Replicator dynamics with diffusion on multiplex networks

R. J. Requejo[*] and A. Díaz-Guilera

Departament de Física Fonamental, Universitat de Barcelona, Martí i Franques 1, 08028 Barcelona, Spain

In this study we present an extension of the dynamics of diffusion in multiplex graphs, which makes the equations compatible with the replicator equation with mutations. We derive an exact formula for the diffusion term, which shows that, while diffusion is linear for numbers of agents, it is necessary to account for nonlinear terms when working with fractions of individuals. We also derive the transition probabilities that give rise to such macroscopic behavior, completing the bottom-up description. Finally, it is shown that the usual assumption of constant population sizes induces a hidden selective pressure due to the diffusive dynamics, which favors the increase of fast diffusing strategies.

I. INTRODUCTION

During the last decade agent-based modeling has increased its importance as a powerful tool to model situations in which the complexity of the interactions of many-agent systems makes it impossible, or at least very difficult, to make analytic predictions of the dynamical behavior of the system [1,2]. Furthermore, the agent-based models complement the analytic approach allowing for an exploration of the coarse-grained dynamics and the connection between microscale and macroscale behavior [3–6]. Such agent-based modeling is of special importance in evolutionary game theoretical studies [7,8], which aim to capture the intricacies of biological, social, and economical systems, where the nonlinearity and feedbacks of the systems cannot be easily foreseen [9–12]. In the middle of such framework [13], and connected with the microevolutionary dynamics [3–6], stands the replicator equation [14,15]

$$(\dot{x}^\alpha)_{\text{rep}} = x^\alpha (f^\alpha - \bar{f}). \tag{1}$$

The replicator equation, which was introduced shortly after the foundation of the evolutionary game theoretical framework [7,16,17], has been extensively used during the last decades to model the evolution of the fractions $x^\alpha = n^\alpha/N$ of traits of type $\alpha = 1, \ldots, L$ in large well-mixed populations with frequency dependent selection, i.e., when the fitness f^α (reproductive potential or capacity) of the agents traits and the mean population fitness $\bar{f} = \sum_\alpha f^\alpha x^\alpha$ depend on the population composition. Such traits may represent different phenotypes and genotypes in biological settings, or different behavioral strategies in a cultural evolutionary framework.

The replicator equation is on the core of the framework of evolution, linked to the quasispecies equation [13], the game dynamical equation [3,4], adaptive dynamics [13], Lotka-Volterra dynamics [18,19], and the Price equation [20–23]. During the last years the physics community has analyzed how such equation is affected by physical entities, as the interplay between different temporal scales of interaction and selection [24,25], the self-organized homeostatic regulation [26] triggered by constraints on resource availability [9,10,12,27], or the existence of network structures [25,28–31]. Further-

more, an extension of the replicator dynamics to regular networks has been developed in a situation where each node is an agent [32], and a second extension of the replicator dynamics has been developed to represent a two-dimensional world where the agents diffuse [33]. However, many real-world situations cannot be modeled as a simple network or a two-dimensional space, and need the introduction of several kinds of links to represent different properties of the system [34], as in air transportation networks, where each airline represents a different network [35].

Let us focus on a cultural evolutionary framework for the remainder of the paper. In such a framework the individual traits being selected are behavioral traits, called strategies. The main aim of this paper is to develop the mathematical tools that allow us to model diffusion of strategies in multiplex networks in a compatible way with the selection dynamics described by the replicator equation (with or without mutations), irrespective of the microscopic dynamics that give rise to the replicator equation (see Appendix A). Let us provide a simple example to illustrate the situation: imagine several cities that are connected by bus, train, and plane, each kind of connection with a different network structure. Individuals in each city are interacting between them, both directly or indirectly: they have different jobs and different incomes, which determines, at least partially, their choices on how to travel, and they also have information about the transportation ways. Based on such interactions and information, they may decide to travel in one or another transport, which determines their individual strategy. As the different transportation ways determine different networks structures, as well as the travel speeds are different, a multilayer network is necessary to represent the full transportation system; hence, as a first approach to such simplified situation, we need to extend the replicator dynamics (describing strategy changes) to diffusive individuals in this kind of multilayer.

We may think that introducing diffusion on multiplex models of evolution is a straightforward task, as the diffusive process has been studied on detail on such networks [36,37]. However, it poses a challenge: as the replicator dynamics describe the evolution of fractions of individuals, the diffusive models need to be rewritten in a compatible way, accounting for the constraint on the addition of the fractions to one, as well as for the conservation of the number of agents. We develop such extension in this paper, showing that working

_____
[*]rrequejo@ffn.ub.edu

with fractions introduces dependencies, which are not present for the diffusion of the numbers of agents. These new dependencies can only be taken into account introducing a nonlinear term, which has to be added to the linear one, in order to represent the general dynamics of diffusion of fractions of individuals in the multiplex. Furthermore, we discuss some situations in which the linear scenario can be recovered, as when population sizes or population size ratios between sites are constant, and show that in such situations hidden selective pressures act on the system, even when they do not appear explicitly on the equations.

The paper is outlined as follows. We start in Sec. II by defining multiplex networks and showing the problem to overcome when working with fractions. After that, in Sec. III we derive the diffusion term compatible with the replicator dynamics and infer the transition probabilities that give rise to it. Then, in Sec. IV we discuss some situations in which the linear dynamics are recovered (and the analytic calculations simplified), and show that some of this situations carry attached the appearance of hidden selective pressures. Finally, in Sec. V we discuss the results.

II. DIFFUSION IN THE MULTIPLEX: THE PROBLEM OF WORKING WITH FRACTIONS

In a general multilayer network [34] each node in one layer can be connected to any node in any other layer. Hence, the connectivity may be given by a tensor with four indices, $M_{ij}^{\alpha\beta}$, whose entries are 1 if nodes i^α—node i in layer α—and j^β are connected and zero otherwise. In multiplex networks [36–40], however, the set of nodes $i = 1, \ldots, N$ is the same in all layers $\alpha = 1, \ldots, L$, and the connectivity of each layer is given by a matrix $A^\alpha = \{a_{ij}^\alpha\}$ (see Fig. 1). Furthermore, only connections among one node i^α and its counterpart in another layer i^β are allowed. This interlayer connectivity is the same for all nodes and is given by the interlayer connectivity matrix $\Lambda^{\alpha\beta}$ [37]. Hence, multiplex networks have a connectivity defined by

$$M_{ij}^{\alpha\alpha} = a_{ij}^\alpha \quad M_{ii}^{\alpha\beta} = \Lambda^{\alpha\beta} \quad M_{ij}^{\alpha\beta} = 0 \quad (2)$$
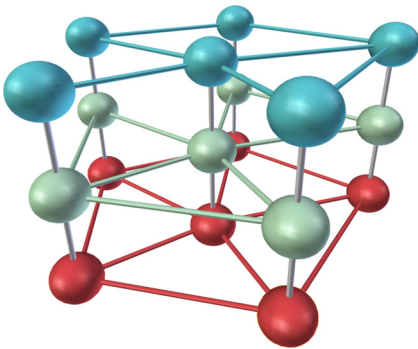
with $i \neq j$ and $\alpha \neq \beta$.

FIG. 1. Example of multiplex network formed by three layers (blue, gray, and red). The set of nodes is the same in all layers and the interlayer links connect them in a one to one basis. For clarity, the interlayer connection between the first and third layer has been omitted.

