FIXATION DYNAMICS ON MULTILAYER NETWORKS

RUODAN LIU* AND NAOKI MASUDA†

Abstract. Network structure has a large impact on constant-selection evolutionary dynamics, with which multiple types of different fitnesses (i.e., strengths) compete on the network. Here we study constant-selection dynamics on two-layer networks in which the fitness of a node in one layer affects that in the other layer, under birth-death processes and uniform initialization, which are commonly assumed. We show mathematically and numerically that two-layer networks are suppressors of selection, which suppresses the effects of the different fitness values between the different types on final outcomes of the evolutionary dynamics (called fixation probability), relative to the constituent one-layer networks. In fact, many two-layer networks are suppressors of selection relative to the most basic baseline, the Moran process. This result is in stark contrast with the results for conventional one-layer networks for which most networks are amplifiers of selection.

Key words. Evolutionary dynamics, fixation probability, constant selection, multilayer networks, amplifier, suppressor

MSC codes. 60J20, 91D30, 92D15, 92D25

1. Introduction.

Evolutionary dynamics is a mathematical modeling framework that allows us to investigate how the composition of different traits in a population changes over time under the assumption that fitter individuals tend to reproduce more often. For example, evolutionary game theory focuses on situations in which the fitness is determined by the game interaction between individuals, such as the prisoner's dilemma game [1–3]. Another example, which we focus on in the present study, is evolutionary graph theory in which one investigates the effects of network structure and possibly its variation over time on evolution of traits [1, 4–7]. In particular, studies of evolutionary games on networks have revealed that the conditions under which cooperation occurs in social dilemma games heavily depend on the network structure and that these conditions can be mathematically derived using random walk theory [8, 9].

Let us consider the constant-selection evolutionary dynamics on networks. In this dynamics, different types are assigned with different constant fitness values, each node of the given network is occupied by either of these types, and the different types compete for survival. One can view this dynamics as competition between resident and mutant phenotypes in structured populations, or social dynamics of opinions in which people switch between different opinions, influenced by their neighbors in the network.

A core property of constant-selection evolutionary dynamics on networks is the fixation probability. It is the likelihood that the mutant type initially occupying a single node of the network ultimately fixates, i.e., the mutant type eventually occupies all the nodes of the network, under the assumption that there is no mutation (i.e., the type on any node does not spontaneously change during the dynamics except due to the influence by their neighbors). The fixation probability depends on the network structure, the fitness of the mutant type, denoted by r, relative to the fitness of the resident type, which is normalized to be 1, as well as the initial condition [1, 4, 6]. The mutant type is more likely to fixate if r is large. The extent to which the fixation

*Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260-2900, USA (rliu8@buffalo.edu).

†Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260-2900, USA (naokimas@gmail.com).

probability of the mutant type increases with rising r hinges on the network structure. Some networks are known to be amplifiers of selection. By definition, in a network amplifying selection, a single mutant has a larger fixation probability than the case of the well-mixed population with the same number of nodes, which is equivalent to the so-called Moran process, at any $r > 1$, and has a lower fixation probability than the case of the Moran process at any $r < 1$. In amplifying networks, the effect of the difference between the mutant and resident type in terms of the fitness (i.e., r versus 1) is magnified by the network. In contrast, other networks are suppressors of selection such that a single mutant has a lower fixation probability than the case of the Moran process at any $r > 1$ and vice versa at any $r < 1$. Under a standard assumption of the birth-death process with selection on the birth and uniform initialization, it has been shown that most networks are amplifiers of selection [10–12]. Suppressors of selection are rare [11, 13].

Studies have shown that the amplifiers of selection under the birth-death process are not necessarily common when we introduce additional factors into evolutionary graph dynamics models, such as the non-uniform initialization [14, 15], directed networks [16], metapopulation models [17,18], temporal (i.e., time-varying) networks [19], and hypergraphs [20]. These results encourage us to study evolutionary dynamics on other extensions of conventional networks with an expectation that the dynamics on them may be drastically different from those on conventional networks.

In the present study, we explore constant-selection evolutionary dynamics on multilayer networks. Multilayer networks express the situation in which the individuals in a population are pairwise connected by different types of edges, such as different types of social relationships; the same pair of individuals may be directly connected by multiple types of edges [21–24]. In evolutionary dynamics on multilayer networks, each layer, corresponding to one type of edge, is a network, and evolutionary dynamics in different network layers are coupled in some manner. This setting has been investigated for evolutionary social dilemma games. See [25] for a review. Earlier work considered two-layer networks in which the game interaction occurs in one network layer and imitation of strategies between players occurs in the other network layer. Cooperation is more enhanced in this model if the edges overlap more heavily between the two layers [26–29] or under other conditions [30, 31] (but see [32]). When players are assumed to be engaged in game interactions, not just imitation of strategies, in the different layers, multilayer networks promote cooperation under some conditions such as positive degree correlation between two layers [33] and asynchronous strategy updating [34]. Cooperation can thrive in this class model even if each network layer in isolation does not support cooperation [35]. However, to the best of our knowledge, constant-selection evolutionary dynamics on multilayer networks have not been studied.

We particularly use two-layer networks. We introduce two models of constant-selection dynamics in multilayer networks, which are analogues of an evolutionary game model in multilayer networks [35], and semi-analytically calculate the fixation probability of mutants for each network layer for two-layer networks with high symmetry. Using martingale techniques, we also theoretically prove that the complete graph layer and the cycle graph layer in a two-layer network are suppressors of selection, and that the star graph layer and the complete bipartite layer in a two-layer network are more suppressing than the corresponding one-layer network. We numerically show that all the two-layer networks that we have numerically investigated are suppressors of selection, except for the coupled star networks. However, the coupled star networks are more suppressing than the one-layer star graphs. In this manner, we conclude that two-layer networks suppress the effects of selection.

2. Moran process. The Moran process is a model of stochastic constant-selection evolutionary dynamics in a well-mixed finite population with N individuals. The population consists of two types of individuals, i.e., the resident and mutant, with constant fitness values, 1 and r, respectively. At each time step, an individual is selected as the parent for reproduction with probability proportional to its fitness and an individual dies uniformly at random. Then, the parent's offspring replaces the dead individual. The fixation probability for a single mutant is given by [1, 4]

$$\rho = \frac{1 - 1/r}{1 - 1/r^N}. \tag{2.1}$$

Extensions of the Moran process to networks depend on specific update rules to be assumed. The network may be directed or weighted. A major variant of the updating rule that we consider in the present paper is the birth-death process with selection on the birth, or the Bd rule [6, 16, 36], which operates as follows. At each time step, an individual is selected as the parent, denoted by u, for reproduction with probability proportional to its fitness. This step is the same as in the Moran process. Then, u's type replaces the type of a neighbor of u, which is selected with probability proportional to the edge weight between u and itself. We use the Bd rule because a majority of work on constant-selection evolutionary dynamics on networks do so [4, 15, 37–43]. However, death-birth processes also give important insights into constant-selection evolutionary dynamics [10, 36, 44, 45], and we briefly examine it with our two-layer network model in section 5.7.

In a directed and weighted network, the edge direction indicates a one-way relationship between the two nodes. A network is an isothermal graph if the weighted in-degree (i.e., sum of the edge weight over all incoming edges to a node) is the same for all nodes. Unweighted regular graphs are examples of isothermal graph. The fixation probability for an isothermal graph is given by Eq. (2.1) [1, 4].

The fixation probability for a single mutant of the Moran process is $1/N$ at $r = 1$ [46–48]. Relative to the Moran process, many networks are either amplifiers or suppressors of selection [4, 10–13, 15, 39–42]. Amplifiers of selection are networks in which the fixation probability is larger than that for the Moran process (i.e., Eq. (2.1)) for any $r > 1$ and smaller than that for the Moran process for any $r < 1$. Suppressors of selection are networks in which the fixation probability is smaller and larger than for the Moran process for any $r > 1$ and $r < 1$, respectively.

3. Models. We introduce two models of constant-selection evolutionary dynamics for a population of N individuals in undirected and possibly weighted multilayer networks. The assumption of the undirected network is for simplicity, and it is straightforward to generalize the following models to the case of directed multilayer networks. We assume a two-layer network as the population structure, whereas it is straightforward to generalize the models to the case of more than two layers. Each layer is assumed to be a connected network with N nodes. It represents one of the two types of relationship between individuals, such as physical proximity contact or online social relationship in the case of human social networks. We call each node in one layer the replica node; there are $2N$ replica nodes in the entire two-layer network. Each replica node has a corresponding replica node in the other layer. A pair of the corresponding replica nodes, one in each layer, represents an individual (see Figure 1 for a schematic). Each edge within a layer represents direct connectivity between two replica nodes in the same layer. Each pair of individuals may be adjacent to each other in both layers, just one layer, or neither layer. For example, two people may directly interact both in person and online, or in only one of the two ways.

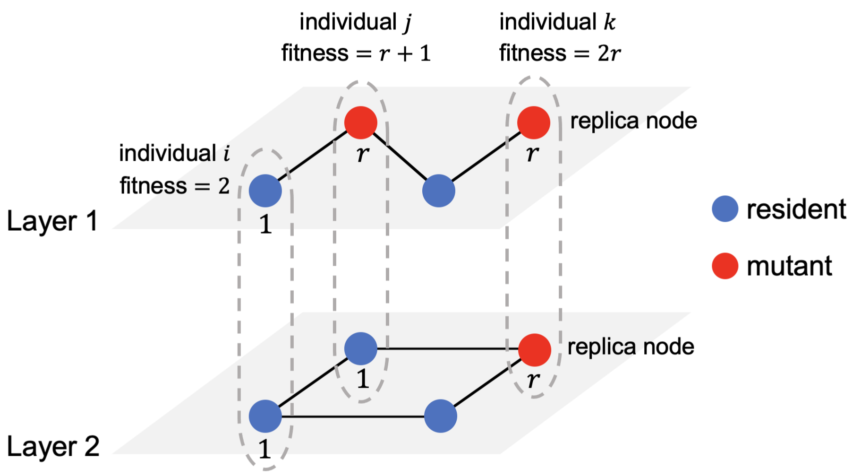

FIG. 1. *An example of a two-layer network. Each individual occupies a replica node in layer 1 and the corresponding replica node in layer 2, as indicated by dashed lines. A resident replica node and a mutant replica node are shown in blue and red, respectively.*

Both models extend the Bd process on conventional (i.e., mono-layer) networks and the Moran process in well-mixed populations to the case of two-layer networks. We assume that each of the $2N$ replica nodes takes either the resident or mutant type at any discrete time. The resident and mutant have fitness 1 and r, respectively, which are constant over time. We define the fitness of each individual by the sum of the fitness of the corresponding replica nodes in both layers [35]. In other words, the individual has fitness 2 if it is of the resident type in both layers, $r + 1$ if it is of the mutant type in one layer and the resident type in the other layer, and $2r$ if it is of the mutant type in both layers. We allow each individual to adopt different types in the opposite layers (i.e., the resident type in one layer and the mutant type in the other layer) because they may behave differently in different types of social relationship. Furthermore, success or failure of an individual in one type of social relationship may affect the same in the other domain, which motivates us to couple the fitness of each individual across the two layers [35].

