PHYSICAL REVIEW E **100**, 022403 (2019)

Solution of the Crow-Kimura model with a periodically changing (two-season) fitness function

David B. Saakian[1,2,*] and Kang Hao Cheong[3,4,†]
[1]*Laboratory of Applied Physics, Advanced Institute of Materials Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[2]*Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[3]*Science and Math Cluster, Singapore University of Technology and Design (SUTD), S487372 Singapore*
[4]*SUTD-MIT International Design Centre, S487372, Singapore*

(Received 14 May 2019; published 5 August 2019)

Since the origin of life, both evolutionary dynamics and rhythms have played a key role in the functioning of living systems. The Crow-Kimura model of periodically changing fitness function has been solved exactly, using integral equation with time-ordered exponent. We also found a simple approximate solution for the two-season case. The evolutionary dynamics accompanied by the rhythms provide important insights into the properties of certain biological systems and processes.

DOI: 10.1103/PhysRevE.100.022403

I. INTRODUCTION

Evolutionary dynamics is the key feature of living systems [1], and the exponential growth of the replicators in the case with sufficient nutrients is perhaps the first feature of living systems. On the other hand, all known biological systems possess rhythms. It is important to look at a model with both features, and this forms the biological motivation of our work here. One of the key findings of Eigen [1,2] was the existence of two phases in an evolution model with a single-peak fitness landscape. When the mutation probability is higher than the threshold value, the system is in a nonselective phase where the population is dissolved in the whole sequence space. In the case of a dynamics fitness landscape, the situation becomes more complicated. In Ref. [3], they consider a single-peak fitness landscape with a peak sequence that changes after a certain period of time. A minimal threshold value for the mutation probability indicating the existence of the selective phase is determined. Some approximate results have been considered in Ref. [4] as well. In Ref. [5], they consider the single-peak Crow-Kimura model [6–9], where the peak sequence changes after some time. In Ref. [10], the mean fitness of the Crow-Kimura model on the static fitness landscape (first calculated in Ref. [9]) has been derived using the ideas of spin-glass models and the methods of quantum statistical physics (the Ising model with the transverse magnetic field). Some approximate solutions for the Eigen model with periodically changing peak sequence have been found in Ref. [11]. In the current study, we have adopted the methods from Ref. [10] for general applicability and used it to solve for the two-season model. The system uses one fitness function for the first part of the period and another one for the second part. Such evolutionary dynamics may be related to seasonal variation [12–15] in the natural world.

The calculation of the mean fitness for the quasispecies model is the first step in the investigation of the biological aspects of the model. First, we will calculate the exact mean fitness for the evolutionary dynamics with the fitness landscape varying periodically. Next, we give the main results of Ref. [10] in a way that can be generalized for the evolution model with a fitness landscape that is periodic in time. We then deduce the exact integral equation for the calculation of mean fitness in the case of general periodic fitness landscape. Finally, we derive a simple approximate solution for two-season evolution model that can be applied to certain biological systems.

II. THE CROW-KIMURA MODEL

Using the Crow-Kimura model, we consider the evolutionary dynamics of a genome modeled as a chain of N genes each having two alleles, ±1, which makes for 2^N genotypes. The Malthusian fitness of a sequence with l mutations from the reference (wild) sequence is Nr_l, and the mutation rate for the one allele change is μ. We collect all the sequences with l number of mutations to the Hamming lth class and define their probabilities as P_l. We then look at the system of equations:

$$\frac{dP_l(t)}{dt} = P_l N(r_l - \mu) + \mu[P_{l-1}(N-l+1) + P_{l+1}(l+1)] - \mu P_l(t) NR, \qquad (1)$$

where $NR = \sum_{l=0}^{N} Nr_l P_l$ is the mean fitness. We drop the P_{l-1} and P_{l+1} terms for the $l = 0$ and $l = N$ cases, respectively.

