Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Replicator--Mutator Dynamics on Networks with Spatial Nonlocality: Theory and Applications

Porikshit Mondal
Institute for Cross-Disciplinary Physics and Complex Systems
Ernesto Estrada
estrada@ifisc.uib-csic.es

Institute for Cross-Disciplinary Physics and Complex Systems

Research Article

Keywords:

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11222200/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

REPLICATOR–MUTATOR DYNAMICS ON NETWORKS WITH SPATIAL NONLOCALITY: THEORY AND APPLICATIONS

PORIKSHIT MONDAL AND ERNESTO ESTRADA

Abstract. The replicator–mutator model provides a general framework for studying the emergence of social norms and dominant behavioral patterns in networked populations. In the traditional network formulation, mutation and observational learning operate through nearest-neighbor interactions. In realistic settings, however, agents may also be influenced by others beyond their immediate neighbors, with the strength of influence typically decaying with distance. Here we investigate how evolutionary dynamics change when these nonlocal interactions are incorporated explicitly. We introduce a generalized interaction operator based on d-path Laplacians, which combines interactions across all shortest-path distances. We then analyze how the range of the coupling affects evolutionary behavior on different network topologies. Our analytical results relate mutation-induced mixing and local convergence to the spectral structure of the corresponding nonlocal random-walk Laplacian. Numerical experiments show that stronger nonlocality reduces the mutation intensity required to attain high diversity and may produce an abrupt transition from behavioral flocking to an almost uniform distribution of behaviors. Thus, long-range influence promotes globalization in the sense of increased effective diversity together with greater homogeneity of behavioral frequencies.

1. Introduction

The replicator–mutator equation constitutes one of the most important mathematical frameworks for describing evolutionary systems in which selection and mutation act simultaneously. In its classical form, the replicator equation models the competition among strategies, types, or behaviors through frequency-dependent selection, such that strategies with fitness larger than the population average increase their relative abundance [36, 24]. The replicator–mutator equation extends this formulation by introducing mutation, exploration, learning errors, or spontaneous switching between strategies, thereby accounting for imperfect replication and stochastic behavioral change [29, 18, 34]. Because of this additional mechanism, the model naturally unifies several important evolutionary descriptions. In the absence of mutation, it reduces to the classical replicator dynamics of evolutionary game theory, while for frequency-independent fitness landscapes it becomes closely related to the quasispecies equation of molecular evolution [18, 40, 38]. More generally, replicator–mutator systems can also be connected to the Price equation and to broader formulations of adaptive evolutionary dynamics [29]. Owing to this flexibility, the replicator–mutator framework has found applications in population genetics, autocatalytic chemical systems, social dilemmas, language evolution, collective decision making, and adaptive social behavior [17, 35, 23]. In recent years, increasing attention has been devoted to the study of replicator–mutator dynamics

in complex networks, where the topology of interactions fundamentally shapes the evolutionary outcomes.

The incorporation of graphs and networks into replicator–mutator dynamics represents a major shift away from the well-mixed approximation traditionally assumed in evolutionary game theory. In graph-structured systems, interactions are constrained by the underlying network topology, which determines who interacts with whom, how information or behaviors propagate, and how mutation redistributes evolutionary mass across the system [30, 22]. In this context, topology is not simply a passive substrate but an active driver of the dynamics. Degree heterogeneity, clustering, directionality, and higher-order motifs can strongly modify fixation probabilities, convergence times, coexistence regimes, and oscillatory behavior [19, 20, 21, 31]. It has been shown that networks may either amplify or suppress selection depending on their structure, while the interplay between mutation and topology can generate rich bifurcation structures, including Hopf bifurcations and stable limit cycles [30, 25]. These phenomena have been observed not only in biological evolution but also in social and behavioral contexts, where networked populations exhibit persistent oscillations, cyclic dominance, and topology-dependent equilibrium selection [23, 35]. Moreover, recent studies have demonstrated that higher-order motifs and rugged fitness landscapes can profoundly influence evolutionary trajectories, particularly in systems involving multiple mutations or valley-crossing processes [19, 20, 21]. These results collectively indicate that realistic evolutionary systems cannot, in general, be understood without explicitly incorporating network structure.

Among graph-based formulations of replicator–mutator dynamics, the model introduced by Olfati-Saber occupies a particularly important position. In this framework, behaviors are represented as nodes of a weighted undirected graph, while edges encode the reward or payoff associated with transitions between behaviors [28]. Mutation is modeled through a graph-dependent stochastic operator, leading naturally to a networked replicator–mutator dynamics capable of describing behavioral evolution in social systems. One of the most remarkable aspects of the Olfati–Saber formulation is the emergence of distinct macroscopic evolutionary phases, including behavioral flocking, partial coexistence, and collapse of dominance, depending on the mutation rate and network structure [28, 16]. The model also establishes deep links between evolutionary dynamics, consensus processes, and multi-agent systems [23]. In this sense, the Olfati–Saber framework provides a natural mathematical setting for studying collective behavior, social learning, and network-mediated adaptation. However, despite its richness, the interactions in the original model remain fundamentally local, being restricted to nearest-neighbor transitions encoded in the adjacency structure of the graph.

In many real systems, however, influence and learning are not purely local. Individuals may imitate behaviors that are not directly adjacent in the behavioral graph, social trends may spread through indirect exposure, and innovations may influence distant regions of a network without requiring immediate local contact. Similar forms of long-range interactions are ubiquitous in biological, technological, and social systems. This observation motivates the introduction of spatial nonlocality into replicator–mutator dynamics. The key question addressed in the present work is therefore the following: how does the evolutionary behavior of the Olfati–Saber model change when mutation and observational learning are allowed to occur

across multiple graph distances rather than only between nearest neighbors? In particular, we are interested in understanding whether long-range interactions can promote globalization in the system, understood as the regime in which no single behavior dominates and the effective number of active behaviors approaches the total number of available behaviors.