The model assumptions up to this point are shared by models 1 and 2. Next, in model 1, in each time step, we select one individual (i.e., parent) for reproduction with probability proportional to its fitness. Then, we select one of the two layers to operate the Bd process with the equal probability, i.e., 1/2. Then, the parent selects one of its neighbors in the selected layer with probability proportional to the weight of the edge between the two individuals. Finally, the parent converts the type of the selected neighbor into the parent's type in the selected layer. This concludes one time step of the Bd process. We repeat this procedure until the entire population settles into an absorbing state in which all individuals are either of the resident or mutant type in each layer. It should be noted that the final state in the two layers may be different, i.e., resident in one layer and mutant in the other layer. This phenomenon may represent the situation in which two opinions or behaviors, O_1 and O_2, are competing in one layer, and two others, O_3 and O_4, are competing in the other layer. Then, all individuals may adopt the combination of O_1 and O_3 in the end, or the combination of O_1 and O_4, for example.

In each time step in model 2, we first select an individual i as the parent with probability proportional to its fitness in each time step. This process is the same as that in model 1. However, differently from model 1, we then do not select the layer but draw a neighbor of i in layer 1, denoted by j, with probability proportional to the edge weight $w_{ij}^{[1]}$, and j copies i's type. At the same time, we select an individual k as another parent with probability proportional to its fitness. Then, we select a neighbor of k in layer 2, denoted by ℓ, with probability proportional to the edge weight $w_{k\ell}^{[2]}$, and ℓ copies k's type. Individual k may be the same as individual i. This model is the same as the main model proposed in [35] except that their model used a death-birth instead of birth-death process and that the fitness for each individual is determined by two-player games in their model and therefore not constant for each type in general. We consider model 2 in addition to model 1 because model 2 is a direct extension of the model proposed in [35]. On the other hand, model 1 is more amenable to mathematical analysis of fixation dynamics than model 2.

4. Theoretical results.

4.1. Neutral drift. In this section, we focus on the case of neutral mutants, i.e., $r = 1$. The fixation probability for the neutral mutant type when there is initially just one mutant node selected uniformly at random must be equal to $1/N$ for one to be able to discuss amplifiers and suppressors of selection. We start by proving this property for two-layer networks.

THEOREM 4.1. *Consider model 1 under $r = 1$. When there are initially i mutants selected uniformly at random from the N replica nodes in one layer, the fixation probability for the mutant for that layer is equal to i/N.*

Proof. When $r = 1$, the fitness of each individual is always equal to 2. Then, the Bd process in layer 1 is independent of that in layer 2. Therefore, the proof is exactly the same as that for conventional networks as shown in [20, 46–48]. ∎

Remark 4.2. This theorem also holds true for model 2 with the proof being unchanged.

4.2. Complete graph layer in a two-layer network is always a suppressor of selection. In this section, we show that the complete graph layer in an arbitrary two-layer network is always a suppressor of selection under model 1. To this end, we let $\xi_t \in \{0,1\}^{2N}$, with $t \in \{0,1,\ldots\}$, be the state of the Bd process on the two-layer network at time t. The initial condition is given by ξ_0. We conveniently define t as the number of the state changes in layer 1, which we assumed to be the complete graph. In other words, when counting t, we ignore the updating steps in which a replica node in layer 1 is selected as the parent but does not induce the actual change of the state of the network (because the child node has the same type as that of the parent) or a replica node in layer 2 is selected as the parent (because there is then no change in the state in layer 1). We consider model 1 in the following text unless we state otherwise.

LEMMA 4.3. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted complete graph and layer 2 is an arbitrary connected network. We let X_t be the number of mutants in the first layer at time t and set $Y_t \equiv r^{-X_t}$. Then, sequence $\{Y_n\}$ is a submartingale for any $r > 0$.*

Proof. Let $\{\mathcal{B}_t\}$ be the filtration, i.e., an increasing sequence of the σ-algebras, generated by the Bd process on the two-layer network. We obtain $X_{t+1} = X_t + 1$ or $X_{t+1} = X_t - 1$ because we count the time t if and only if the number of the mutants changes in the complete graph layer. For an arbitrary state of the two-layer network with X_t mutants, ξ_t, we denote by $p(\xi_t)$ and $q(\xi_t)$ the probabilities with which $X_{t+1} = X_t + 1$ and $X_{t+1} = X_t - 1$, respectively. Note that $p(\xi_t) + q(\xi_t) = 1$.

To calculate $p(\xi_t)$ and $q(\xi_t)$, we denote by N_1 the number of individuals that have the mutant type in both layers, by N_2 the number of individuals that have the mutant type in layer 1 and the resident type in layer 2, by N_3 the number of individuals that have the resident type in layer 1 and the mutant type in layer 2, and by N_4 the number of individuals that have the resident type in both layers. Note that $N_1 + N_2 + N_3 + N_4 = N$. In a single time step of the original Bd process, X_t increases by one with probability

$$p' = \frac{2rN_1 + (r+1)N_2}{2rN_1 + (r+1)(N_2+N_3) + 2N_4} \cdot \frac{1}{2} \cdot \frac{N_3 + N_4}{N_1 + N_2 + N_3 + N_4 - 1} \tag{4.1}$$

and decreases by one with probability

$$q' = \frac{(r+1)N_3 + 2N_4}{2rN_1 + (r+1)(N_2+N_3) + 2N_4} \cdot \frac{1}{2} \cdot \frac{N_1 + N_2}{N_1 + N_2 + N_3 + N_4 - 1}. \tag{4.2}$$

By combining Eqs. (4.1) and (4.2) with $p(\xi_t)/q(\xi_t) = p'/q'$ and $p(\xi_t) + q(\xi_t) = 1$, we obtain

$$p(\xi_t) = \frac{r}{r+1} - \varepsilon, \tag{4.3}$$

$$q(\xi_t) = \frac{1}{r+1} + \varepsilon, \tag{4.4}$$

where

$$\varepsilon = \frac{(r-1)\left[rN_1N_3 + (r+1)N_2N_3 + N_2N_4\right]}{(r+1)\left\{\left[2rN_1 + (r+1)N_2\right](N_3+N_4) + \left[(r+1)N_3 + 2N_4\right](N_1+N_2)\right\}}. \tag{4.5}$$

We obtain

$$\begin{aligned} E[Y_{t+1}|\mathcal{B}_t] &= p(\xi_t) r^{-(X_t+1)} + q(\xi_t) r^{-(X_t-1)} \\ &= \left[\left(\frac{r}{r+1} - \varepsilon\right)\frac{1}{r} + \left(\frac{1}{r+1} + \varepsilon\right) r\right] Y_t \\ &= \left[1 + \left(r - \frac{1}{r}\right)\varepsilon\right] Y_t, \end{aligned} \tag{4.6}$$

where $E[\cdot|\cdot]$ represents the conditional expectation. If $r > 1$, we obtain $E[Y_{t+1}|\mathcal{B}_t] \geq Y_t$ because $r - r^{-1} > 0$ and $\varepsilon \geq 0$. If $r < 1$, we also obtain $E[Y_{t+1}|\mathcal{B}_t] \geq Y_t$ because $r - r^{-1} < 0$ and $\varepsilon \leq 0$. Therefore, in both cases, Y_t is a submartingale. If $r = 1$, we obtain $\varepsilon = 0$ such that Y_t is a martingale, which is a submartingale. ∎

Remark 4.4. Our choice of Y_t is inspired by the construction of a martingale for the biased random walk on $\mathbb{Z}$ (see, e.g., [49, 50]) and its application to constant-selection evolutionary dynamics [14, 43, 51].

THEOREM 4.5. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted complete graph and layer 2 is an arbitrary connected network. Then, the complete graph layer is a suppressor of selection.*

Proof. Sequence $\{Y_t\}$ is a submartingale and bounded because $r^{-N} \le Y_t \le 1$ $\forall t$ when $r \ge 1$ and $1 \le Y_t \le r^{-N}$ $\forall t$ when $r \le 1$. Therefore, Y_t converges almost surely and $E[Y_\infty]$ is finite owing to the martingale convergence theorem [49, 50]. The present Bd process has four absorbing states in which all the nodes in each layer are unanimously occupied by the resident or mutant. The two absorbing states in which all the nodes in layer 1 are occupied by the resident yields $X_t = 0$. The other two absorbing states in which all the nodes in layer 1 are occupied by the mutant yields $X_t = N$. Because an absorbing state is ultimately reached with probability 1,

$$E[Y_\infty] \ge Y_0 \tag{4.7}$$

yields

$$x(\xi_0)r^{-N} + [1 - x(\xi_0)]\, r^{-0} \ge r^{-X_0}, \tag{4.8}$$

where $x(\xi_0)$ is the fixation probability of the mutant under an initial condition ξ_0 with X_0 mutants in layer 1; therefore, there are initially $N - X_0$ residents in the same layer. Equation (4.8) yields

$$\begin{cases} x(\xi_0) \le \frac{1-r^{X_0}}{1-r^N} & (r \ge 1), \\ x(\xi_0) \ge \frac{1-r^{X_0}}{1-r^N} & (r < 1). \end{cases} \tag{4.9}$$

Our goal is to exclude the equalities in Eq. (4.9) for $1 \le X_0 \le N - 1$ because then it will hold true that the complete graph layer is a suppressor of selection. To show this, we distinguish among the following three cases.

To state the first case, we note that $\varepsilon = 0$ for $r \ne 1$ and $1 \le X_0 \le N - 1$ if and only if $N_2 = N_3 = 0$. Therefore, if the initial condition ξ_0 satisfies $N_2 > 0$ or $N_3 > 0$, then Eq. (4.6) implies that

$$E[Y_1|\xi_0] > Y_0 \tag{4.10}$$

for $r \ne 1$. By combining $E[Y_2|\mathcal{B}_t] \ge Y_1$, which follows from Lemma 4.3, with Eq. (4.10), we obtain $E[Y_2|\xi_0] > Y_0$.