We can deduce the solution of Eq. (1) from the solution of the linear version of the same equation (dropping nonlinear terms), using a nonlinear algebraic transformation [16]:

$$\frac{dp_l(t)}{dt} = p_l N(r_l - \mu) + \mu[p_{l-1}(N-l+1) + p_{l+1}(l+1)],$$
$$P_l = \frac{p_l}{\sum_n p_n}. \qquad (2)$$

---

[*] david.saakian@tdtu.edu.vn
[†] Corresponding Author: kanghao_cheong@sutd.edu.sg

We assume the following form of the fitness landscape:

$$r_l = f(x,t), \quad x = 1 - 2l/N.$$

Let us now look at a matrix representation of Eq. (2) as $\frac{d}{dt}\vec{p} = -H\vec{p}$, where $\vec{p} = (p_0 \cdots p_N)$. We can derive the mean fitness of the steady state by inspecting the maximal eigenvalue of the matrix $-H$. The matrix H has a simple presentation as a quantum-mechanical Hamiltonian for the case of imaginary time [7,10]:

$$-H = \mu \sum_k \left(\sigma_k^x - 1\right) + f\left(\frac{\sum_k \sigma_k^z}{N}, t\right). \tag{3}$$

We can obtain a decomposition $H = H_1 + H_2$ by taking the first and second terms in Eq. (3). In the standard Crow-Kimura model, we have the static fitness landscape,

$$f(x,t) = f_0(x). \tag{4}$$

We can define different states S_i by inspecting the eigenstates σ_k^z. In Ref. [10], the mean fitness has been derived using the formula:

$$R = \lim_{\beta \to \infty} \frac{\log \mathrm{Tr} e^{-\beta H}}{\beta}. \tag{5}$$

In Ref. [10], the following expression for the mean fitness has been derived via the methods of quantum statistical physics using large β limit:

$$\begin{aligned} R &= \frac{1}{N\beta} \log[\mathrm{Tr} e^{-H/L} \ldots e^{-\beta H/L}] \\ &= \frac{1}{N\beta} \log[\mathrm{Tr} e^{-\beta H_1/L} e^{-\beta H_2/L} \ldots e^{-\beta H_1/L} e^{-\beta H_2/L}] \\ &= \frac{1}{N\beta} \log\left[\sum_i \langle S_i | e^{-\beta H_1/L} e^{-\beta H_2/L} \ldots e^{-\beta H_1/L} e^{-\beta H_2/L} | S_i \rangle\right]. \end{aligned} \tag{6}$$

Using a Suzuki-Trottere formalism, the Tr has been expressed by the classical statistical physics of the Ising model in the magnetic field $h(x)$. In Ref. [17], the following has been derived:

$$\begin{aligned} &\log \sum_j \langle S_i | e^{-\beta H_1/L} e^{-\beta H_2/L} \ldots e^{-\beta H_1/L} e^{-\beta H_2/L} | S_j \rangle \\ &= N \mathrm{Max}\Bigg[-\int_0^\beta h(x) m(x) dx + f_0[m(x)] \\ &\quad + \log\left[\mathrm{Tr}(1+\sigma_x) \hat{T} e^{\int_0^\beta dx h(x)\sigma_z + \beta\sigma_x}\right]\Bigg], \end{aligned} \tag{7}$$

where $\hat{T}$ means a time-ordered exponent, and we should determine the maximum point via $h(x), m(x)$.

For the static fitness landscape taking $h(x) = h, m(x) = m$ and the limit $\beta \to \infty$, we obtain:

$$NR = \log\{\mathrm{Max}[\exp[-\beta hm + \beta f_0(m) + \phi_1(\beta,h) - \beta]\}|_{h,m}], \tag{8}$$

where

$$\phi_1(\beta,h) = \beta\sqrt{1+h^2}. \tag{9}$$

At the extreme point, we then obtain $h = f_0'(m), f_0'(m) = \frac{m}{\sqrt{1-m^2}}$ and the mean fitness as

$$R = \mathrm{Max}[f_0(x) + \mu(\sqrt{1-x^2} - 1)]. \tag{10}$$

The same result can be determined using the Hamilton-Jacobi equation method [18,19], where the steady-state distribution has been derived. Such an approach is crucial for the admission of exact dynamics [20]. Equation (8) was first derived in Ref. [9] and the corresponding expression for the Eigen model [21] are key mathematical results for quasispecies models.