To introduce such spatial nonlocality, we employ the theory of d-path Laplacians, also known as k-path Laplacians, originally introduced by Estrada [8] and then further developed in several works [13, 14, 9]. The standard graph Laplacian describes local diffusion through nearest-neighbor interactions only. By contrast, the d-path Laplacian encodes interactions between vertices whose shortest-path distance is exactly d, thereby providing a hierarchy of operators that resolve the graph into different distance layers [15, 5]. This framework makes it possible to define generalized diffusion processes with long-range hops on graphs. In particular, transformed sums of d-path Laplacians weighted through Mellin or Laplace kernels interpolate continuously between local and nonlocal regimes [13, 14]. Mellin transforms assign weights of the form d^{-s}, producing power-law decay of long-range interactions, while Laplace transforms assign exponentially decaying weights $e^{-\lambda d}$. These operators have been extensively studied in the context of anomalous diffusion, random walks, synchronization, and quantum transport on graphs [9, 37, 12, 2, 4]. Remarkably, Mellin-transformed d-path Laplacians can generate superdiffusive transport regimes on infinite lattices and complex networks, whereas Laplace-transformed operators generally recover normal diffusion [13, 14]. Related nonlocal diffusion models have also been explored in generalized graph diffusion equations, metaplex networks, and random multi-hopper systems [6, 10, 11].

The present work transfers these ideas from diffusion theory to evolutionary dynamics. Instead of using transformed d-path Laplacians to describe the transport of physical particles or probability densities, we use them to construct a generalized nonlocal mutation operator for the replicator–mutator equation. In this way, observational learning and behavioral transitions are no longer restricted to nearest-neighbor relations but may occur across arbitrary graph distances with strengths controlled by Mellin or Laplace kernels. Small values of the Mellin exponent correspond to strong long-range influence, whereas large exponents recover the local Olfati–Saber model as a limiting case. This construction provides a mathematically controlled framework for investigating how nonlocal interactions reshape evolutionary phases in graph-structured populations. As we show in the computational section, long-range interactions can substantially reduce the mutation rate required to reach high diversity and may induce abrupt transitions from behavioral flocking to complete collapse. The resulting theory therefore establishes a direct connection between replicator–mutator dynamics, networked behavioral evolution, and nonlocal graph diffusion, providing a new framework for understanding how long-range observational learning promotes globalization in evolutionary systems.

2. Preliminaries and setting the language

In this section we describe the fundamental description of the models to be generalized here as well as some of the basic mathematical results needed for the comprehension of our expansion and new results.

2.1. **Olfati–Saber Replicator–Mutator Dynamics on Graphs.** In this section we introduce the replicator–mutator framework proposed by Olfati–Saber [28] and establish the notation used throughout the paper.

Definition 1 (Behavioral state space)**.** Let

$$V = \{b_1, \dots, b_n\} \tag{2.1}$$

be set of possible behaviors (or strategies). The state of the population is represented by the vector

$$\vec{x} = [x_1, \dots, x_n]^T, \tag{2.2}$$

where x_i denotes the fraction of individuals adopting behavior b_i. The state vector satisfies

$$x_i \geq 0, \qquad \sum_{i=1}^{n} x_i = 1. \tag{2.3}$$

Thus, the dynamics evolve on the probability simplex

$$\Delta^{n-1} = \left\{ x \in \mathbb{R}^n : x_i \geq 0,\ \sum_{i=1}^{n} x_i = 1 \right\}. \tag{2.4}$$

Definition 2 (Interaction graph)**.** Behavioral interactions are represented by a weighted undirected graph

$$\Gamma = (V, E, A), \tag{2.5}$$

with weighted adjacency matrix

$$A = [a_{ij}] \in \mathbb{R}^{n \times n}, \qquad a_{ij} \geq 0. \tag{2.6}$$

The quantity a_{ij} represents the payoff associated with an interaction between behaviors b_i and b_j. Throughout this work we assume that self-loops are present:

$$a_{ii} = 1, \qquad \forall i \in V. \tag{2.7}$$

So we can rewrite the weighted adjacency matrix in matrix form,

$$A = [a_{ij}] = a \cdot A_{\text{adj}} + I, \tag{2.8}$$

where $a > 0$ denotes the weight or reward parameter, I represents the identity matrix, and A_{adj} is the adjacency matrix of the graph.

Definition 3 (Normalized interaction matrix)**.** Let

$$k_i = \sum_{j=1}^{n} a_{ij}, \tag{2.9}$$

and define the degree matrix

$$K = \text{diag}(k_1, \dots, k_n). \tag{2.10}$$

The normalized interaction matrix is

$$P = K^{-1} A. \tag{2.11}$$

Equivalently,

$$p_{ij} = \frac{a_{ij}}{\sum_j a_{ij}}. \tag{2.12}$$

The matrix P is row-stochastic:

$$P\mathbf{1} = \mathbf{1}. \tag{2.13}$$

Definition 4 (Fitness and average fitness)**.** The fitness associated with behavior b_i is

$$f_i = \sum_{j=1}^{n} a_{ij} x_j. \tag{2.14}$$

In vector form,

$$\vec{f} = A\vec{x}. \tag{2.15}$$

The average population fitness is

$$\phi = \sum_{i=1}^{n} f_i x_i = \vec{x}^T A \vec{x}. \tag{2.16}$$

Definition 5 (Replicator–mutator dynamics)**.** Let

$$Q = [q_{ij}] \tag{2.17}$$

be a row-stochastic mutation matrix, where q_{ij} denotes the probability that an offspring of an individual adopting behavior b_i adopts behavior b_j.