The second and third cases concern the initial condition ξ_0 satisfying $N_2 = N_3 = 0$ such that each individual has the same type (i.e., resident or mutant) in both layers. Then, we obtain $E[Y_1|\xi_0] = Y_0$ because $\varepsilon = 0$. As the second case, we consider the situation in which ξ_0 satisfies $N_1 \le N - 2$ in addition to $N_2 = N_3 = 0$. In this case, the network's state after the first state transition, ξ_1, satisfies $(N_1, N_2, N_3, N_4) = (N_1, 1, 0, N - N_1 - 1)$ with probability $p(\xi_0) = r/(r+1)$. Conditioned on this transition, we obtain $E[Y_2|\xi_1] > Y_1$ for $r \ne 1$, which is an adaptation of Eq. (4.10). We obtain $E[Y_2|\xi_1] > Y_1$ because this particular ξ_1 yields $N_2 = 1$, which implies $(r - r^{-1})\varepsilon > 0$. If we start from the same ξ_0 and a different ξ_1 is realized with probability $1 - p(\xi_0)$, we still obtain $E[Y_2|\xi_1] \ge Y_1$ owing to Lemma 4.3. Therefore, we obtain $E[Y_2|\xi_0] > Y_0$ when ξ_0 satisfies $N_1 \le N - 2$ and $N_2 = N_3 = 0$.

As the third case, we consider the situation in which ξ_0 satisfies $N_1 = N - 1$, which implies that $N_2 = N_3 = 0$. In this case, ξ_1 satisfies $(N_1, N_2, N_3, N_4) = (N_1 - 2, 0, 1, N - N_1)$ with probability $q(\xi_0) = 1/(r + 1)$. Conditioned on this transition, we obtain $E[Y_2|\xi_1] > Y_1$ for $r \ne 1$ because this particular ξ_1 yields $N_3 = 1$, which implies $(r - r^{-1})\varepsilon > 0$. If we start from the same ξ_0 and a different ξ_1 is realized with probability $1 - q(\xi_0)$, we still obtain $E[Y_2|\xi_1] \ge Y_1$ owing to Lemma 4.3. Therefore, we obtain $E[Y_2|\xi_0] > Y_0$ when ξ_0 satisfies $N_1 = N - 1$.

Because $E[Y_2|\xi_0] > Y_0$ holds true in all the three cases, we obtain $E[Y_2|\mathcal{B}_0] > Y_0$, which together with $E[Y_{t+1}|\mathcal{B}_t] \geq Y_t$, $\forall t \in \{2, 3, \ldots\}$ leads to Eq. (4.7) with the strict inequality. Therefore, Eqs. (4.8) and (4.9) hold true with the strict inequality when $r \neq 1$. □

4.3. Cycle graph layer in a two-layer network is always a suppressor of selection. We use the same method as that for the complete graph layer to show that Lemma 4.3 also holds true when one replaces the complete graph layer by the cycle graph. The cycle graph, which we assumed to form layer 1, is defined by $w_{ij}^{[1]} = 1$ if $j = i \pm 1 \mod N$, and $w_{ij}^{[1]} = 0$ otherwise. For simplicity, we assume that the replica nodes of the mutant type are initially consecutive (i.e., forming just one connected component of mutants) in the cycle graph layer.

LEMMA 4.6. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted cycle graph and layer 2 is an arbitrary connected network. We let X_t the number of mutants in layer 1 at time t and set $Y_t \equiv r^{-X_t}$. The individuals of the mutant type are assumed to be initially located at consecutive replica nodes on the cycle. Then, sequence $\{Y_n\}$ is a submartingale for any $r > 0$.*

We prove Lemma 4.6 in section S1.

THEOREM 4.7. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted cycle graph and layer 2 is an arbitrary connected network. Then, the cycle graph layer is a suppressor of selection, given that the individuals of the mutant type are initially located at consecutive replica nodes on the cycle.*

We prove Theorem 4.7 in section S2.

4.4. Complete bipartite graph layer in a two-layer network. In this section, we consider the two-layer network in which layer 1 is the unweighted complete bipartite graph and layer 2 is an arbitrary connected network. The complete bipartite graph, denoted by K_{N_1,N_2}, where $N_1 + N_2 = N$, consists of two disjoint subsets of nodes V_1 and V_2 with N_1 and N_2 nodes, respectively. It is defined by $w_{ij}^{[1]} = 1$ if $i \in V_1$ and $j \in V_2$, or $i \in V_2$ and $j \in V_1$, and $w_{ij}^{[1]} = 0$ otherwise. We construct a similar proof to that for the complete graph or cycle graph layer to show that the complete bipartite graph layer in an arbitrary two-layer network is more suppressing than the one-layer complete bipartite graph.

LEMMA 4.8. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted complete bipartite graph and layer 2 is an arbitrary connected network. We let $\boldsymbol{X_t} = [X_{1,t}, X_{2,t}]$, where $X_{1,t}$ and $X_{2,t}$ are the number of nodes in V_1 and V_2, respectively, that are occupied by the mutant in layer 1 at time t. We define $Y_t \equiv h_1^{X_{1,t}} h_2^{X_{2,t}}$, where*

$$h_1 = \frac{N_1 + N_2 r}{N_1 r^2 + N_2 r}, \tag{4.11}$$

$$h_2 = \frac{N_2 + N_1 r}{N_2 r^2 + N_1 r}. \tag{4.12}$$

Then, sequence $\{Y_n\}$ is a submartingale for any $r > 0$.

We prove Lemma 4.8 in section S3.

Remark 4.9. Our choice of Y_t is inspired by the application of martingales to the Bd process in one-layer complete bipartite graphs [43].

THEOREM 4.10. *Consider the Bd process on the two-layer network in which layer 1 is the unweighted complete bipartite graph and layer 2 is an arbitrary connected network. Then, the complete bipartite graph layer is more suppressing than the one-layer complete bipartite graph.*

Proof. Equation (4.7) holds true in the present case as well. It is equivalent to

$$x(\xi_0) h_1^{N_1} h_2^{N_2} + [1 - x(\xi_0)]\, h_1^0 h_2^0 \geq h_1^{X_{1,0}} h_2^{X_{2,0}}, \tag{4.13}$$

where $x(\xi_0)$ is the fixation probability of the mutant type under an initial condition ξ_0 with $X_{1,0}$ mutants on the nodes in V_1 and $X_{2,0}$ mutants on the nodes in V_2; there are initially $X_{1,0} + X_{2,0}$ mutants and $N - (X_{1,0} + X_{2,0})$ residents in the complete bipartite graph layer. Equation (4.13) yields

$$\begin{cases} x(\xi_0) \leq \frac{h_1^{X_{1,0}} h_2^{X_{2,0}} - 1}{h_1^{N_1} h_2^{N_2} - 1} & (r \geq 1), \\ \\ x(\xi_0) \geq \frac{h_1^{X_{1,0}} h_2^{X_{2,0}} - 1}{h_1^{N_1} h_2^{N_2} - 1} & (r < 1). \end{cases} \tag{4.14}$$

To exclude the equalities in Eq. (4.14) for $1 \leq X_{1,0} + X_{2,0} \leq N - 1$, we distinguish among 12 cases that are different in terms of the number of individuals in V_1 and that in V_2 with different fitness values. We obtain

$$E[Y_3|\xi_0] > Y_0 \tag{4.15}$$

for all the 12 cases; see section S4 for the proof.

Because Eq. (4.15) holds true in all the cases, we obtain $E[Y_3|\mathcal{B}_0] > Y_0$, which together with $E[Y_{t+1}|\mathcal{B}_t] \geq Y_t$, $\forall t \in \{3, 4, \ldots\}$ leads to Eq. (4.7) with the strict inequality. Therefore, Eqs. (4.13) and (4.14) hold true with the strict inequality when $r \neq 1$, proving that the complete bipartite graph layer in a two-layer network is more suppressing than the mono-layer complete bipartite graph. ∎

Remark 4.11. If $N_1 = 1$ and $N_2 = N - 1$, the complete bipartite graph layer reduces to a star graph. Therefore, Lemma 4.8 and Theorem 4.10 also hold true when one layer of the two-layer network is a star graph.

Remark 4.12. All lemmas and theorems also hold true for model 2 with the proof being essentially unchanged (see section S4 for more).

5. Semi-analytical results for two-layer networks with high symmetry.

5.1. Exact computation of the fixation probability in two-layer networks. In this section, we explain how to exactly calculate the fixation probability for the mutant type when there is initially one replica node of mutant type that is selected uniformly at random in layer 1, and one replica node of mutant type in layer 2. This initial state is the same as that assumed in [35]. Let $s_i^{[1]} \in \{0, 1\}$ and $s_i^{[2]} \in \{0, 1\}$ be individual i's type in layer 1 and layer 2, respectively, where values 0 and 1 represent resident and mutant, respectively. Then, the state of the evolutionary dynamics is specified by a $2N$-dimensional binary vector $\boldsymbol{s} = (s_1^{[1]}, \ldots, s_N^{[1]}, s_1^{[2]}, \ldots, s_N^{[2]})$. Therefore, there are 2^{2N} states in total. We number the states from 1 to 2^{2N} by a bijective map, denoted by φ, given by

$$\begin{aligned} \varphi : S &\to \{1, \ldots, 2^{2N}\}, \\ \boldsymbol{s} &\mapsto \varphi(\boldsymbol{s}), \end{aligned} \tag{5.1}$$

where S is the set of all states. Let $P = [p_{i,j}]$ denote the $2^{2N} \times 2^{2N}$ transition probability matrix, where $p_{i,j}$ is the probability that the state moves from the ith state to the jth state in a time step of the birth-death process. Denote the probability that the mutant fixates in layer 1 by $x_i^{[1]}$ starting from the ith state, where $i \in \{1, \ldots, 2^{2N}\}$. Similarly, denote the probability that the mutant fixates in layer 2 by $x_i^{[2]}$ starting from the ith state. We can obtain $x_i^{[1]}$ by solving the linear system

$$\boldsymbol{x}^{[1]} = P\boldsymbol{x}^{[1]}, \tag{5.2}$$

where $\boldsymbol{x}^{[1]} = (x_1^{[1]}, \ldots, x_{2^{2N}}^{[1]})^\top$, and $^\top$ represents the transposition, with boundary conditions $x^{[1]}_{\varphi((1,\ldots,1,1,\ldots,1))} = 1$, $x^{[1]}_{\varphi((1,\ldots,1,0,\ldots,0))} = 1$, $x^{[1]}_{\varphi((0,\ldots,0,1,\ldots,1))} = 0$, and $x^{[1]}_{\varphi((0,\ldots,0,0,\ldots,0))} = 0$. Similarly, we can obtain $x_i^{[2]}$ by solving the same linear system