Each site i of a multiplex may be regarded as one structural entity, and the different layers α represent different interconnection structures between these entities. Structural entities (sites) may represent specific locations (cities), and the different connectivity of each transportation system or mobility pattern would define the different layers through which the agents move, while the interlayer connections refer to the possibility to reach another transport (layer) from a specific location (for simplicity we will assume that $\Lambda^{\alpha\beta} = 1$). Individuals changing from one layer to another can be represented as changing their strategy due to selection (due to prices, availability, or other competitive reasons) or mutation (random trial) in an evolutionary framework.

In order to describe the state of the entire multiplex system, we may define a set of vectors $\{\vec{n}_i\}$, one per site i, each one with L (number of layers) components. From an evolutionary game theoretical point of view the components of such vectors represent the number of agents n_i^α in a given position i^α of the multiplex, and hence $\vec{n}_i$ is the population composition at i. Then, the evolution of the state of node i^α may be assumed as given by a functional $F_i^\alpha[\{\vec{n}_i(t)\}, a_{ij}^\alpha, \Lambda^{\alpha\beta}]$, which depends, respectively, on the state of the system, and on the intralayer and interlayer connectivities.

As the state of the system is instantaneously defined by a set of quantities $\vec{n}_i$ for each site i, we may write the dynamics as a set of coupled differential equations, one per component (layer),

$$\frac{dn_i^\alpha(t)}{dt} = F_i^\alpha[n_i^\alpha(t), a_{ij}^\alpha n_j^\alpha(t), \Lambda^{\alpha\beta} n_i^\beta(t)], \quad (3)$$

where the terms on which the functional depends are, respectively, the state of i^α, the state of connected nodes in the same layer (intralayer neighborhood of i^α), and the state of equivalent nodes connected through interlayer connections (interlayer neighborhood of i^α).

In order to approach a replicator-equation-like functional [see Eq. (1)], we have to normalize n_i^α with respect to $N_i = \sum_\alpha n_i^\alpha$, the number of agents (population size) on site i, which is only site dependent. The normalized quantity is the fraction $x_i^\alpha = n_i^\alpha/N_i$, which must fulfill the restriction $\sum_\alpha x_i^\alpha = 1$, and hence its derivative satisfies $\sum_\alpha \dot{x}_i^\alpha = 0$. Note that, whenever the agents diffuse from j^β to i^β (or in the opposite direction), they are modifying the value of the fraction in i^α through the modification of the population size N_i. Hence, in order to account for the dynamics in a position i^α it is no longer enough to take into account the direct neighborhood of i^α, given by its connectivity, but it is necessary to take into account the entire neighborhood of i, introducing an extra dependence. The functional in terms of fractions is thus

$$\frac{dx_i^\alpha(t)}{dt} = F_i^\alpha[\vec{x}_i(t), a_{ij}^\alpha x_j^\alpha(t), \Lambda^{\alpha\beta} x_i^\beta(t), a_{ij}^\beta x_j^\beta(t)], \quad (4)$$

where the extra dependencies are given by the first term between brackets, which now accounts for the entire state of i, and the last term, which accounts for the neighborhood of i^β. As we show in the following, such extra dependencies require a modification of the diffusion term, which will no longer be linear, but include a nonlinear term.

III. REPLICATOR DYNAMICS WITH DIFFUSION IN THE MULTIPLEX

In this section we derive a diffusion term in the multiplex compatible with the replicator dynamics, i.e., describing the evolution of the fractions of agents x_i^α at site i with diffusive pattern given by layer α, and keeping the conservation of the total number of agents in the multiplex, $\sum_i N_i = N$ where N is constant, as well as the constraint $\sum_\alpha x_i^\alpha = 1$. We will assume that diffusion and evolution are uncoupled, and hence the diffusion term can simply be added to the dynamics.

As the diffusion of agents has been already studied, let us start trying to derive the equations for diffusion of the fractions from those for the number of agents, and discuss its use in game theoretical studies. If the agents are diffusing across a network structure with adjacency matrix of layer α given by the elements $a_{ij}^\alpha = 1$ for connected nodes and 0 otherwise, the transition probabilities determining the microscopic dynamics of diffusion of (numbers of) individuals i given j in layer α are

$$\begin{aligned}
\left(T_{i|j}^{+\alpha}\right)_{\text{diff}} &= D^\alpha a_{ij}^\alpha n_j^\alpha = D^\alpha a_{ij}^\alpha N_j x_j^\alpha \\
\left(T_{i|j}^{-\alpha}\right)_{\text{diff}} &= D^\alpha a_{ij}^\alpha n_i^\alpha = D^\alpha a_{ij}^\alpha N_i x_i^\alpha,
\end{aligned} \quad (5)$$

where D^α is the diffusion coefficient of layer α and

$$T_{i|j}^{+\alpha} = T\left[n_i^\alpha \to n_i^\alpha + 1 \,\middle|\, n_j^\alpha \to n_j^\alpha - 1\right] \quad (6)$$

is the transition rate of increase (decrease, by changing all signs) of the number of agents in i^α in one unit triggered by the movement of one agent in j^α. Note that, as the process of diffusion implies a redistribution of agents, the aggregated number of agents $N_i + N_j$ is preserved in each diffusive event among i and j, and hence the constraint $(T_{i|j}^{+\alpha})_{\text{diff}} = (T_{j|i}^{-\alpha})_{\text{diff}}$ holds, which also ensures the global conservation of agents in the entire multiplex system by linking the processes $n_i^\alpha \to n_i^\alpha + 1$ and $n_j^\alpha \to n_j^\alpha - 1$. In the following, for simplicity, we will write the transition rates in a simplified form, explicitly stating the process to which it refers in the focal variables and the terms involved, but not the possible linked processes, which will be pointed out when they are relevant for the discussion; in this way, Eq. (6) would be $(T_{i|j}^{+\alpha})_{\text{diff}} = T[n_i^\alpha \to n_i^\alpha + 1 \,|\, n_j^\alpha]$.

The microscopic dynamics can be connected with the macroscopic behavior of the system by expanding a Fokker-Planck equation and truncating high-order terms (this happens naturally for $N \gg 1$ when working with fractions, see Appendix A for the one-dimensional derivation), which results in the Langevin equation,

$$\dot{\eta}_i^\alpha = e + \xi s, \quad (7)$$

where ξ is uncorrelated Gaussian noise, η is either the number of individuals or the fraction of individuals, and the drift and diffusion terms (do not confuse the latter with the diffusive dynamics studied in this section) are respectively

$$e = \sum_j \left(T_{i|j}^{+\alpha} - T_{i|j}^{-\alpha}\right), \quad s = \sqrt{\frac{\sum_j \left(T_{i|j}^{+\alpha} + T_{i|j}^{-\alpha}\right)}{N_i}}, \quad (8)$$

where the transition probabilities have to be written in terms of numbers or fractions depending on the choice in Eq. (7). The first term in Eq. (8) (drift) accounts for the deterministic

behavior of the system and the latter (diffusion term) for stochastic effects. Note that the stochastic effects disappear in the thermodynamic limit $N_i \to \infty$, or whenever $N_i \gg \sum_j(T_{i|j}^{+\alpha} + T_{i|j}^{-\alpha})$, i.e., when the transition rates are small compared to the population size.

Whenever we introduce the transition probabilities Eq. (5) into the drift term in Eq. (8) expressed for numbers of individuals, the deterministic diffusive dynamics for the number of agents are

$$\left(\dot{n}_i^\alpha\right)_{\text{diff}} = D^\alpha \sum_j a_{ij}^\alpha \left(n_j^\alpha - n_i^\alpha\right) = -D^\alpha \sum_j L_{ij}^\alpha n_j^\alpha, \quad (9)$$

a well-known equation for the diffusion of particles on a network, where the tensor $L^\alpha = \{L_{ij}^\alpha\} = \{\delta_{ij} k_i^\alpha - a_{ij}^\alpha\}$ is the graph Laplacian of the corresponding layer α (with $\delta^{\alpha\beta}$ the Kronecker's δ) and k_i^α is the degree of site i in layer α.