III. THE GENERAL PERIODIC FITNESS CASE

Now let us consider the general periodic fitness case. When fitness is periodic in time, i.e.,

$$f(x,t) = f(x,t+T). \tag{11}$$

We are interested in the dynamics for a βT time period. We define the mean fitness as

$$R = \lim_{\beta \to \infty} \frac{\log\left[\mathrm{Tr}\hat{T} e^{-\int_0^{\beta T} dt H(t)}\right]}{\beta T N}, \tag{12}$$

where we consider the time-ordered exponent, and β is a large integer. Then, slightly modifying the results of Ref. [17] given in the previous section, we obtain:

$$\begin{aligned} R = \max\Bigg\{ &\int_0^T [-h(x)m(x)dx + f_0(m(x),x)] \\ &+ \log\left[\mathrm{Tr}\hat{T} e^{\int_0^T dx h(x)\sigma_z + T\sigma_x}\right]\Bigg\}. \end{aligned} \tag{13}$$

We should look at the maximum of the right-hand side of Eq. (13). One simple equation is

$$h(x) = f_0'[m(x),x], \tag{14}$$

where f_0' means the derivative with respect to $m(x)$. We then obtain a different form of the integral equation:

$$\begin{aligned} m(x) &= \frac{\mathrm{Tr} G(x)\sigma_z G^{-x}[G(T)]}{\mathrm{Tr} G(T)}, \\ G(X) &= \hat{T} e^{\int_0^X dx h(x)\sigma_z + X\sigma_x}. \end{aligned} \tag{15}$$

Our Eqs. (13)–(15) are exact equations, and so the theoretical results obtained are widely generalizable. We now investigate an interesting real-world example from here.

IV. THE APPROXIMATE VERSION OF TWO-SEASON EVOLUTION MODEL

A. The approximate formula

Consider evolution in the Crow-Kimura model with mutation rate $\mu = 1$ and fitness,

$$\begin{aligned} f(x,t) &= f_1(x), \mathrm{Mod}(t,T) < T_1, \\ f(x,t) &= f_2(x), \mathrm{Mod}(t,T) > T_1, \end{aligned} \tag{16}$$

where $T = T_1 + T_2$. Let us look the version of Eq. (13) with

$$\begin{aligned} m(x) &= m_1, h(x) = h_1, \mathrm{Mod}(t,T) < T_1, \\ m(x) &= m_2, h(x) = h_2, \mathrm{Mod}(t,T) > T_1. \end{aligned} \tag{17}$$

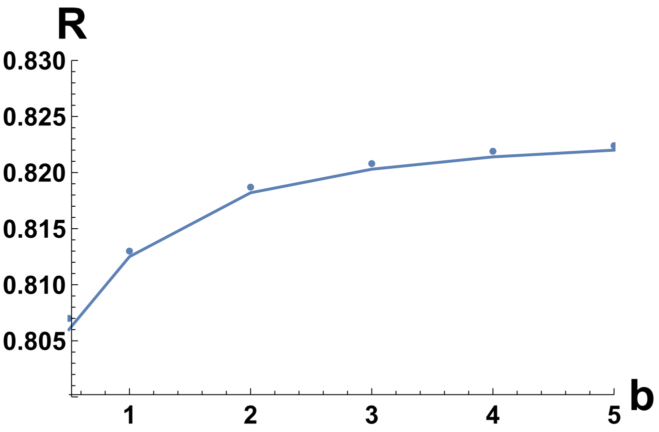

FIG. 1. The mean fitness R versus $b = T_1 = T_2$, $f_1(x) = x$, $f_2(x) = 2x$. The smooth line corresponds to our analytical results by Eq. (12), and the solid dots are the results of numerics.