$$\boldsymbol{x}^{[2]} = P\boldsymbol{x}^{[2]}, \tag{5.3}$$

where $\boldsymbol{x}^{[2]} = (x_1^{[2]}, \ldots, x_{2^{2N}}^{[2]})^\top$, with boundary conditions $x^{[2]}_{\varphi((1,\ldots,1,1,\ldots,1))} = 1$, $x^{[2]}_{\varphi((1,\ldots,1,0,\ldots,0))} = 0$, $x^{[2]}_{\varphi((0,\ldots,0,1,\ldots,1))} = 1$, and $x^{[2]}_{\varphi((0,\ldots,0,0,\ldots,0))} = 0$. Let $C \subset S$ be the set of initial states that contain only one replica node of mutant type in layer 1 and one replica node of mutant type in layer 2. The cardinality of C is N^2. Denote the numerical labels of states in C by $\{k_1, \ldots, k_{N^2}\}$. Then, the fixation probability for the mutant type in layer 1 and 2 starting with the initial configuration with just one mutant in each layer, denoted by $x_C^{[1]}$ and $x_C^{[2]}$, respectively, is given by

$$x_C^{[1]} = \sum_{i \in \{k_1, \ldots, k_{N^2}\}} \frac{x_i^{[1]}}{N^2}, \tag{5.4}$$

$$x_C^{[2]} = \sum_{i \in \{k_1, \ldots, k_{N^2}\}} \frac{x_i^{[2]}}{N^2}. \tag{5.5}$$

For an arbitrary two-layer network, we need to solve a linear system with $2^{2N} - 4$ unknowns to obtain the fixation probability of the mutant type. This is computationally prohibitive when N is large. Although we can exploit that $x^{[1]}_{\varphi((0,\ldots,0,s_1^{[2]},\ldots,s_N^{[2]}))} = 0$ and $x^{[1]}_{\varphi((1,\ldots,1,s_1^{[2]},\ldots,s_N^{[2]}))} = 1$ for any $(s_1^{[2]}, \ldots, s_N^{[2]}) \in \{0,1\}^N$ and similar relationships for $x^{[2]}$, the number of unknowns still scales with 2^{2N} as N increases. Therefore, to drastically reduce the dimension of the linear system to be solved, we analyze two-layer networks with a highly symmetric structure for each layer, in which all or most replica nodes are structurally equivalent to other replica nodes. This strategy has been used for exactly calculating fixation probabilities on conventional networks [1,4], hypergraphs [20], and temporal networks [19].

In the following text, we consider model 1 except in section 5.8, where we briefly consider model 2.

5.2. Coupled complete graphs. We first consider the case in which each layer is the complete graph with N nodes. Because all nodes in each layer are structurally equivalent to each other, we only need to track the number of individuals with the mutant type in both layers, denoted by i_1, the number of individuals with the mutant type in layer 1 and the resident type in layer 2, denoted by i_2, the number of individuals with the resident type in layer 1 and the mutant type in layer 2, denoted by i_3, and the number of individuals with the resident type in both layers, denoted by i_4. One can specify the state of the evolutionary dynamics by a 4-tuple $\boldsymbol{i} = (i_1, i_2, i_3, i_4)$, where $i_1, i_2, i_3, i_4 \in \{0, 1, \ldots, N\}$ and $i_1+i_2+i_3+i_4 = N$. Therefore, there are $\binom{N+3}{3}$ states in total, where $\binom{}{}$ represents the binomial coefficient. For visual clarity, we denote the transition probability matrix by $P = [p_{\boldsymbol{i}\to\boldsymbol{j}}]$, where $p_{\boldsymbol{i}\to\boldsymbol{j}}$ is the probability that the state moves from $\boldsymbol{i} = (i_1, i_2, i_3, i_4)$ to $\boldsymbol{j} = (j_1, j_2, j_3, j_4)$ in a time step. Assume that the current state is $\boldsymbol{i} = (i_1, i_2, i_3, i_4)$. There are nine types of events that can occur next.

In the first type of event, an individual that has the mutant type in layer 1 (and either type in layer 2) is selected as the parent, which occurs with probability $[2ri_1+(1+r)i_2]/[2ri_1+(1+r)(i_2+i_3)+2i_4]$, and layer 1 is selected for the reproduction event with probability 1/2. Then, we select a neighbor of the parent in layer 1 for death, and the selected individual, which we refer to as the child, has the resident type in layer 1 and the mutant type in layer 2 with probability $i_3/(N-1)$. Then, the child copies the parent's type in layer 1. The state after this event is (i_1+1, i_2, i_3-1, i_4). Therefore, we obtain

$$(5.6) \qquad p_{(i_1,i_2,i_3,i_4)\to(i_1+1,i_2,i_3-1,i_4)} = \frac{2ri_1+(1+r)i_2}{2ri_1+(1+r)(i_2+i_3)+2i_4} \cdot \frac{1}{2} \cdot \frac{i_3}{N-1}.$$

In the second type of event, an individual that has the mutant type in layer 1 is selected as the parent, which occurs with probability $[2ri_1 + (1+r)i_2]/[2ri_1 + (1+r)(i_2+i_3)+2i_4]$, and layer 1 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 1 as the child, and the child has the resident type in both layers, which occurs with probability $i_4/(N-1)$. Then, the child copies the parent's type in layer 1. The state after this event is (i_1, i_2+1, i_3, i_4-1). Therefore, we obtain

$$(5.7) \qquad p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2+1,i_3,i_4-1)} = \frac{2ri_1+(1+r)i_2}{2ri_1+(1+r)(i_2+i_3)+2i_4} \cdot \frac{1}{2} \cdot \frac{i_4}{N-1}.$$

In the third type of event, an individual that has the resident type in layer 1 is selected as the parent, which occurs with probability $[(1+r)i_3+2i_4]/[2ri_1+(1+r)(i_2+i_3)+2i_4]$, and layer 1 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 1 as the child, and the child has the mutant type in both layers, which occurs with probability $i_1/(N-1)$. The state after this event is (i_1-1, i_2, i_3+1, i_4). Therefore, we obtain

$$(5.8) \qquad p_{(i_1,i_2,i_3,i_4)\to(i_1-1,i_2,i_3+1,i_4)} = \frac{(1+r)i_3+2i_4}{2ri_1+(1+r)(i_2+i_3)+2i_4} \cdot \frac{1}{2} \cdot \frac{i_1}{N-1}.$$

In the fourth type of event, an individual that has the resident type in layer 1 is selected as the parent, which occurs with probability $[(1+r)i_3+2i_4]/[2ri_1 + (1+r)(i_2+i_3)+2i_4]$, and layer 1 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 1 as the child, and the child has the mutant type in layer 1 and the resident type in layer 2, which occurs with probability $i_2/(N-1)$. The state after this event is (i_1, i_2-1, i_3, i_4+1). Therefore, we obtain

$$(5.9) \qquad p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2-1,i_3,i_4+1)} = \frac{(1+r)i_3+2i_4}{2ri_1+(1+r)(i_2+i_3)+2i_4} \cdot \frac{1}{2} \cdot \frac{i_2}{N-1}.$$

In the fifth type of event, an individual that has the mutant type in layer 2 is selected as the parent, which occurs with probability $[2ri_1 + (1+r)i_3]/[2ri_1 + (1+r)(i_2+i_3) + 2i_4]$, and layer 2 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 2 as the child, and the child has the mutant type in layer 1 and the resident type in layer 2, which occurs with probability $i_2/(N-1)$. The state after this event is (i_1+1, i_2-1, i_3, i_4). Therefore, we obtain

$$p_{(i_1,i_2,i_3,i_4)\to(i_1+1,i_2-1,i_3,i_4)} = \frac{2ri_1 + (1+r)i_3}{2ri_1 + (1+r)(i_2+i_3) + 2i_4} \cdot \frac{1}{2} \cdot \frac{i_2}{N-1}. \tag{5.10}$$

In the sixth type of event, an individual that has the mutant type in layer 2 is selected as the parent, which occurs with probability $[2ri_1 + (1+r)i_3]/[2ri_1 + (1+r)(i_2+i_3) + 2i_4]$, and layer 2 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 2 as the child, and the child has the resident type in both layers, which occurs with probability $i_4/(N-1)$. The state after this event is (i_1, i_2, i_3+1, i_4-1). Therefore, we obtain

$$p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2,i_3+1,i_4-1)} = \frac{2ri_1 + (1+r)i_3}{2ri_1 + (1+r)(i_2+i_3) + 2i_4} \cdot \frac{1}{2} \cdot \frac{i_4}{N-1}. \tag{5.11}$$

In the seventh type of event, an individual that has the resident type in layer 2 is selected as the parent, which occurs with probability $[(1+r)i_2 + 2i_4]/[2ri_1 + (1+r)(i_2+i_3) + 2i_4]$, and layer 2 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 2 as the child, and the child has the mutant type in both layers, which occurs with probability $i_1/(N-1)$. The state after this event is (i_1-1, i_2+1, i_3, i_4). Therefore, we obtain

$$p_{(i_1,i_2,i_3,i_4)\to(i_1-1,i_2+1,i_3,i_4)} = \frac{(1+r)i_2 + 2i_4}{2ri_1 + (1+r)(i_2+i_3) + 2i_4} \cdot \frac{1}{2} \cdot \frac{i_1}{N-1}. \tag{5.12}$$

In the eighth type of event, an individual that has the resident type in layer 2 is selected as the parent, which occurs with probability $[(1+r)i_2 + 2i_4]/[2ri_1 + (1+r)(i_2+i_3) + 2i_4]$, and layer 2 is selected for reproduction with probability 1/2. Then, we select a neighbor of the parent in layer 2 as the child, and the child has the resident type in layer 1 and the mutant type in layer 2, which occurs with probability $i_3/(N-1)$. The state after this event is (i_1, i_2, i_3-1, i_4+1). Therefore, we obtain

$$p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2,i_3-1,i_4+1)} = \frac{(1+r)i_2 + 2i_4}{2ri_1 + (1+r)(i_2+i_3) + 2i_4} \cdot \frac{1}{2} \cdot \frac{i_3}{N-1}. \tag{5.13}$$