In order to derive the equation for the diffusion of fractions instead of numbers, it is important to note that, if we differentiate the fraction $x_i^\alpha = n_i^\alpha/N$, we get

$$\dot{x}_i^\alpha = \frac{\dot{n}_i^\alpha}{N_i} - x_i^\alpha \frac{\dot{N}_i}{N_i} \quad (10)$$

and hence a nonlinear term that depends on x_i^α appears. Since we are trying to find the exact description of the diffusive process which is compatible with the replicator equation for mobile agents in a multiplex, we can follow this approach: first, construct a compatible macroscopic equation, and then infer the transition probabilities that give rise to it. In order to do this we may introduce Eq. (9) into Eq. (10) [note that the latter is a replicator equation of the form of Eq. (A1)], and use $\sum_\beta n_i^\beta = N_i$, obtaining a term of the form,

$$\left(\dot{x}_i^\alpha\right)_{\text{diff}} = -D^\alpha \sum_j \rho_{ij} L_{ij}^\alpha x_j^\alpha + x_i^\alpha \sum_\beta \sum_j D^\beta \rho_{ij} L_{ij}^\beta x_j^\beta, \quad (11)$$

where $\rho_{ij} = N_j/N_i$ is a population size dependence between neighboring sites. As it can be observed, in addition to a linear term, the first one, a second term appears. This term ensures that the normalization of the fractions is the proper one, i.e., fractions always add up to one, as shown in Figs. 2(a), 2(b) for a system of two layers and two nodes in each layer. Note that, as previously foreseen in Sec. II, this term implies a dependence of the dynamics at i^α on positions j^β to which it is not directly connected, but represents the neighborhood of i^β nodes.

The diffusion term (11) may be rewritten in a more compact form

$$\left(\dot{x}_i^\alpha\right)_{\text{diff}} = -\sum_\beta \sum_j D^\beta x_j^\beta \rho_{ij} \left(\delta^{\alpha\beta} - x_i^\alpha\right)\left(k_i^\beta \delta_{ij} - a_{ij}^\beta\right). \quad (12)$$

Note that, if one extracts the node degree from the second parenthesis—the network Laplacian term—then it becomes $k_i^\beta(\delta_{ij} - a_{ij}^\beta/k_i^\beta)$, which is similar to the first parenthesis term, $\rho_{ij}(\delta^{\alpha\beta} - x_i^\alpha)$, but with constant values instead of variables. In this way the latter term may be interpreted as a population-composition-dependent Laplacian, which accounts for the instantaneous heterogeneity in population sizes and fractions of individuals in the network.
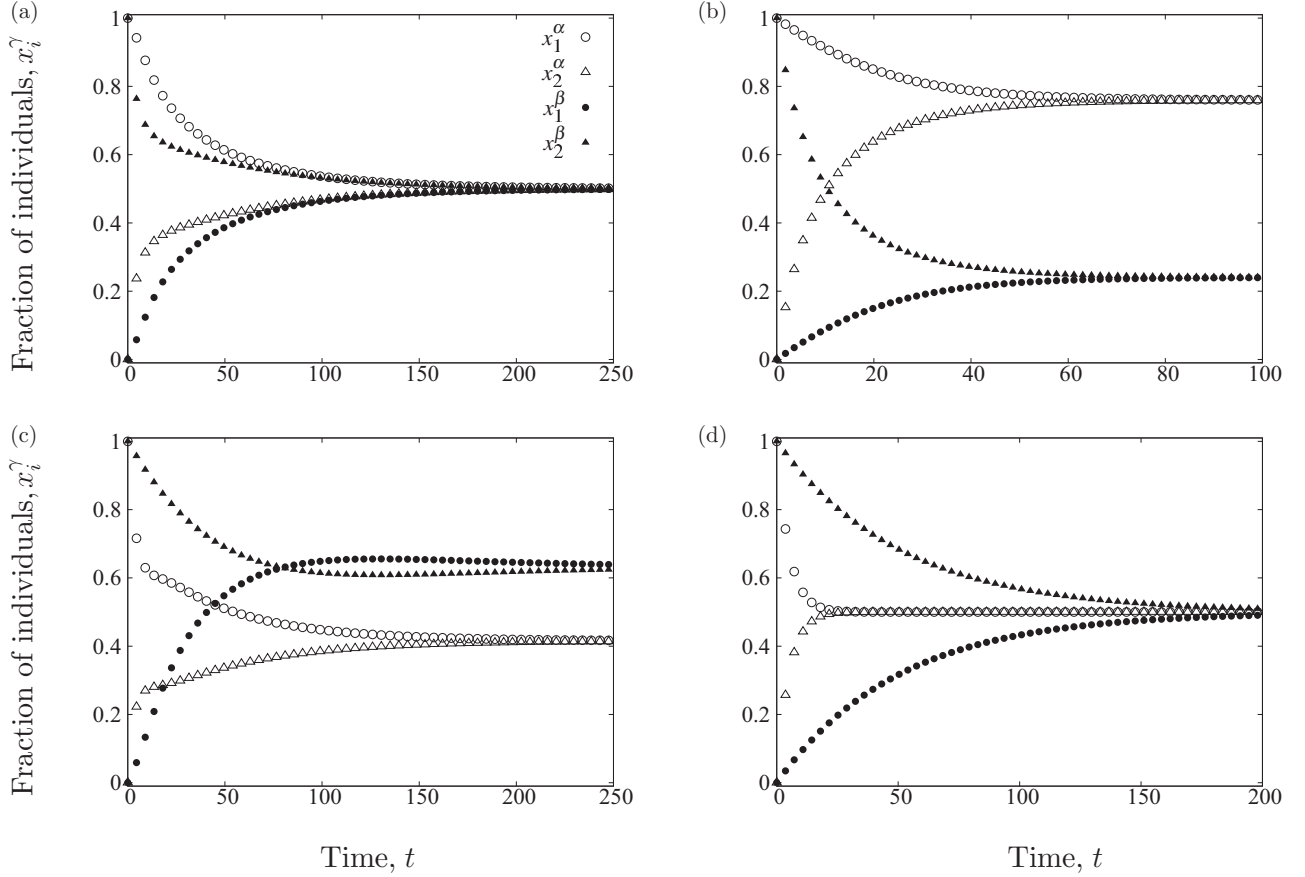
FIG. 2. Intralayer diffusive dynamics of a system consisting of two layers with diffusion coefficients $D^\alpha = 10^{-1}$ and $D^\beta = 10^{-2}$, and two sites, $i = 1, 2$. Initially there are N individuals in nodes 1^α and 2^β, which are allowed to diffuse within its own layer (note that N may take any value in the situation depicted, provided that it is the same in both layers, as the fractional character of ρ makes it vanish); no interlayer process is acting on the system. In order to fulfill the constraint $\sum_\alpha x_i^\alpha = 1$, filled and empty circles, as well as filled and empty triangles, have to add up to one. (a), (b) dynamics of Eq. (11), with (a) ρ_{ij} calculated exactly analytically and (b) setting $\rho_{ij} = 1$; (b), (c) dynamics of Eq. (17) with (c) ρ_{ij} calculated exactly analytically and (d) $\rho_{ij} = 1$; Cases (c) and (d) do not keep the normalization of the fractions. The dynamics given in (a) and (b) maintain the restriction $\sum_\gamma x_i^\gamma = 1$ due to the quadratic term in Eq. (11). However, while case (a) describes accurately the expected dynamics, (b) shows an asymmetry in the final state, which is not expected due to pure intralayer diffusion, but the result of an induced evolutionary pressure due to the different diffusive velocities and the restriction $\rho_{ij} = 1$.