We then derive the mean fitness

$$R(T_1 + T_2) = -T_1h_1m_1 + T_1f_1(m_1) - T_2h_2m_2 + T_2f_2(m_2) + \phi_2(T_1, h_{1,2}, h_2), \tag{18}$$

where $e^{\phi_2(T_1,h_1,T_2,h_2)}$ is the maximal eigenvalue of the linear operator.

$$e^{T_1(h_1\sigma_z+\sigma_x)}e^{T_2(h_2\sigma_z+\sigma_x)} \tag{19}$$

and is derived in the Appendix:

$$\begin{aligned}
\phi_2 &= \ln\frac{\cosh\big(\beta_1\sqrt{1+h_1^2}\big)\cosh\big(\beta_2\sqrt{1+h_2^2}\big)X}{\cosh(u)} + u\\
u &= \operatorname{arctanh}\frac{\sqrt{Y^2-Z^2}}{X},\\
X &= 1 + t_1t_2\frac{1+h_1h_2}{\sqrt{1+h_1^2}\sqrt{1+h_2^2}},\\
Y &= t_1t_2\frac{h_1-h_2}{\sqrt{1+h_1^2}\sqrt{1+h_2^2}},\\
Z &= \sqrt{t_1^2 + t_2^2 + 2t_1t_2\frac{1+h_1h_2}{\sqrt{1+h_1^2}\sqrt{1+h_2^2}}}
\end{aligned} \tag{20}$$

and we denote $t_1 = \tanh(T_1\sqrt{1+h_1^2})$, $t_2 = \tanh(T_2\sqrt{1+h_2^2})$.

B. Two different linear fitness functions

Considering evolution in the T_1 period via a fitness $f_1(x) = k_1x$ and T_2 period via $f_2(x) = k_2x$. Then we can derive the following:

$$R = \phi_2(T_1, k_1, T_2, k_2). \tag{21}$$

In Fig. 1 we have compared our analytical results with the numerics.

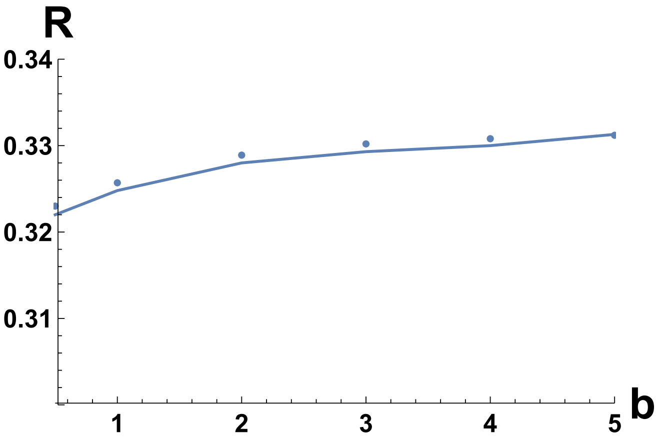

FIG. 2. The mean fitness R versus $b = T_1 = T_2$, $f_1(x) = x$, $f_2(x) = x^2$. The smooth line corresponds to our analytical results by Eq. (12), and the solid dots are the results of numerics.

C. The case when one of the fitnesses is a linear function

Consider another simple case when one of the fitnesses is a linear function. In order to calculate the mean fitness, we should find the maximum

$$R = -T_2h_2m_2 + T_2f_2(m_2) + \phi_2(T_1, T_2, h_2). \tag{22}$$

In Figs. 2 and 3, we have given the comparison of our results with the numerics. Clearly, there is excellent agreement between the two.