If any other event occurs, the state remains unchanged. Therefore, we obtain

$$\begin{aligned} p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2,i_3,i_4)} = 1 &- p_{(i_1,i_2,i_3,i_4)\to(i_1+1,i_2,i_3-1,i_4)} - p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2+1,i_3,i_4-1)} \\ &- p_{(i_1,i_2,i_3,i_4)\to(i_1-1,i_2,i_3+1,i_4)} - p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2-1,i_3,i_4+1)} \\ &- p_{(i_1,i_2,i_3,i_4)\to(i_1+1,i_2-1,i_3,i_4)} - p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2,i_3+1,i_4-1)} \\ &- p_{(i_1,i_2,i_3,i_4)\to(i_1-1,i_2+1,i_3,i_4)} - p_{(i_1,i_2,i_3,i_4)\to(i_1,i_2,i_3-1,i_4+1)}. \end{aligned} \tag{5.14}$$

By slightly adapting the notations introduced in section 5.1, we denote by $x_{\boldsymbol{i}}^{[1]}$ and $x_{\boldsymbol{i}}^{[2]}$ the fixation probability of the mutant type in layer 1 and layer 2, respectively, when the initial state is $\boldsymbol{i} = (i_1, i_2, i_3, i_4)$. To obtain fixation probabilities in layer 1, we solve Eq. (5.2), where $\boldsymbol{x}^{[1]}$ is a column vector of which each entry is the fixation probability for the mutant type starting from one of the $\binom{N+3}{3}$ initial states. The boundary conditions are given by $x^{[1]}_{(N,0,0,0)} = 1$, $x^{[1]}_{(0,N,0,0)} = 1$, $x^{[1]}_{(0,0,N,0)} = 0$, and $x^{[1]}_{(0,0,0,N)} = 0$. To obtain fixation probabilities in layer 2, we solve Eq. (5.3) with boundary conditions $x^{[2]}_{(N,0,0,0)} = 1$, $x^{[2]}_{(0,N,0,0)} = 0$, $x^{[2]}_{(0,0,N,0)} = 1$, and $x^{[2]}_{(0,0,0,N)} = 0$. There are two initial states with one mutant in each layer, i.e., $(1, 0, 0, N-1)$ and $(0, 1, 1, N-2)$. These initial states occur with probability $1/N$ and $(N-1)/N$, respectively. Therefore, we obtain

$$x^{[\ell]}_C = \frac{1}{N} x^{[\ell]}_{(1,0,0,N-1)} + \frac{N-1}{N} x^{[\ell]}_{(0,1,1,N-2)}, \quad \ell \in \{1, 2\}, \tag{5.15}$$

where we remind that $x^{[\ell]}_C$ is the fixation probability for the mutant type in layer ℓ when there is initially one mutant in each layer.

We obtained $x^{[1]}_C$ $(= x^{[2]}_C)$ by numerically solving Eq. (5.2) for $N = 6$ and $N = 30$. The results shown in Figure 2(a) and 2(b) for $N = 6$ and $N = 30$, respectively, indicate that these coupled complete graphs are suppressors of selection. This result is consistent with Theorem 4.5.

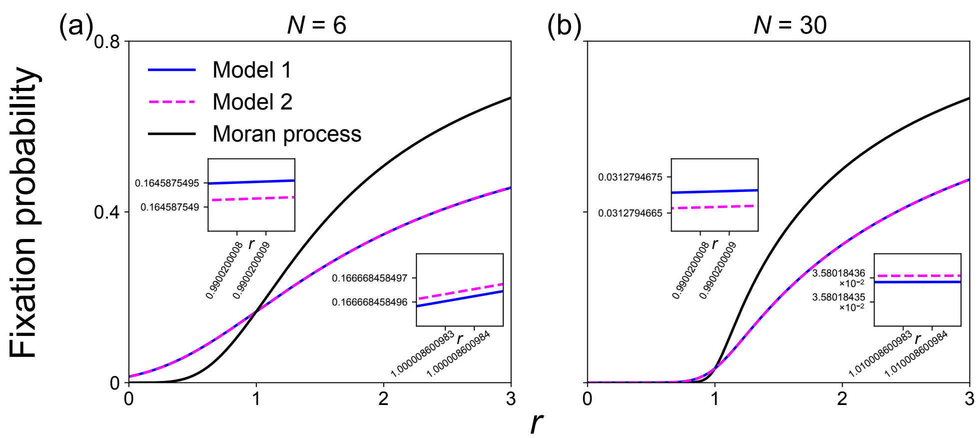

FIG. 2. *Fixation probability for coupled complete graphs under models 1 and 2. (a) $N = 6$. (b) $N = 30$. The insets to the left within each panel magnify the results for r values less than and close to $r = 1$. Those to the right within each panel magnify the results for r values greater than and close to $r = 1$.*

5.3. Combination of the complete graph and star graph. Next, we consider the two-layer network in which layer 1 is the complete graph and layer 2 is the star graph. All the N nodes in the complete graph are structurally equivalent, and so are all the $N-1$ leaf nodes (i.e., nodes with degree 1) in the star graph. Therefore, we represent the state of the evolutionary dynamics by $\boldsymbol{i} = (h_1, h_2, i_1, i_2, i_3, i_4)$, i.e., an ordered 6-tuple, where $h_1 = 0$ or 1 if the individual that is the hub node (i.e., the replica node with degree $N-1$) in layer 2 is of resident or mutant type in layer 1, respectively; $h_2 = 0$ or 1 if the hub node in layer 2 is of the resident or mutant type, respectively; i_1 is the number of the remaining $N-1$ individuals that have the mutant type in both layers, i_2 is the number of the remaining $N-1$ individuals that have the mutant type in layer 1 and the resident type in layer 2; i_3 is the number of the remaining $N-1$ individuals that have the resident type in layer 1 and the mutant type in layer 2; i_4 is the number of the remaining $N-1$ individuals that have the resident type in both layers. There are $2^2\binom{N+2}{3}$ states in total.

Similarly to the case in which both layers are the complete graph, we distinguish nine types of state transitions from each state. We derive the probability of each state transition in section S5.

We use the same numerical method for solving Eqs. (5.2) and (5.3) as that for the coupled complete graphs. We show the fixation probability for the mutant type for $N = 6$ and $N = 30$ in Figure 3(a) and 3(b), respectively. The figure suggests that the two-layer networks composed of a complete graph layer and a star graph layer are suppressors of selection.

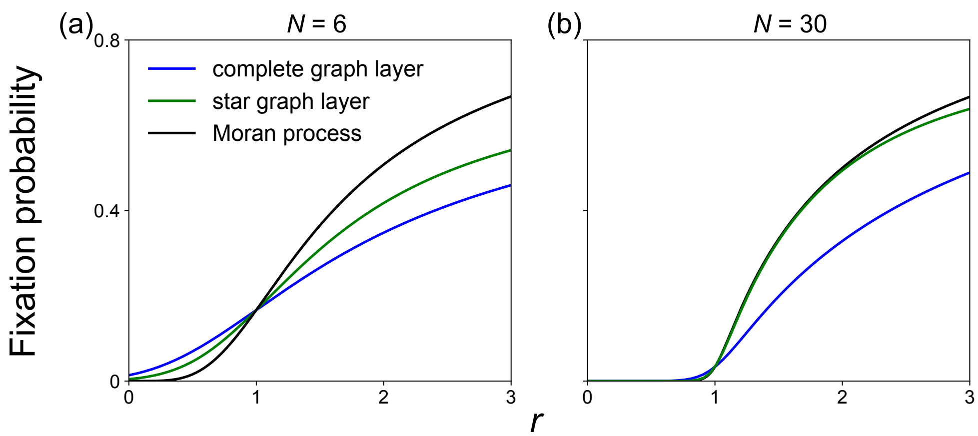

FIG. 3. *Fixation probability for two-layer networks composed of a complete graph layer and a star graph layer under model 1. (a) $N = 6$. (b) $N = 30$.*

5.4. Coupled star graphs. Here we consider the case in which each layer is the star graph with N nodes. We assume that the hub replica node in layer 1 corresponds to a leaf replica node in layer 2, and vice versa. Then, we can specify the network's state by an ordered 8-tuple $\boldsymbol{i} = (h_1, h_2, h_3, h_4, i_1, i_2, i_3, i_4)$, where $h_1 = 0$ or 1 if the hub node in layer 1 is of resident or mutant type, respectively; $h_2 = 0$ or 1 if the individual that is the hub node in layer 1 is of resident or mutant type in layer 2, respectively; $h_3 = 0$ or 1 if the individual that is the hub node in layer 2 is of resident or mutant type in layer 1, respectively; $h_4 = 0$ or 1 if the hub node in layer 2 is of resident or mutant type, respectively; We reuse i_1, i_2, i_3, and i_4 defined in section 5.3 with a slight difference. Here, we count i_1, i_2, i_3, and i_4 among the $N-2$ individuals that are leaf nodes in both layers. There are $2^4\binom{N+1}{3}$ states in total.

We distinguish all types of state transitions from each state and derive the probability of each state transition in section S6. The number of the type of state transitions varies between seven and nine and depends on the current state.

We use the same numerical method for solving Eqs. (5.2) and (5.3) as that for the coupled complete graphs. We show the fixation probability for the mutant type for $N = 6$ and $N = 30$ in Figure 4(a) and 4(b), respectively. Figure 4 indicates that the coupled star graph is a suppressor of selection when $N = 6$, but is neither a suppressor nor an amplifier of selection when $N = 30$. However, in both cases, the coupled star graph is more suppressing than the one-layer star graph, which is a strong amplifier of selection. This last result is consistent with Theorem 4.10 and Remark 4.11.

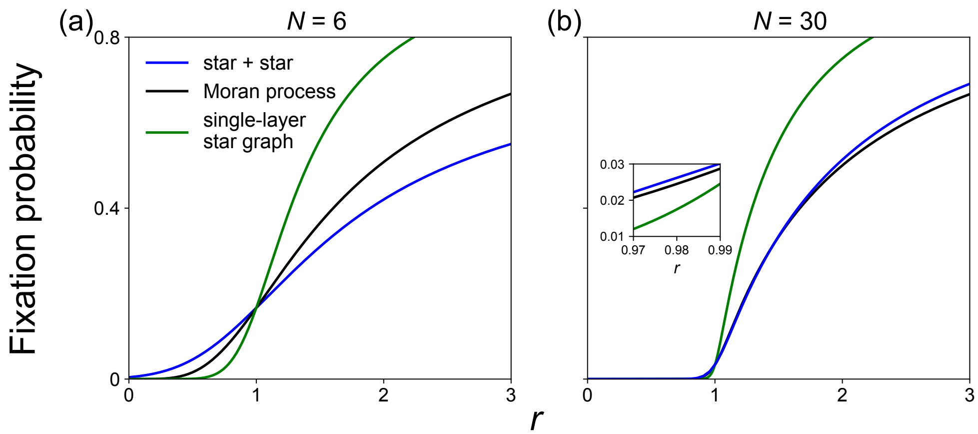

FIG. 4. *Fixation probability for coupled star graphs under model 1. We compare the results for the coupled star graphs, shown in blue, with those for the Moran process, shown in black, and those for the single-layer star graphs, shown in green. (a) $N = 6$. (b) $N = 30$. The inset in (b) magnifies the result for r values less than and close to $r = 1$.*

5.5. Combination of the complete graph and complete bipartite graph. Consider two-layer networks in which layer 1 is the complete graph and layer 2 is the complete bipartite graph K_{N_1,N_2}, where $N_1 + N_2 = N$. The complete bipartite graph K_{N_1,N_2} has two disjoint subsets of nodes V_1 and V_2 with N_1 and N_2 nodes, respectively. Each node in V_1 is adjacent to each node in V_2. We can describe the state of the evolutionary dynamics by an 8-tuple. For each 8-tuple state, we distinguish 17 types of transition events and can derive the transition probability of each state to each state. We show the calculations of the transition probability matrix in section S7.