Now, it is possible to infer the transition probabilities of the microscopic process from the emergent macroscopic dynamics [Eq. (12)], which may be used in Markovian analysis [12]. The macroscopic deterministic dynamics emerge from the drift term in Eq. (8), which by extrapolation to two dimensions results in

$$\dot{x}_i^\alpha = \sum_\beta \sum_j \left(T_{i|j}^{+\alpha|\beta} - T_{i|j}^{-\alpha|\beta} \right), \tag{13}$$

where

$$T_{i|j}^{+\alpha|\beta} = T\left[x_i^\alpha \to x_i^\alpha + \Delta x_i^\alpha \,\middle|\, x_j^\beta \right] \tag{14}$$

is the transition rate of increase (decrease, by changing all signs) of the fraction of agents in i^α given the fraction of agents in j^β (see Appendix C). Note that x_j^β may also vary in the same process, but such variation is specified in its related transition rate, and not explicitly written here for simplicity.

Then, we can compare Eq. (12) and Eq. (13) (see Appendix C for the derivation), and infer the bidimensional transition probabilities describing the variation of the fractions

of individuals x_i^α, which give rise to Eq. (12), resulting in

$$
\begin{aligned}
T_{i|i}^{+\alpha|\alpha} &= 0 \quad & T_{i|i}^{-\alpha|\alpha} &= D^\alpha k_i^\alpha x_i^\alpha \left(1 - x_i^\alpha \right) \\
T_{i|j}^{+\alpha|\alpha} &= D^\alpha \rho_{ij} a_{ij}^\alpha x_j^\alpha \left(1 - x_i^\alpha \right) \quad & T_{i|j}^{-\alpha|\alpha} &= 0 \\
T_{i|i}^{+\alpha|\beta} &= D^\beta k_i^\beta x_i^\alpha x_i^\beta \quad & T_{i|i}^{-\alpha|\beta} &= 0 \\
T_{i|j}^{+\alpha|\beta} &= 0 \quad & T_{i|j}^{-\alpha|\beta} &= D^\beta \rho_{ij} a_{ij}^\beta x_i^\alpha x_j^\beta,
\end{aligned}
\tag{15}
$$

where we have omitted the subscript "diff" for simplicity.

Note that the transition probabilities depend on both the layers (α) and the sites (i), and that the symmetry, which is present for mutation transition probabilities, which makes them keep local population sizes constant (see Appendix B), is broken in general settings, $T^{-\alpha|\beta} \neq T^{+\beta|\alpha}$. Such asymmetry is however expected, as diffusion implicitly needs variable local population sizes (although the total number of particles in the multiplex is conserved), and assuming constant population sizes locally may have unexpected effects, as shown in the next section.

The transition probabilities in Eq. (15), together with those that give rise to the replicator equation with mutations (Appendixes A, B) complete the microscopic description of the evolutionary process for agents in a multiplex. Let us finally write down the deterministic replicator dynamics (see Appendix A) with additive mutations (see Appendix B) and diffusion in a multiplex:

$$\dot{x}_i^\alpha = x_i^\alpha \left(f_i^\alpha - \bar{f}_i \right) + \sum_\beta \left(x_i^\beta q_i^{\beta\alpha} - x_i^\alpha q_i^{\alpha\beta} \right)$$
$$- \sum_\beta \sum_j D^\beta x_j^\beta \rho_{ij} \left(\delta^{\alpha\beta} - x_i^\alpha \right) \left(k_s^\beta \delta_{ij} - a_{ij}^\beta \right), \quad (16)$$

where f_i^α is the fitness of individuals in position i^α, $\bar{f}_i = \sum_\alpha f_i^\alpha$ is the mean fitness of individuals in i across layers and $q_i^{\alpha\beta}$ is the mutation rate, which accounts for transitions of agents between α and β layers in site i. For the case of mutations coupled to reproduction substitute the first two terms in the previous equation by the so called replicator-mutator equation [Eq. (B7)].

Remarkably, the nonlinear effects due to diffusion are all contained in the x_i^α in the first parenthesis of the latter term in Eq. (16). There are, however, some situations in which the nonlinearity disappears, and in which the analytic calculations can be simplified. Let us discuss them, as well as their implications, with special emphasis on the appearance of hidden selective pressures.

IV. RECOVERING LINEAR DIFFUSION AND SIMPLIFYING THE ANALYTICS

Some situations allow us to recover linear diffusion, as well as to simplify the analytic calculations. Here, we analyze three of such scenarios. The first one refers to a situation in which population size ratios are all forced to be constant; note that this includes as a subcase having constant populations, situation when the second term in Eq. (10) disappears. Remarkably, in this situation a hidden selective pressure appears favoring the increase of fast diffusing strategies. The second scenario refers to slow population change, a situation that approximates the constant population sizes scenario. The last subsection explores the situation in which the population ratio can be expressed as some analytic function. In that case, time scales separation allows us to write some formulas, which simplify the calculations.

A. Constant population ratios (or sizes) induce hidden selective pressure

Let us first study the case in which some mechanism makes $\rho_{ij} = N_j/N_i \to c$, where c is a constant; for simplicity, we will assume $c = 1$, which includes the usual assumption in evolutionary game theoretical studies of constant population sizes [3–5,41–43]). This case is shown in Fig. 2(b), where the restriction on the addition of fractions of individuals to one during the dynamical evolution of the system is fulfilled. However, the final state is not a symmetric one in which there are $N/2$ individuals in each node, as it happens without the restriction on the quotient of population sizes. Why does this happen?

The case in which ρ_{ij} is forced to be equal to a certain value in Eq. (11) may happen only if the conservation of particles due to the diffusive process does no longer hold. In such case, which may be due to a fast (compared to diffusion) and neutral evolutionary process acting on the system, the difference in diffusive velocities transforms into different diffusive pressures, which induce different selective pressures while the system decays to the equilibrium.

The different selective pressures are induced by the fact that fast diffusing individuals are increasing their frequency in a site compared to slower diffusing ones, and then the increased fraction is fixated by the neutral selective process, which only renormalizes the population size without altering the proportions. This is similar to a Wright-Fisher process, in which the population reproduces during the reproduction period according to their fitness and then, keeping the proportions of individuals, the population size is renormalized to its initial value. In our case, however, the variation in the fractions is due to the diffusion of the agents and the continuous renormalization of the population size is due to the neutral evolutionary process, thus favoring the fast diffusive strategy, as shown in Fig. 2(b).