The replicator–mutator dynamics are defined by

$$\dot{x}_i = \sum_{j=1}^{n} f_j x_j q_{ji} - x_i \phi, \qquad i = 1, \ldots, n. \tag{2.18}$$

Remark 6 (Replicator limit). In the absence of mutation,

$$Q = I, \tag{2.19}$$

equation (2.18) reduces to the classical replicator equation

$$\dot{x}_i = x_i(f_i - \phi). \tag{2.20}$$

2.1.1. *Matrix formulation.*

Definition 7 (Diagonal fitness operator)**.** Define the diagonal fitness matrix

$$F = \operatorname{diag}(f_1, \ldots, f_n). \tag{2.21}$$

Proposition 8 (Matrix form of the RM dynamics)**.** *The replicator–mutator dynamics can be written compactly as*

$$\dot{\vec{x}} = Q^T F \vec{x} - \phi \vec{x}. \tag{2.22}$$

In the mutation-free case,

$$Q = I, \tag{2.23}$$

equation (2.22) *reduces to*

$$\dot{\vec{x}} = F\vec{x} - \phi\vec{x}. \tag{2.24}$$

Proof. Observe that

$$(F\vec{x})_i = f_i x_i. \tag{2.25}$$

Therefore,

$$(Q^T F\vec{x})_i = \sum_{j=1}^{n} q_{ji} f_j x_j, \tag{2.26}$$

which reproduces equation (2.18). □

Proposition 9 (Mass conservation)**.** *The simplex Δ^{n-1} is invariant under the replicator–mutator dynamics.*

Proof. Summing equation (2.18) over all i,

$$\sum_{i=1}^{n} \dot{x}_i = \sum_{i,j} f_j x_j q_{ji} - \phi \sum_i x_i. \tag{2.27}$$

Since Q is row-stochastic,

$$\sum_{i=1}^{n} q_{ji} = 1, \tag{2.28}$$

and therefore

$$\sum_{i,j} f_j x_j q_{ji} = \sum_j f_j x_j = \phi. \tag{2.29}$$

Using also

$$\sum_i x_i = 1, \tag{2.30}$$

we obtain

$$\sum_{i=1}^{n} \dot{x}_i = 0. \tag{2.31}$$

Hence the total mass is conserved. □

2.1.2. *Olfati–Saber mutation model.*

Definition 10 (Olfati–Saber mutation operator)**.** Following Olfati–Saber, the mutation matrix is defined as

$$q_{ij} = \begin{cases} \mu p_{ij}, & i \neq j, \\ 1 - \mu(1 - p_{ii}), & i = j, \end{cases} \tag{2.32}$$

where $\mu \geq 0$ is the mutation-intensity parameter. The operator has unit row sums for every μ. It is a nonnegative row-stochastic mutation matrix only when

$$0 \leq \mu \leq \min_i (1 - p_{ii})^{-1}. \tag{2.33}$$

For larger values, as used in some numerical experiments below, Q is understood as a mass-preserving mixing operator rather than as a matrix of transition probabilities. Further discussion and details regarding numerical tractability are provided in Section S4 of the Supplementary Information.

Definition 11 (Random-walk Laplacian)**.** The random-walk Laplacian associated with the interaction graph is

$$L_{\mathrm{rw}} = I - K^{-1}A = I - P. \tag{2.34}$$

Proposition 12 (Mutation operator representation)**.** *The mutation matrix admits the representation*

$$Q = (1-\mu)I + \mu P = I - \mu L_{\mathrm{rw}}. \tag{2.35}$$

Proof. Using

$$L_{\mathrm{rw}} = I - P, \tag{2.36}$$

we obtain

$$I - \mu L_{\mathrm{rw}} = I - \mu(I-P) = (1-\mu)I + \mu P. \tag{2.37}$$

The componentwise form coincides with the Olfati–Saber definition above. □

Corollary 13. *The replicator–mutator dynamics become*

$$\dot{\vec{x}} = (I - \mu L_{\mathrm{rw}})^T F\vec{x} - \phi\vec{x}. \tag{2.38}$$

Definition 14 (Discrete-time RM dynamics)**.** A discrete-time version of the model is given by

$$\vec{x}(t+h) = \vec{x}(t) + h\left[(I - \mu L_{\mathrm{rw}})^T F\vec{x}(t) - \phi\vec{x}(t)\right], \qquad h > 0. \tag{2.39}$$

Definition 15 (Effective number of behaviors)**.** The effective number of active behaviors is defined by

$$n_e(\vec{x}) = \left(\sum_{i=1}^{n} x_i^2\right)^{-1}. \tag{2.40}$$

Remark 16. The quantity n_e measures the effective diversity of the population. Its values range from

$$1 \le n_e \le n. \tag{2.41}$$

Small values correspond to concentration around a few dominant behaviors, while large values indicate a more homogeneous distribution over the strategy space.

Definition 17 (Order parameter)**.** An order parameter commonly used in the literature is

$$\rho = \left|\sum_{j=1}^{n} x_j e^{i\theta_j}\right|, \qquad \theta_j = \frac{2\pi j}{n}. \tag{2.42}$$

Remark 18 (Asymptotic regimes). Depending on the asymptotic value of n_e, the following qualitative regimes are been identified:

Behavioral flocking:

$$n_e \approx 1. \tag{2.43}$$

Cohesion:

$$1 < n_e \ll n. \tag{2.44}$$

Collapse:

$$1 \ll n_e < n. \tag{2.45}$$

Complete collapse:

$$n_e \approx n. \tag{2.46}$$

Note that the term "collapse" refers to the collapse of behavioural dominance, rather than to a collapse in diversity. Throughout the manuscript, we follow Olfati-Saber's convention for naming the different phases.