V. THE TWO-DIMENSIONAL TWO-SEASON FITNESS LANDSCAPE

For the case of the two-dimensional two-season fitness landscape, let us consider the generalization of Eq. (1), when the genome has two parts: lengths $L_1 = \lambda_1N$ and $L_2 = \lambda_2N$ and the total fitness depends on the number of mutations in

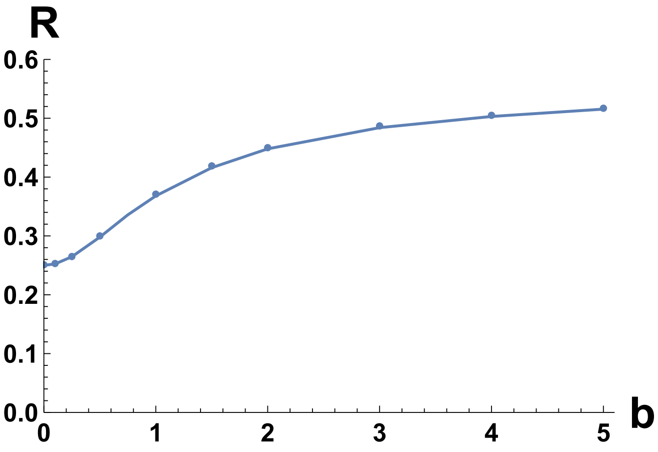

FIG. 3. The mean fitness R versus $b = T_1 = T_2$, $f_1(x) = 0$, $f_2(x) = x^2$. The smooth line corresponds to our analytical results by Eq. (12), and the solid dots are the results of numerics.

each part [22],

$$\begin{aligned}\frac{dP_{ln}(t)}{dt} &= P_{ln}(r_{ln}-\mu)N + \mu[P_{l-1,n}(L_1-l+1)\\ &\quad + P_{l,n-1}(L_2-n+1) + P_{l+1,n}(l+1)\\ &\quad + P_{l,n+1}(n+1)] - P_{ln}(t)NR, \end{aligned} \tag{23}$$

where $RM = \sum_{lnN} r_{ln}P_{ln}$ is the mean fitness and $m_1 = 1-2l/L_1, m_2 = 1-2n/L_2$. We have a mutation rate per gene μ/N. The fitness is described here via two functions $f_1(m_1,m_2)$ and $f_2(m_1,m_2)$ during the period β_1 and β_2.

Such a model adequately describes the situation whereby, after some period of time, the fitness peak moves from one peak sequence to the other one at some Hamming distance. For the static fitness landscape with fitness function $f_1(m_1,m_2)$, the model has been solved in Refs. [22,23],

$$\begin{aligned} RT &= -\lambda_1 T h_1 m_1 - \lambda_2 T h_2 m_2 + T f_1(m_1,m_2)\\ &\quad + \lambda_1\phi_1(T,h_1) + \lambda_2\phi_1(T,h_2), \end{aligned} \tag{24}$$

where $\lambda_1 = L_1/N, \lambda_2 = L-2/N, N = L_1+L_2$. For the two-season fitness landscape, we have the following expression for the mean fitness:

$$\begin{aligned} R(T_1+T_2) &= -\lambda_1(T_1h_1m_1+T_2h_2m_2) - \lambda_2(T_1h_3m_3+T_2h_4m_4)\\ &\quad + \phi_2(T_1,h_1,T_2,h_2) + \phi_2(T_1,h_3,T_2,h_4)\\ &\quad + T_1f_1(m_1,m_3) + T_2f_2(m_2,m_4). \end{aligned} \tag{25}$$

We should focus on the maximum of the latter expression. Thus for the general case of two dimensional two-season evolution, we have a system of four nonlinear algebraic equations.