We use the same numerical method to solve Eqs. (5.2) and (5.3) for this two-layer network. We show the fixation probability for $N = 6$ in Figure 5(a) and 5(b), and $N = 20$ in Figure 5(c) and 5(d), respectively. We reduce the larger N value to 20 due to large memory requirement for this network. We set $N_1 = N_2 = N/2$ in Figure 5(a) and 5(c). In this case, a one-layer network K_{N_1,N_2} is a regular graph and therefore an isothermal graph. We set $N_2 \approx 2N_1$, where $\approx$ represents "approximately equal to", in Figure 5(b) and 5(d). Figure 5 shows that these two-layer networks are suppressors of selection.

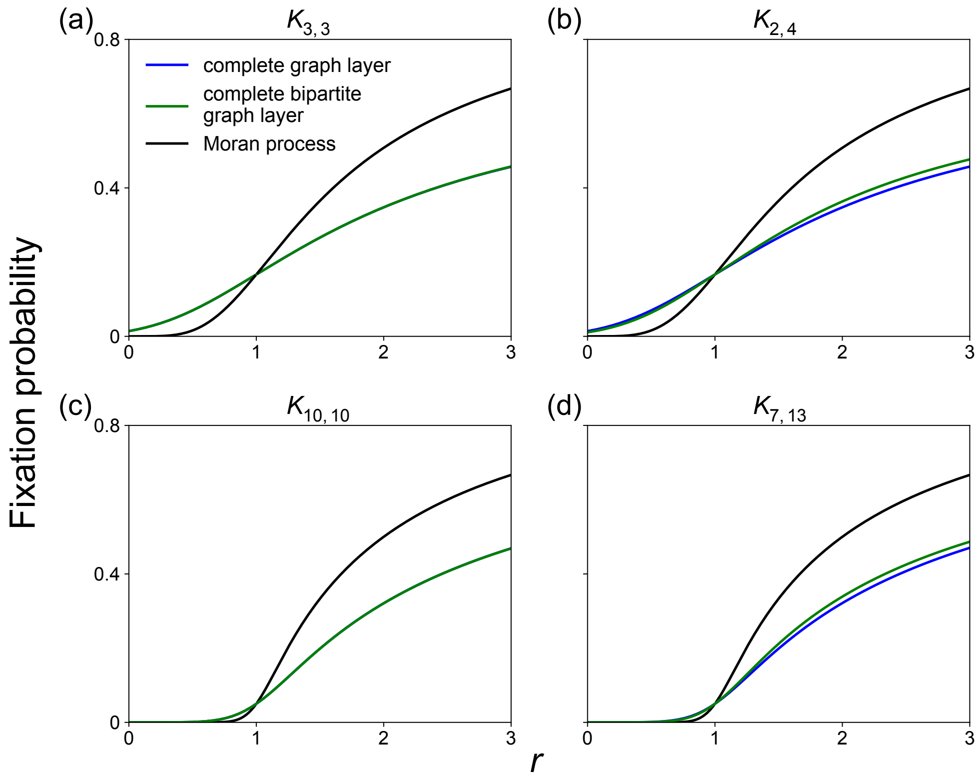

FIG. 5. *Fixation probability for two-layer networks composed of a complete graph layer and a complete bipartite graph layer under model 1. (a) $N = 6$ with $K_{3,3}$. (b) $N = 6$ with $K_{2,4}$. (c) $N = 20$ with $K_{10,10}$. (d) $N = 20$ with $K_{7,13}$. In panels (a) and (c), the results for the complete graph layer are close to those for the complete bipartite graph layer such that the blue lines are almost hidden behind the green lines.*

5.6. Combination of the complete graph and two-community networks. Empirical networks often have community (i.e., group) struture [52]. Therefore, in this section, we consider two-layer networks in which layer 1 is the complete graph and layer 2 is a weighted network with two communities. Specifically, layer 2 is composed of two disjoint sets of nodes V_1 and V_2 with N_1 and N_2 nodes, respectively, where $N_1 + N_2 = N$. Each set of nodes forms a clique (i.e., complete graph as a subgraph) with edge weight 1. In addition, each node in V_1 is connected with each node in V_2 with edge weight $\bar{\epsilon}$. A small $\bar{\epsilon}$ implies a strong community structure. It should also be noted that the combination of the complete graph and the complete bipartite graph corresponds to this model in the limit of $\bar{\epsilon} \to \infty$. Similar to the case of combination of the complete graph and the complete bipartite network, we use an 8-tuple and distinguish 17 types of transition events from each state to another state. We show the calculations of the transition probability matrix in section S8.

We numerically solve Eqs. (5.2) and (5.3) for $N = 6$ and $N = 20$ with $\bar{\epsilon} = 0.1$. We show the results for $N = 6$ in Figure 6(a) and 6(b), and $N = 20$ in Figure 6(c) and 6(d), respectively. In Figure 6(a) and 6(c), we set $N_1 = N_2 = N/2$, and the two-community network is an isothermal graph. In Figure 6(b) and 6(d), we set $N_2 \approx 2N_2$. Figure 6 indicates that these two-layer networks are suppressors of selection.

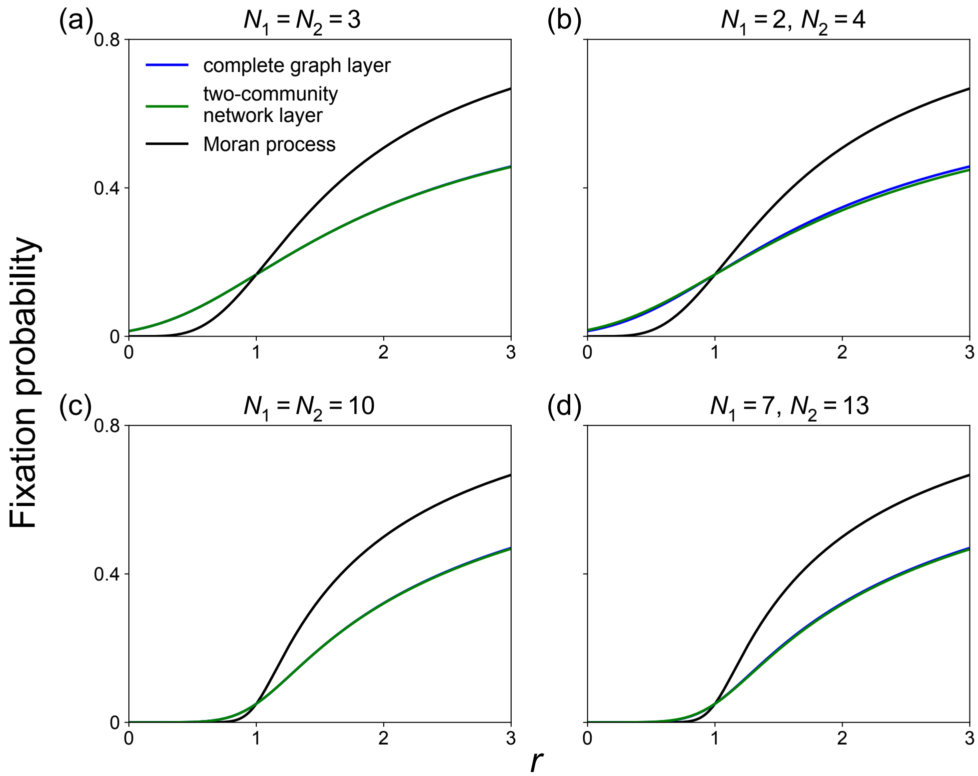

FIG. 6. *Fixation probability for two-layer networks composed of a complete graph layer and a two-community network layer under model 1. (a) $N_1 = N_2 = 3$ and $\bar{\epsilon} = 0.1$. (b) $N_1 = 2$, $N_2 = 4$, and $\bar{\epsilon} = 0.1$. (c) $N_1 = N_2 = 10$ and $\bar{\epsilon} = 0.1$. (d) $N_1 = 7$, $N_2 = 13$, and $\bar{\epsilon} = 0.1$.*

5.7. Death-birth process variant of model 1. We have analyzed model 1, which is a two-layer Bd process. To assess the robustness of our main result that two-layer networks are mostly suppressors of selection, in this section, we consider a variant of model 1 in which we replace the Bd updating rule by the death-birth updating rule with selection on the birth, often referred to as the dB rule [6,8,16,36]. According to the dB rule, we select an individual uniformly at random for death in each time step. Then, the neighbors of the dying individual compete to reproduce its type on the vacant site with probability proportional to their fitness. The fixation probability for this death-birth process in the case of the well-mixed population (i.e., unweighted complete graph) is [44]

$$\rho^{\mathrm{dB}} = \left(1 - \frac{1}{N}\right) \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^{N-1}}}. \tag{5.16}$$

We extend the dB process to the case of two-layer networks. For simplicity, we only consider the case in which both layers are complete graphs (i.e., coupled complete graph). In each time step, an individual selected uniformly at random (i.e., with probability $1/N$) dies. We then select one of the two layers to operate the dB process with equal probability, i.e., $1/2$. The neighbors of the dying individual in the selected layer compete for filling the empty site with probability proportional to the product of their fitness and the edge weight (which we set to 1 because we are considering unweighted complete graphs for both layers). As we show in section S9, we can derive the set of $\binom{N+3}{3} - 4$ linear equations with which to calculate the fixation probability similarly to the case of the Bd process on the coupled complete graph.