Figure (2.1) illustrates the evolution of behavior frequencies. In this experiment, a ring lattice topology is considered with number of nodes $n = 50$, reward parameter $a = 0.3$, and number of nearest neighbours $k = 10$. For different values of the mutation parameter, distinct evolutionary phases are observed: (a) behavioural flocking ($\mu = 0.01$), (b) cohesion ($\mu = 0.15$), (c) collapse ($\mu = 1$), and (d) complete collapse ($\mu = 12.5$). The corresponding variations in diversity (n_e) and the order parameter (ρ) are also shown in the figure.

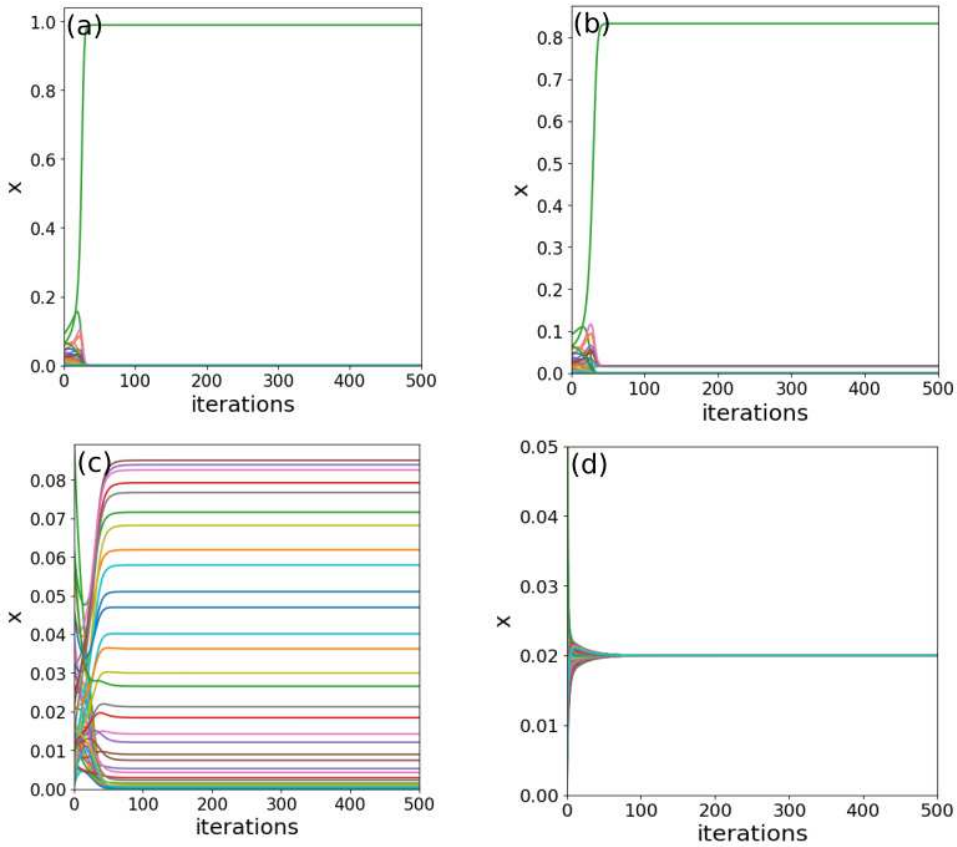

FIGURE 2.1. Evolution of behaviour frequencies on a ring lattice with $n = 50$, $k = 10$, and $a = 0.3$. Different evolutionary phases are observed for varying values of the mutation parameter μ: (a) behavioural flocking ($\mu = 0.01$), (b) cohesion ($\mu = 0.15$), (c) collapse ($\mu = 1$), and (d) complete collapse ($\mu = 12.5$). The corresponding values of diversity (n_e) and the order parameter (ρ) are shown at the top of each subplot.

Moreover, we consider a Watts–Strogatz small-world network obtained from a ring lattice by randomly rewiring one end of each link with probability p. Figure (2.2(a)) presents the diversity–mutation curve, which exhibits an S-shaped profile and displays phase transitions across four regimes, ranging from the emerging phase of behavioural flocking ($n_e = 1$) to complete collapse ($n_e \approx n$). The parameters used are $n = 50$, the initial number of nearest neighbours $k = 10$, $a = 0.3$, and $p = 0.5$.

Furthermore, Figure (2.2(b)) shows a heat-map, where the rewiring probability p and mutation rate μ are plotted against the resulting diversity n_e. Typically, small values of p are considered in order to capture the small-world property of the Watts–Strogatz network [39]. However, in our analysis we consider the full range (0 to 1) to explore the complete spectrum of network structures, from a regular ring lattice to a random network. The results indicate that as the rewiring probability increases, the system requires a smaller mutation rate μ to produce comparable levels of diversity.

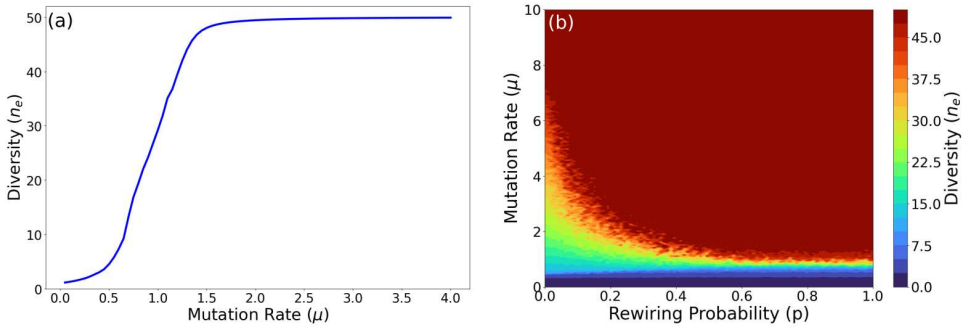

FIGURE 2.2. (a) Phase transitions in the evolution of behaviours for a Watts–Strogatz small-world network are illustrated through the S-shaped diversity–mutation curve. The rewiring probability is $p = 0.5$, while the remaining parameters are $n = 50$, $k = 10$, and $a = 0.3$. (b) Heat-map showing the effect of the rewiring probability p, varied from 0 to 1, while the other parameters remain the same.