B. Slow population size change and quasineutral selection

The equation describing the evolution of the fraction of particles present at each point may in principle be simplified whenever, starting from a situation near the equilibrium, the variation of the number of agents is so slow that it can be neglected, $\dot{N}_i/N_i \to 0$. This is equivalent to assuming that $\dot{x}_i^\alpha \approx \dot{n}_i^\alpha/N_i$, which results in

$$\left(\dot{x}_i^\alpha \right)_{\text{diff}} = D^\alpha \sum_j a_{ij}^\alpha \left(\frac{N_j}{N_i} x_j^\alpha - x_i^\alpha \right) = -D^\alpha \sum_j \rho_{ij} L_{ij}^\alpha x_j^\alpha. \quad (17)$$

However, Eq. (17) does not generally keep the proper normalization for the fractions $\sum_\beta x_i^\beta = 1$, as it can be easily proven with a simple example. Let us assume a multiplex network formed by identical networks with identical diffusion coefficients (for two-dimensional spatial networks, this equals the Fisher-Kolmogorov reaction-diffusion scenario for gene wave-front propagation [44–46]). In this case, which is analogous to a monolayer network, there is no dependence of the adjacency matrix terms and diffusion coefficients on the layer index α, which now serves only to identify the different strategies present in each node. Hence, by summing Eq. (17) over layers (or strategies) and noting that $\sum_\beta x_i^\beta = 1$, we obtain the condition $\sum_\beta \dot{x}_i^\beta = (D/N_i) \sum_j a_{ij}(N_j - N_i) = (D/N_i) \sum_j a_{ij} N_j - k_i N_i$, which is only equal to zero, i.e., satisfies the dynamical constraint, whenever

$$N_i = \frac{\sum_j a_{ij} N_j}{k_i} \quad \text{or} \quad N_i = N_j. \quad (18)$$

These two restrictions are equivalent to requiring that $\sum_j L_{ij} N_j = 0$, which could be implemented or engineered in the system, but it is not *a priori* expected to happen as a self-organizing feature. Furthermore, simulations using a system with two nodes and two layers confirm that the

normalization is not fulfilled using Eq. (17), as shown in Figs. 2(c), 2(d). Hence, the construction of a microscopic model that describes the macroscopic diffusive dynamics of fractions of individuals cannot be done by simply rewriting the transition probabilities in Eq. (5) in terms of fractions.

The case of slow population change is equivalent to assuming that the second term in Eq. (11) vanishes

$$\sum_\beta \sum_j D^\beta \rho_{ij} L_{ij}^\beta x_j^\beta \to 0 \tag{19}$$

and is hence only slightly influencing the dynamics described by the first term. In such case, the diffusion term approaches Eq. (17). Note however that, if we assume that the term above is strictly zero, we recover the case of strictly constant population sizes, and hence hidden selective pressures may appear, as explained in the previous subsection.

From an evolutionary perspective, small perturbations introduced in the neutral selection limit satisfy the conditions leading to the slow population size change approximation whenever diffusion is slow. If we define the quasineutral selection limit as represented by $f_i^\alpha \to 0$, then the population size at each site varies slowly, given that $\dot{N}_i = N_i \bar{f}_i \to 0$ and $D^\gamma \to 0$. The quasineutral selection limit is important for two reasons: first, it allows for analytic calculations to be carried out in a similar way to the weak selection limit [47,48] and, second, this limit approaches the neutral theory of molecular evolution proposed by Kimura [49]. However, care should be taken when approaching this limit, as explained above.

C. Time-scales separation

Let us finish exploring the situations in which ρ_{ij} can be written as a function of the fractions of individuals (or their fitness, which are determined by such fractions once the payoff matrix is known) and time due to time-scales separation.

Whenever we are able to write the population size dependence of the diffusion term as

$$\rho_{ij} = \rho_{ij}(x_k^\gamma, t) \tag{20}$$

the entire replicator dynamics in the multiplex, including replication and deaths, mutations, and diffusion, can be written as a function of the state of the system, given by the fractions of individuals in each node, and of the structural terms (diffusion coefficients and Laplacian of the multiplex).

In the cases in which it is not possible to write an explicit dependence for ρ_{ij} as above, there are at least two situations in which time-scales separation allows for approximations that take such form. The first one is whenever diffusion is slow and most of the population size change is due to replication and death. In this case it is easy to prove that the differential equation

$$\dot{\rho}_{ij} \approx (\bar{f}_j - \bar{f}_i)\rho_{ij} \tag{21}$$

governs the evolution of population sizes. Note that the solution of this equation is an exponential integral,

$$\rho_{ij}(t) = \rho_{ij}^0 e^{\int_0^t (\bar{f}_j - \bar{f}_i)dt'}, \tag{22}$$

which implies that the system has memory. More precisely, the entire history of the difference of mean population fitness differences, which depends on the population compositions,

is influencing the present state. Hence, Eq. (22) is a memory kernel of the past states of the system.

As the memory kernel has an exponential form, although the entire history is contained in it, the influence of past states decays very fast with time compared to present states. This can be easily proved noting that, given two time lapses beginning at $t = 0$ and ending at t_1 and $t_2 > t_1$, the memory kernels are $\rho_{ij}(t_1) = \rho_{ij}^0 e^{\int_0^{t_1}(\bar{f}_j - \bar{f}_i)dt'}$ and $\rho_{ij}(t_2) = \rho_{ij}^0 e^{\int_0^{t_1}(\bar{f}_j - \bar{f}_i)dt'} e^{\int_{t_1}^{t_2}(\bar{f}_j - \bar{f}_i)dt'}$. The quotient between them is $\rho_{ij}(t_2)/\rho_{ij}(t_1) = e^{\int_{t_1}^{t_2}(\bar{f}_j - \bar{f}_i)dt'}$, which is a memory kernel of the time lapse between t_1 and t_2 and does not take into account the time lapse between 0 and t_1. Hence, we may always make $t_2 = t_1 + dt$ to express the quotient as an instantaneous integral of the fitness differences, which allows for computation of the dynamics without keeping track of the entire history of the system.

The opposite limit to slow diffusion is the fast diffusion limit: In this case we may assume that, after a short transient, the population size variations fulfill $\dot{N}_i/\dot{N}_j = \kappa_{ij}(t) \neq 0,\infty$; then the approximation

$$\rho_{ij} = (\bar{f}_i/\bar{f}_j)\kappa_{ij}(t) \tag{23}$$

can be used. Furthermore, in the fast diffusion limit we may assume that the population size variations due to replication and death may be negligible compared to diffusion, as in the previous subsection; in such case the conditions in Eq. (18)—equal populations across all sites or a local mean field describing the population sizes—may be prone to happen, although again, this may introduce hidden selective pressures.

V. DISCUSSION AND CONCLUSIONS

When dealing with replicator dynamics the equations are written in terms of the fractions of strategies in the population in every node, although diffusion is usually described in terms of the populations themselves. We have solved this arbitrariness by writing the complete diffusion equations for the fractions, but this makes the diffusion process have nonlinear contributions. On the one hand, it introduces the ratio of population sizes, which is clearly state dependent. On the other hand, it generates an additional quadratic term in the population fractions.

The extra nonlinear diffusive term in Eq. (11) takes into account the state of the vicinity of the focal node and its equivalent nodes in all layers (extended neighborhood), and not only its direct neighbors (directly connected nodes). This is relevant for the calculation of the evolution of the fractions of individuals at each site, and contrasts with the usual diffusion term, which only takes into account the state of directly connected nodes within each layer.

We have finally explored the recovery of the linear scenario, finding some situations in which the analytics are simplified. One of such scenarios happens whenever population sizes are constant across sites (due to environmental saturation, for instance). In that case the nonlinearity disappears, but the diffusion process induces an extra evolutionary pressure acting while the system is out of equilibrium. This has been argued to happen because a hidden selective process induces a constant renormalization of the population size, which favors

the increase of fast diffusing strategies. This latter result suggests that, depending on the network architecture, the induced evolutionary selective pressures may work so as to create extra gradients of selection acting while the system is out of equilibrium, which may induce, depending on the multiplex architecture, an unexpectedly complex phenomenology.