We show the fixation probability for this death-birth process on coupled complete graphs with $N = 6$ and $N = 30$ in Figure 7(a) and 7(b), respectively. We find that these coupled complete graphs are suppressors of selection under the dB rule, relative to the Moran process. The green lines in Figure 7 represent Eq. (5.16). We find that the coupled complete graphs under the dB rule are also more suppressing than the one-layer complete graphs under the same dB rule.

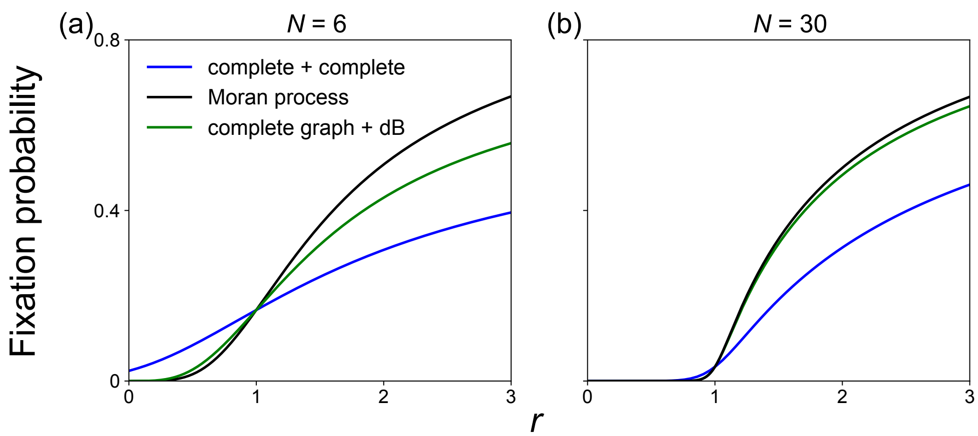

FIG. 7. *Fixation probability for coupled complete graphs under the variant of model 1 with the dB updating. (a) $N = 6$. (b) $N = 30$. The green lines represent Eq.* (5.16).

5.8. Model 2 on coupled complete graphs. In this section, we consider model 2 in which both layers are complete graphs. We can describe the state of the network by the same 4-tuple $\boldsymbol{i} = (i_1, i_2, i_3, i_4)$ as that we used in section 5.2 and derive a set of $\binom{N+3}{3} - 4$ linear equations to determine the fixation probability. For each state (i_1, i_2, i_3, i_4), we distinguish 21 types of events and obtain the transition probability from each state to another state, as shown in section S10.

We show the fixation probability for the mutant type for $N = 6$ and $N = 30$ by the dashed lines in Figure 2(a) and 2(b), respectively. We find that these coupled complete graphs are suppressors of selection. Furthermore, the results for model 2 are close to those for model 1, while model 2 is slightly less suppressing than model 1.

6. Numerical results. In this section, we carry out numerical simulations of the Bd process on four two-layer networks without particular symmetry, i.e., a coupled Erdős-Rényi (ER) random graph, a coupled Barabási-Albert (BA) network, and two empirical two-layer networks. To generate a two-layer ER random graph with $N = 100$ individuals, in each layer, we connected each pair of nodes with probability 0.1. We iterated generating networks from the ER random graph with $N = 100$ nodes until we obtained two connected networks, which we used as two layers. The two generated networks had $M_1 = 498$ edges and $M_2 = 500$ edges, respectively. To generate a two-layer BA network, we sampled two networks with $N = 100$ nodes each from the BA model [53]. In the network growth process of the BA model, each incoming node is connected to five already existing nodes according to the linear preferential attachment rule. We use the star graph on 6 nodes as the initial network in each layer. Each of the two generated networks was more heterogeneous than the ER graph in terms of the node's degree, was connected, and had $M = 475$ edges. Without loss of generality, we uniformly randomly permuted the label of all nodes in layer 2. Otherwise, there would be strong positive correlation between the degree of the two replica nodes of the same individual. One empirical network is the Vickers-Chan 7thGraders (VC7) network, which is a two-layer network of scholastic and friendship relationships among $N = 29$ seventh grade students in a school in Victoria, Australia, with $M_1 = 126$ edges in layer 1 and $M_2 = 152$ edges in layer 2 [54]. The second empirical network is the Lazega Law Firm (LLF) network, which is a two-layer network of professional and cooperative relationships among $N = 71$ partners at the LLF, with $M_1 = 717$ edges in layer 1 and $M_2 = 726$ edges in layer 2 [55].

We focus on model 1 and examine whether these two-layer networks tend to be suppressors of selection. We initially placed a mutant on just one replica node in each layer. Therefore, there are $N \times N$ possible initial states. To calculate the fixation probability for a single mutant for each layer, we run the Bd process until the mutant type or the resident type fixates in the selected layer. For each value of r, we run $20N^2$ simulations starting from each of the N^2 initial conditions 20 times. We obtain the fixation probability of the mutant type for each layer as the number of the runs in which the mutant type has fixated in the selected layer divided by $20N^2$.

Figure 8 shows the relationship between the fixation probability for a single mutant and r for the four two-layer networks, one per panel. The figure shows that both layers are suppressors of selection in all the four two-layer networks. Unexpectedly, we also find that the fixation probability as a function of r is similar between the two layers, which are different networks in terms of edges, in all the four two-layer networks.

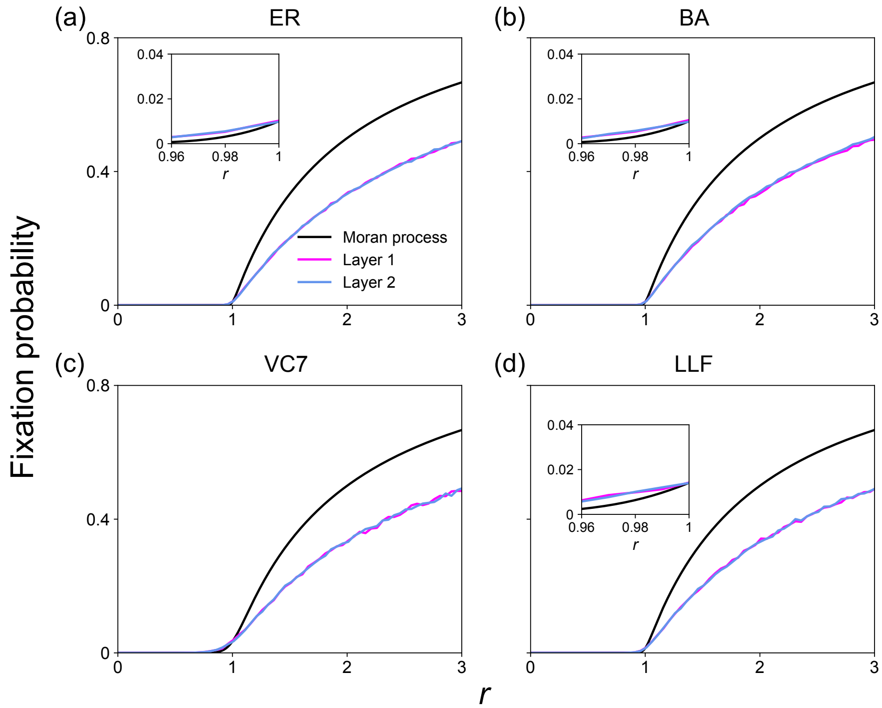

FIG. 8. *Fixation probability in model and empirical two-layer networks. (a) Two-layer ER graph. (b) Two-layer BA model. (c) Vickers-Chan 7thGraders network (VC7). (d) Lazega Law Firm network (LLF). The insets of (a), (b), and (d) magnify the results for values of r less than and close to* $r = 1$.

7. Discussion. Inspired by an evolutionary game model on two-layer networks [35], we formulated and analyzed constant-selection dynamics on two-layer networks in which each individual's fitness is defined to be the sum of the fitness of the replica node over the two layers. Using martingales, we proved that two-layer networks are suppressors of selection if one layer is a particular network with high symmetry, at least relative to that network considered as a single-layer network. The single-layer regular graphs, including the complete graph, cycle, and bipartite complete graphs in which the two parts have same number of nodes, are isothermal graphs [4, 56], i.e., equivalent to the Moran process. Therefore, these theorems show that two-layer networks that have any of these networks as one layer are suppressors of selection regardless of the second layer. Furthermore, we semi-analytically analyzed some two-layer networks in which both layers are highly symmetric networks to show that they are also suppressors of selection except the coupled star graph with $N = 30$ nodes. Nonetheless, the couple star graph with $N = 30$ is more suppressing than the single star graph with $N = 30$. Numerical simulations of stochastic evolutionary dynamics on four larger two-layer networks without particular symmetry have also shown that these networks are suppressors of selection. Overall, we have provided mathematical results and compelling numerical evidence that two-layer networks are suppressors of selection unless both layers are strong amplifiers of selection such as the star graph (see Figure 4).

We argue that the intuitive reason behind this result is the key assumption of our model that the total fitness of a replica node depends on the fitness of the corresonding replica node in the other layer as well as its own fitness. Suppose that $r > 1$ and that a replica node i in layer 1 is of resident type. Then, intuitively, it is more likely to be invaded by a mutant type, if a neighbor is a mutant because the mutant's fitness ($= r$) is higher than the resident's fitness ($= 1$). However, if the replica node i in layer 2 is of mutant type, the total fitness for the ith individual is equal to $1 + r$. Therefore, the mutant type in layer 2 boosts the likelihood that ith individual reproduces in layer 1 relative to the case of a single-layer network. In this manner, the two-layer nature of the model blurs the effect of fitness due to the interference of one layer into constant-selection dynamics in the other layer. This is why two-layer networks are expected to be suppressors of selection, at least relative to their one-layer counterparts. We note that we exploited this intuition in formulating and proving our theorems using martingales.

As we reviewed in section 1, most networks are amplifiers of selection under the Bd process and uniform initialization. However, small directed networks [16], metapopulation model networks [17,18], a type of temporal network called a switching network (i.e., in which the network switches between two static network with regular or irregular time intervals) when N is small [19], and hypergraphs [20] tend to be suppressors of selection under the same conditions (i.e., the Bd process with uniform initialization) even if the undirected variant of them is an amplifier of selection. Here we add two-layer networks as another case in which suppressors of selection are common. These results altogether suggest that amplifiers of selection under the Bd process with uniform initialization are not so common as was initially considered. It is straightforward to extend our models to the case of more than two layers. Constant-selection dynamics under adaptive networks (i.e., time-varying networks in which network changes are induced by the state, or type, of the nodes) are underexplored [57,58]. Whether or not these networks are amplifying or suppressing would be a worthwhile investigation.