3. Nonlocal Replicator–Mutator Dynamics on Graphs

We now generalize the local replicator–mutator model by incorporating distance-dependent interactions across the graph.

Definition 19 (d-path adjacency matrices)**.** Let

$$\operatorname{dist}(i, j) \tag{3.1}$$

denote the graph-theoretic distance between vertices v_i and v_j. For each integer

$$d = 1, \ldots, l, \tag{3.2}$$

where l is the graph diameter, define the d-path adjacency matrix

$$A_d = [(A_d)_{ij}] \tag{3.3}$$

by

$$(A_d)_{ij} = \begin{cases} 1, & \operatorname{dist}(i,j) = d, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

Remark 20. The matrices A_d encode interactions occurring at exactly distance d in the graph. If $d = 1$, then we get back to our primary adjacency matrix of a graph, i.e., $A_1 = A_{\text{adj}}$.

Definition 21 (Laplace-transformed adjacency)**.** The Laplace-type nonlocal weighted adjacency matrix is

$$A_\lambda = a \cdot \left(A_1 + \sum_{d=2}^{l} e^{-\lambda(d-1)} A_d \right) + I, \qquad \lambda > 0, \tag{3.5}$$

where $a > 0$ represents the weight or reward parameter, A_d represents the interaction contribution associated with paths of length d, and I denotes the identity matrix, which retains a unit intrinsic (self-interaction) contribution for each behavior.

Note that the identity contribution is retained so that the generalized interaction matrix preserves the unit diagonal contribution of the original local replicator–mutator formulation. Thus, the nonlocal extension modifies the off-diagonal interaction structure through the path-dependent terms while retaining the baseline self-interaction of each behavior.

Definition 22 (Mellin-transformed adjacency)**.** The Mellin-type nonlocal wighted adjacency matrix is

$$A_s = a \cdot \left(\sum_{d=1}^{l} d^{-s} A_d \right) + I, \qquad s > 0. \tag{3.6}$$

Proposition 23 (Local interaction limit)**.** *Both nonlocal transformations recover the local interaction structure in the limit of strong decay:*

$$A_\lambda \to a \cdot A_1 + I \quad \text{as } \lambda \to \infty, \tag{3.7}$$

and

$$A_s \to a \cdot A_1 + I \quad \text{as } s \to \infty. \tag{3.8}$$

Proof. By definition,

$$A_\lambda = a \cdot \left(A_1 + \sum_{d=2}^{l} e^{-\lambda(d-1)} A_d \right) + I. \tag{3.9}$$

For $\lambda \to \infty$, $\sum_{d=2}^{l} e^{-\lambda(d-1)} A_d \to 0$. Therefore,

$$A_\lambda \to a \cdot A_1 + I \qquad \text{as } \lambda \to \infty. \tag{3.10}$$

For the Mellin case,

$$A_s = a \cdot \left(\sum_{d=1}^{l} d^{-s} A_d \right) + I = \{ a \cdot A_1 + I \} + \left\{ a \cdot \left(\sum_{d=2}^{l} d^{-s} A_d \right) \right\}. \tag{3.11}$$

Since, for every $d \geq 2$,

$$d^{-s} \to 0 \qquad \text{as } s \to \infty, \tag{3.12}$$

it follows that

$$A_s \to a \cdot A_1 + I. \tag{3.13}$$

Therefore, the Mellin transform recovers the local weighted adjacency $(a \cdot A_1 + I)$ as $s \to \infty$. □

Remark 24 (Mean-field limit). In the opposite regime,

$$\lambda \to 0 \quad \text{or} \quad s \to 0, \tag{3.14}$$

all path distances contribute approximately equally, yielding an effective mean-field coupling.

3.1. **Nonlocal replicator–mutator dynamics.**

Definition 25 (Nonlocal weighted adjacency operator)**.** Define

$$A_\star = \begin{cases} A_\lambda, & \text{Laplace coupling,} \\ A_s, & \text{Mellin coupling.} \end{cases} \tag{3.15}$$

Definition 26 (Nonlocal random-walk Laplacian)**.** Let

$$K_\star = \operatorname{diag}(A_\star \mathbf{1}), \tag{3.16}$$

and define

$$L_\star = I - K_\star^{-1} A_\star. \tag{3.17}$$

Definition 27 (Nonlocal replicator–mutator dynamics)**.** The nonlocal replicator–mutator dynamics are defined by

$$\dot{\vec{x}} = (I - \mu L_\star)^T F_\star \vec{x} - \phi_\star \vec{x}, \tag{3.18}$$

where

$$F_\star = \operatorname{diag}(A_\star \vec{x}), \tag{3.19}$$

and

$$\phi_\star = \vec{x}^T A_\star \vec{x}. \tag{3.20}$$

Proposition 28 (Mass conservation)**.** *The simplex Δ^{n-1} is invariant under the nonlocal replicator–mutator dynamics.*

Proof. The proof follows exactly as in the local case, using the fact that

$$L_\star \mathbf{1} = 0. \tag{3.21}$$

□

Remark 29. The operator $A_\star$ may be interpreted as an effective payoff kernel incorporating interactions across all graph distances. The parameters λ and s control the spatial range of the interaction:

large λ or s: predominantly local interactions,
small λ or s: strong long-range interactions.

4. Spectral Interpretation of the Replicator–Mutator Dynamics

Although the replicator–mutator dynamics introduced by Olfati–Saber are nonlinear and do not admit a closed-form analytical solution in general, important qualitative and structural insights can be obtained from the spectral properties of the underlying graph Laplacian. In particular, the mutation operator defining the dynamics is directly related to the random-walk Laplacian of the interaction graph. Consequently, the eigenvalues and eigenvectors of the Laplacian determine the characteristic time scales over which mutation redistributes the population across the network and smooths heterogeneous behavioral patterns.