APPENDIX A: MICROSCOPIC LOCAL DERIVATION OF THE REPLICATOR EQUATION

The replicator equation in a well-mixed population [no network structure, Eq. (1)] can be derived—for large population sizes accepting smooth derivatives—by differentiating the fraction of individuals x^α,

$$\frac{d(x^\alpha)_{\rm rep}}{dt} = \frac{d(n^\alpha/N)_{\rm rep}}{dt} = x^\alpha \left(\frac{\dot{n}^\alpha}{n^\alpha} - \frac{\dot{N}}{N} \right) \qquad (A1)$$

and assuming that the fitness of the individuals corresponds to the instantaneous per-capita growth rate of each strategy due to replication and deaths [41],

$$f^\alpha(\mathbf{x}) = \frac{\dot{n}^\alpha}{n^\alpha}, \qquad (A2)$$

where $\mathbf{x} = \{x^\alpha\}$ is the state vector of the population, and the mean fitness of the aggregated population fulfills $\bar{f} = \sum_\alpha x^\alpha f^\alpha = \dot{N}/N$. If the population size is constant, $\dot{N} = 0$, then the per-capita growth rate is proportional to the difference between trait and mean population fitness,

$$f^\alpha(\mathbf{x}) - \bar{f}(\mathbf{x}) = \frac{\dot{n}^\alpha}{n^\alpha}, \qquad (A3)$$

being hence the replicator equation applicable for constant and variable populations with slightly different fitness definitions. As it will be proven in the following, the replicator dynamics also emerge as the macroscopic description of some microscopic dynamics.

Let us start the derivation of the replicator equation by assuming that there is a microscopic process, which can be described by some transition probabilities between states, and that such states are well defined. As we will assume that individuals of different types diffuse through different network architectures, each of such networks being part of a multiplex structure, let us introduce now the related notation: As before, each agent type will be labeled by the superscript α of the layer to which it belongs (related with its strategy), and the subscript i will refer to a site in such layer.

The probability for node i^α to be at time t in a state with n_i^α individuals of α type will be denoted as $P(n_i^\alpha, t)$, and the probability of increasing or decreasing such number

of individuals by one individual will be $T^+(n_i^\alpha)$ and $T^-(n_i^\alpha)$ respectively, where

$$T^+(n_i^\alpha) = T[n_i^\alpha \to n_i^\alpha + 1] \qquad (A4)$$

(similarly for T^- with a sign change). With this notation, it is possible to write the master equation

$$\begin{aligned} P(n_i^\alpha, & t+1) - P(n_i^\alpha, t) \\ &= P(n_i^\alpha - 1, t-1) T^+(n_i^\alpha - 1) + P(n_i^\alpha + 1, t-1) \\ &\quad \times T^-(n_i^\alpha + 1) - P(n_i^\alpha, t-1)[T^-(n_i^\alpha) + T^+(n_i^\alpha)], \end{aligned}$$
$$(A5)$$

which describes the evolution of n_i^α. Now, let us assume that the total number of agents in site i is $N_i = \sum_\alpha n_i^\alpha \gg 1$. We do not require it to be infinite, but just large enough, so that we can make a continuous approach without neglecting finite-size fluctuations. In this case it is possible to define the rescaled variables $x_i^\alpha = n_i^\alpha/N_i$, $\tau = t/N_i$ and $\rho(x_i^\alpha, \tau) = N_i P(n_i^\alpha, t)$. For simplicity, let us assume that there are only two strategies present in the population, and the constraint $\sum_\alpha x_i^\alpha = 1$; we can then expand the master equation in a one-dimensional Taylor expansion for $N_i \gg 1$ (the derivation is similar for the three- [42] and n-strategies cases [43] by using a multivariate Taylor expansion), giving rise to

$$\frac{d}{dt} \rho(x_i^\alpha, t) = -\frac{d}{dx_i^\alpha}[e(x_i^\alpha)\rho(x_i^\alpha, t)]$$
$$+ \frac{1}{2}\frac{d^2}{(dx_i^\alpha)^2}[s^2(x_i^\alpha)\rho(x_i^\alpha, t)]. \qquad (A6)$$

As the previous equation has the form of a Fokker-Planck equation, it is possible to transform it into the Langevin equation

$$\dot{x}_i^\alpha = e(x_i^\alpha) + s(x_i^\alpha)\xi, \qquad (A7)$$

where ξ is uncorrelated Gaussian noise and the drift and diffusion terms are

$$\begin{aligned} e(x_i^\alpha) &= T^+(x_i^\alpha) - T^-(x_i^\alpha), \\ s(x_i^\alpha) &= \sqrt{\frac{T^+(x_i^\alpha) + T^-(x_i^\alpha)}{N_i}}, \end{aligned} \qquad (A8)$$

respectively accounting for the deterministic behavior and the stochastic effects.

Note that whenever the transition probabilities can be written as

$$T^+(x_i^\alpha) = x_i^\alpha R(x_i^\alpha), \quad T^-(x_i^\alpha) = x_i^\alpha D(x_i^\alpha), \qquad (A9)$$

the drift term $e(x_i^\alpha)$ looks like a replicator equation. This is indeed the only term acting in the thermodynamic limit $N \to \infty$, and whenever the size of large populations does not increase or decrease too fast, $N_i \gg x_i^\alpha[R(x_i^\alpha) + D(x_i^\alpha)]$. In this cases the Langevin equation simplifies to

$$\dot{x}_i^\alpha = x_i^\alpha \cdot [R(x_i^\alpha) - D(x_i^\alpha)] \qquad (A10)$$

and the terms $R(x_i^\alpha)$ and $D(x_i^\alpha)$ act as the replication and death components of the fitness difference $f_i^\alpha - \bar{f}_i$ [compare Eqs. (A10) and (1)]. Hence, any process in which the transition probabilities can be factorized as shown in Eq. (A9), can be

written as a replicatorlike system with fitness obtained as the solution of the equation

$$f_i^\alpha - \bar{f}_i = R(x_i^\alpha) - D(x_i^\alpha). \tag{A11}$$

This property has a particularly useful value: by using the latter equation and decomposing fitness in payoff components, it is possible to combine several processes, as virus spread (linear models) and cultural reproduction (frequency dependence, usually nonlinear), into a unified evolutionary framework.

APPENDIX B: MICROSCOPIC LOCAL UPDATING RULES AND MUTATIONS

Whenever there is an arbitrary number of strategies in the population, the transition rates in Eq. (A9) can be decomposed in additive terms in a way in which each term relates to the contribution of each of the strategies. In this way Eq. (A7) can be generalized to

$$\dot{x}_i^\alpha = \sum_\beta \left(T_i^{+\alpha|\beta} - T_i^{-\alpha|\beta} \right), \tag{B1}$$

where $T_i^{+\alpha|\beta}$ is the transition rate of increase of the fractions of individuals in i,α corresponding to the increase in one unit due to the action of agents in β,

$$T_i^{+\alpha|\beta} = T\left[n_i^\alpha \rightarrow n_i^\alpha + 1 \,\middle|\, n_i^\beta \right], \tag{B2}$$

and decreases them by one unit for the minus sign (note that this may imply a variation in the local population size N_i). The exact choice of the transition rates depends on the microscopic dynamics. These transition rates may be written as [42],

$$T_i^{+\alpha|\beta} = \sum_\gamma x_i^\gamma x_i^\beta g_i^{+\gamma|\beta} q_i^{\gamma\alpha}, \quad T^{-\alpha|\beta} = T^{+\beta|\alpha} \tag{B3}$$

whenever the transitions depend on a big number of random interactions between agents in the same site (local well-mixing assumption) in which one individual replaces another (defined by the second condition) and there are mutations between γ and α individuals at a rate $q_i^{\gamma\alpha}$. Factor $g^{\pm\gamma|\beta}$ contains the information about how the microscopic updating rule acts: it states the exact mechanism by which one strategy increases or decreases due to the action of another.