VI. CONCLUSIONS

Both evolutionary dynamics and rhythms have played a key role in the functioning of all forms of life. Hence, it is important to investigate the evolution on a periodic landscape and solve for such a model involving both features. We have formulated the Crow-Kimura model with fitness landscape that is periodic in time and deduced an exact integral equation for a periodic fitness landscape with time-ordered exponent. We have also formulated the two-season evolution model and derived a simple approximate solution. Our analytical results are well corroborated by numerical results as they show excellent agreement. Our work has great relevance to the quasispecies model which has far more applications than the traditional population genetics [24], for example, in the modeling of viruses [25] to cancer biology, genetic algorithms and learning [26], and evolvability [27]. We perceive these results to be of general significance for the understanding and management of seasons-related infections, too. Critically, we have obtained the first result to quasispecies involving a smooth fitness landscape. The present study suggests a framework for developing mathematical tools to realistically capture the discontinuous characteristics in Hamilton-Jacobi equation theory. A similar method may also be applied to the two-season Eigen model.

ACKNOWLEDGMENTS

D.B.S. acknowledges financial support from the Russian Science Foundation Grant No. 19-11-00008. K.H.C. acknowledges support from the SUTD Start-up Research Grant No. SRG SCI 2019 142.

APPENDIX: THE DETERMINATION OF ϕ_2

We should investigate the operator

$$e^{T_1(h\sigma_z+\sigma_x)}e^{T_2(v\sigma_z+\sigma_x)}. \tag{A1}$$

Let us introduce the unit vectors

$$\vec{a} = \left(\frac{h}{\sqrt{1+h^2}}, \frac{1}{\sqrt{1+h^2}}\right), \quad \vec{b} = \left(\frac{v}{\sqrt{1+v^2}}, \frac{1}{\sqrt{1+v^2}}\right), \tag{A2}$$

$$\begin{aligned} \exp[T_1(h\sigma_x+\sigma_z)] &= \cosh(T_1\sqrt{h^2+1})[1+\tanh(T_1\sqrt{h^2+1})(\vec{a}\cdot\vec{\sigma})]\exp[T_2(v\sigma_x+\sigma_z)]\\ &= \cosh(T_1\sqrt{v^2+1})[1+\tanh(T_1\sqrt{v^2+1})(\vec{b}\cdot\vec{\sigma})]. \end{aligned} \tag{A3}$$

We should find the maximum eigenvalue of the matrix,

$$\begin{aligned} \exp[T_1(h\sigma_x+\sigma_z)]\exp[T_2(v\sigma_x+\sigma_z)] &= \cosh(T_1\sqrt{h^2+1})\cosh(T_2\sqrt{v^2+1})[1+\tanh_1\tanh_2(\vec{a}\cdot\vec{b})+\tanh_1\tanh_2 i(a_1b_2-a_2b_1)\sigma_y]\\ &\quad + (\tanh_1\vec{a}+\tanh_2\vec{b})\vec{\sigma} = \cosh(T_1\sqrt{h^2+1})\cosh(T_2\sqrt{v^2+1})X + Yi\sigma_y + Z(\vec{c}\vec{\sigma}),\\ X &= 1+\tanh_1\tanh_2(\vec{a}\cdot\vec{b}), \quad Y = \tanh_1\tanh_2(a_1b_2-a_2b_1),\\ Z &= \sqrt{(\tanh_1)^2+(\tanh_2)^2+2(\tanh_1)(\tanh_2)(\vec{a}\cdot\vec{b})},\\ \vec{c} &= \frac{\tanh_1\vec{a}+\tanh_2\vec{b}}{\sqrt{(\tanh_1)^2+(\tanh_2)^2+2(\tanh_1)(\tanh_2)(\vec{a}\cdot\vec{b})}}. \end{aligned} \tag{A4}$$

Then we use

$$\exp\{u[(\vec{n}\cdot\vec{\sigma})+si\sigma_y]\} = \exp\left[u\sqrt{1-s^2}\left(\frac{\vec{n}\cdot\vec{\sigma}}{\sqrt{1-s^2}}+\frac{si}{\sqrt{1-s^2}}\sigma_y\right)\right]$$

$$= \cosh(u\sqrt{1-s^2})\left[1+\tanh(u\sqrt{1-s^2})\left(\frac{\vec{n}\cdot\vec{\sigma}}{\sqrt{1-s^2}}+\frac{si}{\sqrt{1-s^2}}\sigma_y\right)\right]. \tag{A5}$$