4.1. **Mutation Operator and Random-Walk Laplacian.**

Definition 30 (Random-walk matrix and Laplacian)**.** Let

$$A \in \mathbb{R}^{n\times n} \tag{4.1}$$

be a nonnegative weighted adjacency (payoff) matrix associated with a graph $G = (V, E)$. Define the degree matrix

$$K = \operatorname{diag}(A\mathbf{1}), \tag{4.2}$$

and the row-stochastic random-walk matrix

$$P = K^{-1}A. \tag{4.3}$$

The associated random-walk Laplacian is

$$L_{\mathrm{rw}} = I - P. \tag{4.4}$$

The matrix P defines a random walk on the graph, where the transition probability from one vertex to another is proportional to the corresponding payoff weight. In the Olfati–Saber formulation, the mutation or dispersal operator is constructed directly from this random-walk Laplacian.

Definition 31 (Mutation operator)**.** The mutation matrix is defined as

$$Q = I - \mu L_{\mathrm{rw}} = (1-\mu)I + \mu P, \tag{4.5}$$

where $\mu \geq 0$ is the mutation or mixing parameter.

For $0 \leq \mu \leq 1$, the above expression interpolates continuously between the identity operator and the random-walk transition matrix. When $\mu = 0$, there is no mutation, while increasing μ strengthens the mixing induced by the graph topology. For $\mu > 1$, Q remains mass preserving but is no longer a convex combination of I and P.

4.2. **Spectral Relation Between Q and L_{rw}.**

Proposition 32 (Shared eigenvectors)**.** *Let*

$$L_{\mathrm{rw}} v_i = \lambda_i v_i \tag{4.6}$$

be an eigenpair of the random-walk Laplacian. Then v_i is also an eigenvector of the mutation matrix Q, with eigenvalue

$$1 - \mu\lambda_i. \tag{4.7}$$

Proof. Using the definition

$$Q = I - \mu L_{\mathrm{rw}}, \tag{4.8}$$

we obtain

$$Qv_i = (I - \mu L_{\mathrm{rw}})v_i = v_i - \mu\lambda_i v_i = (1 - \mu\lambda_i)v_i. \tag{4.9}$$

Hence v_i is simultaneously an eigenvector of Q and L_{rw}, and the corresponding eigenvalues are related by the affine transformation

$$\lambda_i \mapsto 1 - \mu\lambda_i. \tag{4.10}$$

□

Thus, the mutation matrix and the random-walk Laplacian possess the same eigenvectors. Mutation modifies only the amplification or attenuation associated with each Laplacian mode.

Lemma 33 (Constant mode)**.** *The constant vector* $\mathbf{1}$ *is an eigenvector of* P*,* L_{rw}*, and* Q*.*

Proof. Since P is row stochastic,

$$P\mathbf{1} = \mathbf{1}. \tag{4.11}$$

Therefore,

$$L_{\mathrm{rw}}\mathbf{1} = (I - P)\mathbf{1} = 0, \tag{4.12}$$

which implies

$$Q\mathbf{1} = (I - \mu L_{\mathrm{rw}})\mathbf{1} = \mathbf{1}. \tag{4.13}$$

□

The vector $\mathbf{1}$ is the constant right mode of these row-normalized operators. For the transpose dynamics acting on column population vectors, the zero mode is instead the stationary distribution of P.

Proposition 34 (Spectrum for undirected graphs)**.** *Suppose that* A *is symmetric and all degrees are positive. Then* L_{rw} *is similar to the symmetric normalized Laplacian*

$$L_{\mathrm{sym}} = I - K^{-1/2}AK^{-1/2}, \tag{4.14}$$

and consequently has a real spectrum satisfying

$$0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le 2. \tag{4.15}$$

The eigenvalues of Q *therefore lie in the interval*

$$1 - 2\mu \le \sigma(Q) \le 1. \tag{4.16}$$

Proof. Since A is symmetric and all degrees are positive, the degree matrix

$$K = \mathrm{diag}(A\mathbf{1}) \tag{4.17}$$

is positive diagonal and therefore $K^{1/2}$ and $K^{-1/2}$ are well defined.

The random-walk Laplacian is

$$L_{\mathrm{rw}} = I - K^{-1}A. \tag{4.18}$$

Consider the similarity transformation

$$K^{1/2}L_{\mathrm{rw}}K^{-1/2}. \tag{4.19}$$

Then

$$K^{1/2}L_{\mathrm{rw}}K^{-1/2} = K^{1/2}(I - K^{-1}A)K^{-1/2}. \tag{4.20}$$

Expanding,

$$K^{1/2}L_{\mathrm{rw}}K^{-1/2} = I - K^{-1/2}AK^{-1/2}. \tag{4.21}$$

Hence

$$K^{1/2}L_{\mathrm{rw}}K^{-1/2} = L_{\mathrm{sym}}. \tag{4.22}$$

Therefore L_{rw} is similar to L_{sym}, and both matrices have the same eigenvalues.

Since $A = A^T$, the matrix

$$K^{-1/2}AK^{-1/2} \tag{4.23}$$

is symmetric. Thus

$$L_{\mathrm{sym}} = I - K^{-1/2}AK^{-1/2} \tag{4.24}$$

is also symmetric, and all its eigenvalues are real. Moreover, L_{sym} is the standard normalized Laplacian of an undirected graph, so its spectrum satisfies

$$0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n \le 2. \tag{4.25}$$

Because L_{rw} is similar to L_{sym}, the same spectral bounds hold for L_{rw}.