Two probabilistic microscopic dynamics are usually investigated, which give rise to the replicator dynamics (assuming $q_i^{\gamma\alpha} = \delta^{\gamma\alpha}$). The first one is the modified Moran process, where one randomly chosen individual is assumed to die, and another individual, chosen according to a probability proportional to its fitness, reproduces. This process is defined by $g_i^{+\alpha|\beta} = f_i^\alpha / \bar{f}_i$ and $g_i^{-\alpha|\beta} = g_i^{+\beta|\alpha}$. The second process is proportional imitation, where one individual compares its strategy to another one, both chosen at random in the mean field limit, and changes its strategy with a probability increasing linearly with the difference of the payoffs between them. This process is defined by $g_i^{+\alpha|\beta} = (1/2)[1 + (f_i^\alpha - f_i^\beta)/\Delta f_{s,\max}]$ and $g_i^{-\alpha|\beta} = g_i^{+\beta|\alpha}$, where $\Delta f_{s,\max}$ is the maximum fitness difference and keeps the proper normalization. In both cases the symmetry condition $g_i^{-\alpha|\beta} = g_i^{+\beta|\alpha}$ ensures that the evolutionary process maintains a constant population, and the dynamics result in the replicator equation, up to a

multiplicative factor, which relates to the temporal scale of the dynamics.

Whenever mutations happen as a strategy change at any point during the lifetime of the individuals, they can be introduced as additive terms [4] of the form

$$(T^{+\alpha|\beta})_{\text{mut}} = x^\beta q^{\beta\alpha}, \quad (T^{-\alpha|\beta})_{\text{mut}} = (T^{+\beta|\alpha})_{\text{mut}} \tag{B4}$$

to the transition rates in Eq. (B3), where the coupled mutations may be eliminated by setting $q_i^{\alpha\beta} = \delta^{\alpha\beta}$ (with $\delta^{\alpha\beta}$ the Kronecker's δ). The introduction of the additive term in the transition rates gives rise to the extra additive term in the replicator dynamics

$$\left(\dot{x}_i^\alpha \right)_{\text{mut}} = \sum_\beta \left(x_i^\beta q_i^{\beta\alpha} - x_i^\alpha q_i^{\alpha\beta} \right). \tag{B5}$$

This term describes the effect of random mutations or equivalently of random exploration of strategies [50] in the evolutionary process. For the case of equal symmetric mutations, i.e., $q_i^{\alpha\beta} = q_i^{\beta\alpha} = \mu$, between all strategies this term simplifies to

$$\left(\dot{x}_i^\alpha \right)_{\text{mut}} = \mu \left(1 - L x_i^\alpha \right), \tag{B6}$$

where L is the number of strategies [51].

Let us finally recall that, when mutations are coupled to the reproductive dynamics as in Eq. (B3), then Eq. (B1) gives rise to the replicator-mutator equation,

$$\left(\dot{x}_i^\alpha \right)_{\text{rep,mut}} = \sum_\beta x_i^\beta f_i^\beta q_i^{\beta\alpha} - x_i^\alpha \bar{f}_i. \tag{B7}$$

The diffusion term in Eq. (12) could also be added to this equation to represent situations where only newborns mutate.

APPENDIX C: DERIVATION OF THE TRANSITION PROBABILITIES FOR THE NONLINEAR DIFFUSION TERM

The microscopic dynamics that give rise to the nonlinear equation describing the dynamics of diffusion for fractions of individuals [Eq. (12)] is defined by the transition probabilities determining the increase or decrease of the number of individuals, as shown in Eq. (13). In order to infer such transition probabilities, which complete the bottom-up description, we can start expanding Eq. (12) as

$$\left(\dot{x}_i^\alpha \right)_{\text{diff}} = -\sum_\beta \sum_j D^\beta x_j^\beta \rho_{ij} \left(\delta^{\alpha\beta} k_i^\beta \delta_{ij} - \delta^{\alpha\beta} a_{ij}^\beta - k_i^\beta x_i^\alpha \delta_{ij} \right.$$
$$\left. + a_{ij}^\beta x_i^\alpha \right) \tag{C1}$$

Then, we can split the double sum on the four contributions corresponding to the terms arising from i^α, j^α, i^β and j^β (assuming that $\alpha \neq \beta$ and $i \neq j$), obtaining

$$\left(\dot{x}_i^\alpha \right)_{\text{diff}} = -D^\alpha k_i^\alpha x_i^\alpha \left(1 - x_i^\alpha \right) + \sum_{j \neq i} D^\alpha a_{ij}^\alpha \rho_{ij} x_j^\alpha \left(1 - x_i^\alpha \right)$$
$$+ \sum_{\beta \neq \alpha} D^\beta k_i^\beta x_i^\alpha x_i^\beta - \sum_{j \neq i} \sum_{\beta \neq \alpha} D^\beta a_{ij}^\beta \rho_{ij} x_i^\alpha x_j^\beta, \tag{C2}$$

respectively.

By performing the same kind of split in Eq. (13) we find

$$
\dot{x}_i^\alpha = T_{i|i}^{+\alpha|\alpha} - T_{i|i}^{-\alpha|\alpha} + \sum_{j\neq i}\left(T_{i|j}^{+\alpha|\alpha} - T_{i|j}^{-\alpha|\alpha}\right)
$$
$$
+ \sum_{\beta\neq\alpha}\left(T_{i|i}^{+\alpha|\beta} - T_{i|i}^{-\alpha|\beta}\right) + \sum_{j\neq i}\sum_{\beta\neq\alpha}\left(T_{i|j}^{+\alpha|\beta} - T_{i|j}^{-\alpha|\beta}\right).
$$
$$(C3)$$

Then, since the terms relate to the influence of different nodes k^γ on the focal node i^α, we can compare term by term both equations above, finding that each term in the first equation corresponds to the substraction of a pair of terms in the second.

In order to finally find the right transition terms, we need to take into account the physical constraints in the system. Let us analyze them one by one.

First, we can focus on the first term on the above equations. This term relates to the influence of i^α on its own dynamics. Since the agents in i^α are diffusing away from that position, the influence is necessarily negative, and hence $T_{i|i}^{+\alpha|\alpha} = 0$, and

$$
T_{i|i}^{-\alpha|\alpha} = T\left[\frac{n_i^\alpha}{N_i} \to \frac{n_i^\alpha - 1}{N_i - 1}\,\middle|\, x_i^\alpha\right] = D^\alpha k_i^\alpha x_i^\alpha\left(1 - x_i^\alpha\right),
$$
$$(C4)$$

where $|x_i^\alpha$ denotes a dependence, not implying that such term is constant (indeed, it varies in the process). Note also that whenever $x_i^\alpha = 0$, the transition rate is zero, as expected due to the absence of agents diffusing away. Furthermore, if $x_i^\alpha = 1$, the transition rate is again zero, as expected due to the fact that agents diffusing away decrease the number of agents, but leave the fraction unchanged.

Now, let us focus on the influence of j^α on the dynamics of i^α ($j \neq i$). Since agents diffusing away from j^α are increasing the number of agents in i^α, the influence is necessarily positive,

and the transition rates are $T_{i|j}^{-\alpha|\alpha} = 0$ and

$$
T_{i|j}^{+\alpha|\alpha} = T\left[\frac{n_i^\alpha}{N_i} \to \frac{n_i^\alpha + 1}{N_i + 1}\,\middle|\, x_j^\alpha\right] = D^\alpha a_{ij}^\alpha \rho_{ij} x_j^\alpha\left(1 - x_i^\alpha\right).
$$
$$(C5)$$

The limits, as before, can be easily proven to behave in the correct way.