Finally, using Eq. (A5), we calculate

$$u = A\tanh(\sqrt{Z^2-Y^2}/X). \tag{A6}$$

Thus we obtain for large n,

$$\phi_2(T_1,h,T_2,v) = \ln(\mathrm{Tr}A^n)^{1/n} = \ln\frac{\cosh(T_1\sqrt{h^2+1})\cosh(T_2\sqrt{v^2+1})X}{\cosh(u)} + u. \tag{A7}$$

[1] M. Eigen, Naturwiss. **58**, 465 (1971).

[2] M. Eigen, J. McCaskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).

[3] M. Nilsson and N. Snoad, Phys. Rev. Lett. **84**, 191 (2000).

[4] C. Wilke, C. Ronnewinkel, and T. Martimez, Phys. Rep. **349**, 395 (2001).

[5] M. Ancliff and J.-M. Park, Phys. Rev. E **80**, 061910 (2009).

[6] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper Row, New York, 1970).

[7] E. Baake, M. Baake, and H. Wagner, Phys. Rev. Lett. **78**, 559 (1997).

[8] E. Baake and H. Wagner, Genet. Res. **78**, 93 (2001).

[9] J. Hermisson, O. Redner, H. Wagner, and E. Baake, Theor. Popul. Biol. **62**, 9 (2002).

[10] D. B. Saakian and C. K. Hu, Phys. Rev. E **69**, 021913 (2004).

[11] A. G. M. Neves, Phys. Rev. E **82**, 031915 (2010).

[12] A. O. Bergland *et al.*, PLoS Genet **10**, e1004775 (2014).

[13] E. L. Behrman, V. M. Howick, M. Kapun, F. Staubach, A. O. Bergland, D. A. Petrov, B. P. Lazzaro, and P. S. Schmidt, Proc. R. Soc. Lond. B **285**, 1870 (2018).

[14] M. J. Wittmann, A. O. Bergland, M. W. Feldman, P. S. Schmidt, and D. A. Petrov, Proc. Natl. Acad. Sci. USA **114**, E9932 (2017).

[15] A. O. Bergland, M. Agotsch, D. Mathias, W. E. Bradshaw, and C. M. Holzapfel, J. Ecol. Entomol. **30**, 129 (2005).

[16] C. J. Thompson and J. L. McBride, Math. Biosci. **21**, 127 (1974).

[17] D. B. Saakian, C. K. Hu, and H. Khachatryan, Phys. Rev. E **70**, 041908 (2004).

[18] D. B. Saakian, J. Stat. Phys. **128**, 781 (2007).

[19] K. Sato and K. Kaneko, Phys. Rev. E **75**, 061909 (2007).

[20] D. B. Saakian, O. Rozanova, and A. Akmetzhanov, Phys. Rev. E **78**, 041908 (2008).

[21] D. B. Saakian and C.-K. Hu, Proc. Natl. Acad. Sci. USA **103**, 4935 (2006).

[22] D. B. Saakian, Z. Kirakosyan, and C.-K. Hu, Phys. Rev. E **86**, 031920 (2012).

[23] D. B. Saakian, E. Munoz, C.-K. Hu, and M. W. Deem, Phys. Rev. E **73**, 041913 (2006).

[24] N. Goldenfeld and C. Woese, Annu. Rev. Condens. Matt. Phys. **2**, 375 (2011).

[25] M. Eigen, Proc. Natl. Acad. Sci. USA **99**, 13374 (2002).

[26] V. G. Red'ko, Procedia Comput. Sci. **71**, 215 (2015).

[27] D. J. Earl and M. W. Deem, Proc. Natl. Acad. Sci. USA **101**, 11531 (2004).