Finally, the mutation matrix is

$$Q = I - \mu L_{\mathrm{rw}}. \tag{4.26}$$

If λ_i is an eigenvalue of L_{rw}, then the corresponding eigenvalue of Q is

$$1 - \mu\lambda_i. \tag{4.27}$$

Since

$$0 \le \lambda_i \le 2, \tag{4.28}$$

it follows that

$$1 - 2\mu \le 1 - \mu\lambda_i \le 1. \tag{4.29}$$

Therefore,

$$1 - 2\mu \le \sigma(Q) \le 1. \tag{4.30}$$

□

4.3. **Connection with the Replicator–Mutator Dynamics.** The continuous-time replicator–mutator equation can be written compactly as

$$\dot{x} = Q^T\big[(Ax) \odot x\big] - (x^TAx)x, \tag{4.31}$$

where $x \in \mathbb{R}^n$ satisfies

$$x \ge 0, \qquad x^T\mathbf{1} = 1, \tag{4.32}$$

and $\odot$ denotes the Hadamard product.

The nonlinear term

$$(Ax) \odot x \tag{4.33}$$

describes local replication driven by payoffs, while the matrix Q^T redistributes this replication flux according to the graph structure.

Although the full dynamics are nonlinear, the mixing operator Q^T acts linearly on the replication term. Consequently, the Laplacian eigenvectors provide a natural basis for understanding how network structure shapes the propagation and attenuation of evolutionary modes.

4.4. **Role of the Spectral Gap.**

Definition 35 (Spectral gap)**.** Let

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n \tag{4.34}$$

be the eigenvalues of L_{rw}. The quantity

$$\lambda_2 \tag{4.35}$$

is called the spectral gap of the graph.

The spectral gap determines the characteristic time scale over which heterogeneity is smoothed by mutation and dispersal.

Remark 36. Modes associated with large eigenvalues λ_i correspond to rapidly varying, high-frequency patterns on the graph and are strongly attenuated through the factor

$$1 - \mu\lambda_i. \tag{4.36}$$

Conversely, modes associated with small eigenvalues correspond to smooth spatial patterns that persist for longer times in the dynamics.

Proposition 37 (Mixing time scale)**.** *The characteristic mixing time induced by mutation scales as*

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu\lambda_2}. \tag{4.37}$$

Consequently, graphs with small spectral gap exhibit slow mixing and long-lived heterogeneity, whereas graphs with large spectral gap homogenize rapidly.

Proof. Consider first the pure mutation dynamics

$$\dot{y} = -\mu L_{\mathrm{rw}}^T y. \tag{4.38}$$

This linear system isolates the mixing effect of the mutation operator from the nonlinear selection term in the full replicator–mutator dynamics.

Assume that the graph is connected and undirected. Then the eigenvalues of L_{rw} satisfy

$$0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n. \tag{4.39}$$

The eigenvalue $\lambda_1 = 0$ corresponds to the consensus mode, while the remaining eigenvalues correspond to nontrivial heterogeneous modes.

Expanding the initial condition in Laplacian modes, we may write

$$y(0) = c_1 w_1 + \sum_{i=2}^{n} c_i w_i, \tag{4.40}$$

The solution is

$$y(t) = e^{-\mu L_{\mathrm{rw}}^T t} y(0). \tag{4.41}$$

Therefore, each nontrivial mode decays with rate $\mu\lambda_i$, so that

$$y(t) - y_\infty = \sum_{i=2}^{n} c_i e^{-\mu\lambda_i t} w_i, \tag{4.42}$$

where w_i are right eigenvectors of L_{rw}^T and y_∞ denotes the limiting stationary component.

The slowest decaying nontrivial mode is the one associated with the smallest positive eigenvalue λ_2. Hence, for large time,

$$\|y(t) - y_\infty\| \sim Ce^{-\mu\lambda_2 t}, \tag{4.43}$$

for some constant $C > 0$ depending on the initial condition.

The characteristic mixing time is the inverse of the slowest exponential decay rate. Thus,

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu\lambda_2}. \tag{4.44}$$

Consequently, if λ_2 is small, then the decay rate

$$\mu\lambda_2 \tag{4.45}$$

is small and the system mixes slowly, producing long-lived heterogeneous patterns. Conversely, if λ_2 is large, the slowest nontrivial mode is attenuated rapidly, and the graph homogenizes quickly under mutation. □

4.5. **Linearized Dynamics and Mode Decomposition.** Consider a perturbation

$$\delta x \tag{4.46}$$

around an interior equilibrium x^*. Linearizing Eq. (4.31) yields

$$\delta\dot{x} = J(x^*)\delta x, \tag{4.47}$$

where the Jacobian matrix contains contributions involving Q^T, A, and diagonal matrices depending on x^*.

In regimes such as weak selection or equilibria close to the uniform distribution, the transient dynamics can be approximately decomposed into Laplacian eigenmodes. Each mode evolves with an effective growth or decay rate influenced by the quantity

$$\mu\lambda_i. \tag{4.48}$$

Thus, the Laplacian eigenvectors define the natural spatial modes of the evolutionary dynamics, while the eigenvalues determine their characteristic attenuation time scales.

4.6. **Pure Mutation Benchmark.** To isolate the effect of mutation alone, consider the linear system

$$\dot{y} = -\mu L_{\mathrm{rw}}^T y. \tag{4.49}$$

Proposition 38 (Pure mutation dynamics)**.** *The solution of the pure mutation system is*

$$y(t) = \exp(-\mu L_{\mathrm{rw}}^T t)\, y(0). \tag{4.50}$$

Moreover, each Laplacian eigenmode decays independently as

$$\exp(-\mu\lambda_i t). \tag{4.51}$$

Proof. Since the system is linear with constant coefficients, its solution is given by the matrix exponential:

$$y(t) = e^{-\mu L_{\mathrm{rw}}^T t} y(0). \tag{4.52}$$

If w_i is a right eigenvector of L_{rw}^T with eigenvalue λ_i, then

$$e^{-\mu L_{\mathrm{rw}}^T t} w_i = e^{-\mu\lambda_i t} w_i. \tag{4.53}$$

Thus each eigenmode evolves independently with exponential decay rate $\mu\lambda_i$. □

This benchmark highlights the central role of the Laplacian spectrum in determining the characteristic time scales of mutation and mixing. Although the full replicator–mutator dynamics are nonlinear, the same spectral intuition carries qualitatively to the complete system.