Then, for the influence of i^β on i^α ($\beta \neq \alpha$), we have to note that agents diffusing away from i^β decrease the denominator of the fraction $x_i^\alpha = n_i^\alpha / \sum_\gamma n_i^\gamma$, since they are decreasing n_i^β, and hence the fraction x_i^α increases. Therefore, the transition rates are $T_{i|i}^{-\alpha|\beta} = 0$ and

$$
T_{i|i}^{+\alpha|\beta} = T\left[\frac{n_i^\alpha}{N_i} \to \frac{n_i^\alpha}{N_i - 1}\,\middle|\, x_i^\beta\right] = D^\beta k_i^\beta x_i^\alpha x_i^\beta.
$$
$$(C6)$$

Finally, individuals diffusing away from j^β ($j \neq i$ and $\beta \neq \alpha$) are increasing the number of individuals in i^β, and hence increasing the denominator in the fraction of x_i^α (opposite as in the previous case). The influence is then so as to decrease the fraction of individuals at i^α and hence $T_{i|j}^{+\alpha|\beta} = 0$ and

$$
T_{i|j}^{-\alpha|\beta} = T\left[\frac{n_i^\alpha}{N_i} \to \frac{n_i^\alpha}{N_i + 1}\,\middle|\, x_j^\beta\right] = D^\beta a_{ij}^\beta \rho_{ij} x_i^\alpha x_j^\beta.
$$
$$(C7)$$

From the transition probabilities above we can derive the diffusion term corresponding to the stochastic effects in the Langevin equation [extrapolating s to two dimensions in Eq. (8)], which is

$$
s = \sqrt{\frac{\sum_\beta \sum_j D^\beta x_j^\beta \rho_{ij}\left[x_i^\alpha + \left(1 - 2x_i^\alpha\right)\delta^{\alpha\beta}\right]\left(a_{ij}^\beta + k_i^\beta \delta_{ij}\right)}{N_i}}.
$$
$$(C8)$$

This term can be added to Eq. (12)—multiplied by white Gaussian noise—in order to account for the stochastic effects introduced by the diffusive process in situations in which the thermodynamic limit cannot be assumed, or population sizes vary fast.

[1] E. Bonabeau, Proc. Nat. Acad. Sci. USA **99**, 7280 (2002).

[2] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis, Science **310**, 987 (2005).

[3] D. Helbing, Theory and Decision **40**, 149 (1996).

[4] D. Helbing, in *Game Theory, Experience, Rationality*, edited by W. Leinfellner and E. Köhler (Kluwer Academic, Dordrecht, 1998), pp. 211–214.

[5] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. Lett. **95**, 238701 (2005).

[6] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. E **74**, 011901 (2006).

[7] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[8] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, 2010).

[9] R. J. Requejo and J. Camacho, J. Theor. Biol. **272**, 35 (2011).

[10] R. J. Requejo and J. Camacho, Phys. Rev. Lett. **108**, 038701 (2012).

[11] D. Vilone, A. Robledo, and A. Sánchez, Phys. Rev. Lett. **107**, 038101 (2011).

[12] R. J. Requejo, J. Camacho, J. A. Cuesta, and A. Arenas, Phys. Rev. E **86**, 026105 (2012).

[13] K. M. Page and M. A. Nowak, J. Theor. Biol. **219**, 93 (2002).

[14] P. D. Taylor and L. B. Jonker, Math. Biosc. **40**, 145 (1978).

[15] J. Hofbauer, P. Schuster, and K. Sigmund, J. Theor. Biol. **81**, 609 (1979).

[16] J. Maynard Smith and G. R. Price, Nature (London) **246**, 15 (1973).

[17] J. Maynard Smith, J. Theo. Biol. **47**, 209 (1974).

[18] J. Hofbauer, Nonlinear Anal. Theor. Meth. Appl. **5**, 1003 (1981).

[19] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

[20] G. R. Price, Nature (London) **227**, 520 (1970).

[21] G. R. Price, Ann. Hum. Genet. **35**, 485 (1972).

[22] G. R. Price, Ann. Hum. Genet. **36**, 129 (1972).

[23] A. Traulsen, Evolution **64**, 316 (2010).

[24] C. P. Roca, J. A. Cuesta, and A. Sánchez, Phys. Rev. Lett. **97**, 158701 (2006).

[25] C. P. Roca, J. A. Cuesta, and A. Sánchez, Phys. Life Rev. **6**, 208 (2009).

[26] J. Lovelock and L. Margulis, TELLUS **26**, 2 (1974).

[27] R. J. Requejo and J. Camacho, Phys. Rev. E **85**, 066112 (2012).

[28] M. A. Nowak and R. M. May, Nature (London) **359**, 826 (1992).

[29] C. P. Roca, J. A. Cuesta, and A. Sánchez, Phys. Rev. E **80**, 046106 (2009).

[30] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).

[31] F. C. Santos, M. D. Santos, and J. M. Pacheco, Nature (London) **454**, 213 (2008).

[32] H. Ohtsuki and M. Nowak, J. Theor. Biol. **243**, 86 (2006).

[33] J. Y. Wakano, M. A. Nowak, and C. Hauert, Proc. Nat. Acad. Sci. USA **106**, 7910 (2009).

[34] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. Porter, J. Complex Networks **2**, 203 (2014).

[35] A. Cardillo, J. Gómez-Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo, and S. Boccaletti, Sci. Rep. **3**, 1344 (2013).

[36] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, Phys. Rev. Lett. **110**, 028701 (2013).

[37] A. Solé-Ribalta, M. De Domenico, N. E. Kouvaris, A. Díaz-Guilera, S. Gómez, and A. Arenas, Phys. Rev. E **88**, 032807 (2013).

[38] M. DeDomenico, A. Sole-Ribalta, E. Cozzo, M. Kivela, Y. Moreno, M. A. Porter, S. Gomez, and A. Arenas, Phys. Rev. X **3**, 041022 (2013).

[39] C. Granell, S. Gómez, and A. Arenas, Phys. Rev. Lett. **111**, 128701 (2013).

[40] N. E. Kouvaris, S. Hata, and A. Díaz-Guilera, Sci. Rep. **5**, 10840 (2015).

[41] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).

[42] A. J. Bladon, T. Galla, and A. J. McKane, Phys. Rev. E **81**, 066122 (2010).

[43] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. E **85**, 041901 (2012).

[44] R. A. Fisher, Ann. Eugenics **7**, 355 (1937).

[45] A. N. Kolmogorov, I. G. Petrovsky, and N. S. Piskunov, Bull. Univ. Moscow, Ser. Int. A **1**, 1 (1937) [reprinted in V. M. Tikhomirov (ed.), *Selected Works of A. M. Kolmogorov* (Kluwer, Amsterdam, 1991)].

[46] J. Fort and V. Mèndez, Rep. Prog. Phys. **65**, 895 (2002).

[47] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg, Nature (London) **428**, 646 (2004).

[48] A. Traulsen and M. A. Nowak, PLoS ONE **2**, e270 (2007).

[49] M. Kimura, Nature (London) **217**, 624 (1968).

[50] A. Traulsen, C. Hauert, H. De Silva, M. Nowak, and K. Sigmund, PNAS **106**, 709 (2009).

[51] A. Arenas, J. Camacho, J. A. Cuesta, and R. J. Requejo, J. Theor. Biol. **279**, 113 (2011).