community shows a normal spreading behavior [13–15].

After this work was submitted, we learned of a possible application of our model to evolution, when mutation of genes is considered as diffusion in the high-dimensional genetic space. The asymmetry between the death and birth of individuals leads to a dynamical localization such that different individuals of a given species will have similar genetic code. Indeed, Tsimring, Levine, and Kessler [16] used a similar model to explain the recent experimental observation by Novella *et al.* [17] of virus evolution in fitness space. In the work by Tsimring, Levine, and Kessler, a model for virus mutation, which can be regarded as diffusion in fitness space, was studied. They found that the viruses become localized in a single cluster in fitness space, where the center of mass performs a *directed* motion due to an anisotropy of the genetic space. Although this spatial anisotropy was not included here, both models show a strict analogy, since for both the localization effect is caused by the asymmetry between death and birth. In our model, we expect a directed motion of the center of mass, if the spatial anisotropy is included.

ACKNOWLEDGMENTS

We thank L. Benguigui, Y. Levi, S. Rabinovich, and G. H. Weiss for many helpful discussions. We thank the Alexander-von-Humboldt Foundation, the Minerva Foundation, and the Deutsche Forschungsgemeinschaft (SFB 173) for financial support.

[1] K. B. Althreya and P. E. Ney, *Branching Processes* (Springer, Berlin, 1972).
[2] S. Asmussen, and N. Kaplan, Stoch. Proc. Appl. **4**, 1 (1976); S. Asmussen and N. Kaplan, *ibid.* **4**, 15 (1976).
[3] J. D. Biggins, in *Spatial Spread in Branching Processes* (Springer, Heidelberg, 1980).
[4] Y. Iwasa and E. Teramoto, J. Math. Biol. **19**, 109 (1984).
[5] P. Grassberger and A. de la Torre, Ann. Phys. (N.Y.) **122**, 373 (1979).
[6] A. S. Mikhailov, Phys. Rep. **184**, 307 (1989), and references therein.
[7] C. E. Pielou, *Mathematical Ecology* (Wiley, New York, 1977).
[8] C. L. Bos, H. J. Mandersloot, F. Vader, and B. Steenbergen, Neth. J. Plant Pathol. **95**, 225 (1989).
[9] C. E. van den Lande, Neth. J. Plant Pathol. **99**, Suppl. 3, 129 (1993).
[10] R. Lecoustre, D. Fargette, C. Fauquet, and P. de Reffye, Phytopatology **79**, 913 (1989).
[11] Y.-C. Zhang, M. Serva, and M. Polikarpov, J. Stat. Phys. **58**, 849 (1990).
[12] The growth rate α_+ and the diffusion constant D are model parameters that can be used to define the length and the time scales.
[13] J. G. Skellam, Biometrika **38**, 196 (1951).
[14] W. E. Fry, S. B. Goodwin, A. T. Dyer, J. M. Matuszak, A. Drenth, P. W. Tooley, L. S. Sujkowski, Y. J. Koh, B. A. Cohen, L. J. Spielman, K. L. Deahl, D. A. Inglis, and K. P. Sandlan, Plant Disease **77**, 653 (1993).
[15] H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss, Nature **355**, 423 (1992); H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss, Phys. Rev. A **45**, 7128 (1992).
[16] L. Tsimring, H. Levine, and D. A. Kessler, Phys. Rev. Lett. (to be published).
[17] I. S. Novella, E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland, Proc. Natl. Acad. Sci. U.S.A. **92**, 5841 (1995).