We have exploited some highly symmetric networks to be used as network layers with the aim of reducing the number of linear equations to be solved from $O(2^{2N})$ to a polynomial order of N. We used the same strategy to analyze hypergraphs [20] and switching temporal networks [19]. The same technique was exploited for analytically solving the fixation probability in the complete bipartite graphs [59,60], stars [4,56], and so-called superstars [4]. However, the size of the two-layer networks for which we exactly calculated the fixation probability is still modest, i.e., up to $N = 20$ or $N = 30$ depending on the network. This is because, the network in our models has two layers, and we need to track the type (i.e., resident or mutant) of the replica nodes in both layers to specify the state of an individual. It would be ideal if this type of mathematical technique leads to analytical solutions of the fixation probability, not just to reduce the dimension of the problem. This is left for future work.

Acknowledgments. We thank Alex McAvoy for fruitful discussion. N.M. acknowledges support from the Japan Science and Technology Agency (JST) Moonshot R&D (under Grant No. JPMJMS2021), the National Science Foundation (under Grant Nos. 2052720 and 2204936), and JSPS KAKENHI (under grant Nos. JP 21H04595 and 23H03414).

REFERENCES

[1] Nowak MA, Evolutionary Dynamics, Harvard Press, Cambridge, MA, 2006.
[2] Sigmund K, The calculus of selfishness, Princeton University Press, Princeton, NJ, 2010.
[3] Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, and Szolnoki A, Statistical physics of human cooperation, Phys. Rep., 687 (2017), 1–51.

[4] Lieberman E, Hauert C, and Nowak MA, Evolutionary dynamics on graphs, Nature, 433 (2005), 312–316.
[5] Szabó G and Fáth G, Evolutionary games on graphs, Phys. Rep., 446 (2007), 97–216.
[6] Shakarian P, Roos P, Johnson A, A review of evolutionary graph theory with applications to game theory, Biosystems, 107 (2012), 66–80.
[7] Perc M, Gómez-Gardeñes J, Szolnoki A, Floría LM, and Moreno Y, Evolutionary dynamics of group interactions on structured populations: a review, J. R. Soc. Interface., 10 (2013), 20120997.
[8] Ohtsuki H, Hauert C, Lieberman E, Nowak MA, A simple rule for the evolution of cooperation on graphs and social networks, Nature, 441 (2006), 502–505.
[9] Allen B, Lippner G, Chen Y, Fotouhi B, Momeni N, Yau S, et al, Evolutionary dynamics on any population structure, Nature, 544 (2017), 227–230.
[10] Hindersin L and Traulsen A, Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics, PLoS Comput. Biol., 11 (2015), e1004437.
[11] Cuesta FA, Sequeiros PG, and Rojo ÁL, Evolutionary regime transitions in structured populations, PLoS ONE, 13 (2018), e0200670.
[12] Allen B, Sample C, Steinhagen P, Shapiro J, King M, Hedspeth T, et al, Fixation probabilities in graph-structured populations under weak selection, PLoS Comput. Biol., 17 (2021), e1008695.
[13] Cuesta FA, Sequeiros PG, and Rojo ÁL, Suppressors of selection, PLoS ONE, 12 (2017), e0180549.
[14] Adlam B, Chatterjee K, Nowak MA, Amplifiers of selection, Proc. Math. Phys. Eng. Sci., 471 (2015), 20150114.
[15] Pavlogiannis A, Tkadlec J, Chatterjee K, and Nowak MA, Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory, Commun. Biol., 1 (2018), 71.
[16] Masuda N, Directionality of contact networks suppresses selection pressure in evolutionary dynamics, J. Theoret. Biol., 258 (2009), 323–334.
[17] Yagoobi S, Traulsen A, Fixation probabilities in network structured metapopulations, Sci. Rep., 11 (2021), 17979.
[18] Marrec L, Lamberti I, Bitbol A, Toward a universal model for spatially structured populations, Phys. Rev. Lett., 127 (2021), 218102.
[19] Bhaumik J and Masuda N, Fixation probability in evolutionary dynamics on switching temporal networks, J. Math. Biol., 87 (2023), 64.
[20] Liu R and Masuda N, Fixation dynamics on hypergraphs, PLoS Comput. Biol., 19 (2023), e1011494.
[21] Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, and Porter MA, Multilayer networks, J. Complex Netw., 2 (2014), 203–271.
[22] Boccaletti S, Bianconi G, Criado R, del Genio CI, Gómez-Gardeñes J, Romance M, et al, The structure and dynamics of multilayer networks, Phys. Rep., 544 (2014), 1–122.
[23] Bianconi C, Multilayer networks, Oxford University Press, Oxford, UK, 2018.
[24] De Domenico M, More is different in real-world multilayer networks, Nat. Phys., 19 (2023), 1247–1262.
[25] Wang Z, Wang L, Szolnoki A, and Perc M, Evolutionary games on multilayer networks: a colloquium, Eur. Phys. J. B, 88 (2015), 124.
[26] Ohtsuki H, Nowak MA, and Pacheco JM, Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs, Phys. Rev. Lett., 98 (2007), 108106.
[27] Ohtsuki H, Pacheco JM, and Nowak MA, Evolutionary graph theory: Breaking the symmetry between interaction and replacement, J. Theoret. Biol., 246 (2007), 681–694.
[28] Wang Z, Wang L, Perc M, Degree mixing in multilayer networks impedes the evolution of cooperation, Phys. Rev. E, 89 (2014), 052813.
[29] Su Q, Li A, Wang L, and Eugene Stanley H, Spatial reciprocity in the evolution of cooperation, Proc. Royal Soc. B, 286 (2019), 20190041.
[30] Li G and Sun X, Evolutionary game on a growing multilayer network, Physica A, 578 (2021), 126110.
[31] Chen W and Wu T, Adaptive multilayer networks resolve the cooperation dilemma induced by breaking the symmetry between interaction and learning, New J. Phys., 23 (2021), 093019.
[32] Inaba M and Akiyama E, Evolution of cooperation in multiplex networks through asymmetry between interaction and replacement, Sci. Rep., 13 (2023), 9814.
[33] Duh M, Gosak M, Slavinec M, and Perc M, Assortativity provides a narrow margin for enhanced cooperation on multilayer networks, New J. Phys., 21 (2019), 123016.

[34] Allen JM and Hoyle RB, Asynchronous updates can promote the evolution of cooperation on multiplex networks, Physica A, 471 (2017), 607–619.
[35] Su Q, McAvoy A, Mori Y, and Plotkin JB, Evolution of prosocial behaviours in multilayer populations, Nat. Hum. Behav., 6 (2022), 338–348.
[36] Pattni K, Broom M, Rychtář J, and Silvers J, Evolutionary graph theory revisited: when is an evolutionary process equivalent to the Moran process? Proc. Math. Phys. Eng. Sci., 471 (2015), 20150334.
[37] Chalub FACC, An asymptotic expression for the fixation probability of a mutant in star graphs, J. Dyn. Games, 3 (2016), 217–223.
[38] Askari M, Samani KA, Analytical calculation of average fixation time in evolutionary graphs, Phys. Rev. E, 92 (2015), 042707.
[39] Pavlogiannis A, Tkadlec J, Chatterjee K, and Nowak MA, Amplification on undirected population structures: Comets beat stars, Sci. Rep., 7 (2017), 82.
[40] Giakkoupis G, Amplifiers and suppressors of selection for the Moran process on undirected graphs, arXiv:1611.01585 (2016).
[41] Galanis A, Göbel A, Goldberg LA, Lapinskas J, and Richerby D, Amplifiers for the Moran process, J. ACM, 64 (2017), 5, 1–90.
[42] Goldberg LA, Lapinskas J, Lengler J, Meier F, Panagiotou K, and Pfister P, Asymptotically optimal amplifiers for the Moran process, Theor. Comput. Sci., 758 (2019), 73–93.
[43] Monk T, Green P, and Paulin M, Martingales and fixation probabilities of evolutionary graphs, Proc. Math. Phys. Eng. Sci., 470 (2014), 20130730.
[44] Kaveh K, Komarova NL, and Kohandel M, The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem, R. Soc. Open Sci., 2 (2015), 140465.
[45] Allen B, Sample C, Jencks R, Withers J, Steinhagen P, Brizuela L, et al, Transient amplifiers of selection and reducers of fixation for death-Birth updating on graphs, PLoS Comput. Biol., 16 (2020), e1007529.
[46] Donnelly P, Welsh D, Finite particle systems and infection models, Math Proc. Camb. Phil. Soc., 94 (1983), 167–182.
[47] Masuda N, Ohtsuki H, Evolutionary dynamics and fixation probabilities in directed networks, New J. Phys., 11 (2009), 033012.
[48] Broom M, Hadjichrysanthou C, and Rychtář J, Two results on evolutionary processes on general non-directed graphs, Proc. Math. Phys. Eng. Sci., 466 (2010), 2795–2798.
[49] Feller W, An Introduction to Probability Theory and Its Applications, Volume II. John Wiley & Sons, New York, NY, 1971.
[50] Durrett R, Probability Theory and Examples, Duxbury Press, Belmont, CA, 1996.
[51] Monk T, Martingales and the fixation probability of high-dimensional evolutionary graphs, J. Theoret. Biol., 451 (2018), 10–18.
[52] Fortunato S, Community detection in graphs, Phys. Rep., 486 (2010), 75–174.
[53] Barabási A-L and Albert R, Emergence of scaling in random networks, Science, 286 (1999), 509–512.
[54] Vickers M and Chan S, Representing Classroom Social Structure, Melbourne: Victoria Institute of Secondary Education, 1981. Available at https://networks.skewed.de/net/7th_graders, accessed on July 20, 2023.
[55] Emmanuel L, The Collegial Phenomenon: The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership, Oxford University Press, Oxford, UK, 2001.
[56] Broom M, Rychtář J, An analysis of the fixation probability of a mutant on special classes of non-directed graphs, Proc. Math. Phys. Eng. Sci., 464 (2008), 2609–2627.
[57] Tkadlec J, Kaveh K, Chatterjee K, Nowak MA, Natural selection of mutants that modify population structure, arXiv preprint, arXiv: 2111.10890 (2021).
[58] Melissourgos T, Nikoletseas SE, Raptopoulos CL, and Spirakis PG, An extension of the moran process using type-specific connection graphs, J. Comput. System Sci., 124 (2022), 77–96.
[59] Zhang C, Wu Y, Liu W, and Yang X, Fixation probabilities on complete star and bipartite digraphs, Disc. Dyn. Nat. Soc., 2012 (2012), 940465.
[60] Voorhees B and Murray A, Fixation probabilities for simple digraphs, Proc. Math. Phys. Eng. Sci., 469 (2013), 20120676.