4.7. **Canonical Graph Families and Dynamical Implications.** We now summarize the spectral properties of several canonical, loopless graph families and discuss their implications for the replicator–mutator dynamics.

4.7.1. *Cycle Graph C_n.*

Proposition 39 (Spectrum of the cycle graph)**.** *For the unweighted cycle graph C_n,*

$$P = \frac{1}{2}A, \qquad L_{\mathrm{rw}} = I - \frac{1}{2}A. \tag{4.54}$$

The Laplacian eigenvalues are

$$\lambda_k = 1 - \cos\left(\frac{2\pi k}{n}\right), \qquad k = 0, \dots, n-1. \tag{4.55}$$

In particular,

$$\lambda_2 = \Theta(n^{-2}), \tag{4.56}$$

and therefore

$$\tau_{\mathrm{mix}} = \Theta\left(\frac{n^2}{\mu}\right). \tag{4.57}$$

Proof. The cycle graph C_n is 2-regular, hence

$$P = K^{-1}A = \frac{1}{2}A, \tag{4.58}$$

and therefore

$$L_{\mathrm{rw}} = I - P = I - \frac{1}{2}A. \tag{4.59}$$

The adjacency matrix of C_n is circulant, so its eigenvalues are

$$\eta_k = 2\cos\left(\frac{2\pi k}{n}\right), \qquad k = 0, \dots, n-1. \tag{4.60}$$

Hence the eigenvalues of L_{rw} are

$$\lambda_k = 1 - \frac{1}{2}\eta_k = 1 - \cos\left(\frac{2\pi k}{n}\right). \tag{4.61}$$

For the first nonzero eigenvalue,

$$\lambda_2 = 1 - \cos\left(\frac{2\pi}{n}\right) \sim \frac{1}{2}\left(\frac{2\pi}{n}\right)^2 = \frac{2\pi^2}{n^2}, \tag{4.62}$$

so

$$\lambda_2 = \Theta(n^{-2}). \tag{4.63}$$

Using

$$\tau_{\text{mix}} \sim \frac{1}{\mu\lambda_2}, \tag{4.64}$$

we obtain

$$\tau_{\text{mix}} = \Theta\left(\frac{n^2}{\mu}\right). \tag{4.65}$$

□

Hence cycle graphs possess very slow mixing and support long-lived low-frequency spatial modes.

4.7.2. *Path Graph P_n.* For the path graph, the spectral gap also satisfies

$$\lambda_2 = \Theta(n^{-2}), \tag{4.66}$$

implying slow homogenization comparable to that of cycles. The dominant modes correspond to smooth spatial variations constrained by the graph boundaries.

4.7.3. *Star Graph S_n.*

Proposition 40 (Spectrum of the star graph)**.** *For the star graph S_n, the normalized Laplacian spectrum is*

$$\lambda_1 = 0, \qquad \lambda = 1 \quad \textit{with multiplicity } (n-2), \qquad \lambda_n = 2. \tag{4.67}$$

Hence

$$\lambda_2 = 1, \qquad \tau_{\text{mix}} \sim \frac{1}{\mu}. \tag{4.68}$$

Proof. Let the center of the star be vertex 1, and let vertices $2, \ldots, n$ be the leaves. The normalized Laplacian is

$$L_{\text{sym}} = I - K^{-1/2} A K^{-1/2}. \tag{4.69}$$

For the star graph, the center has degree $n-1$, while every leaf has degree 1.

The vector

$$K^{1/2}\mathbf{1} \tag{4.70}$$

is an eigenvector of L_{sym} with eigenvalue 0. Hence

$$\lambda_1 = 0. \tag{4.71}$$

Now consider vectors supported only on the leaves and satisfying

$$\sum_{i=2}^{n} v_i = 0, \qquad v_1 = 0. \tag{4.72}$$

There are $n-2$ linearly independent vectors of this form. For such vectors, the normalized adjacency term vanishes, so

$$L_{\text{sym}} v = v. \tag{4.73}$$

Thus,

$$\lambda = 1 \tag{4.74}$$

has multiplicity $n-2$.

Since the normalized Laplacian of a connected bipartite graph has largest eigenvalue 2, and the star graph is connected and bipartite, we also have

$$\lambda_n = 2. \tag{4.75}$$

Therefore,

$$\sigma(L_{\mathrm{sym}}) = \{0, \underbrace{1, \ldots, 1}_{n-2 \text{ times}}, 2\}. \tag{4.76}$$

Hence

$$\lambda_2 = 1. \tag{4.77}$$

Using the mixing-time estimate

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu \lambda_2}, \tag{4.78}$$

we obtain

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu}. \tag{4.79}$$

Mutation therefore redistributes mass very efficiently through the central hub, producing rapid homogenization. □

4.7.4. *Complete Graph K_n.*

Proposition 41 (Spectrum of the complete graph)**.** *For the complete graph K_n,*

$$P = \frac{1}{n-1} A, \qquad L_{\mathrm{rw}} = I - \frac{1}{n-1} A. \tag{4.80}$$

The Laplacian eigenvalues are

$$\lambda_1 = 0, \qquad \lambda = \frac{n}{n-1} \quad \textit{with multiplicity } (n-1). \tag{4.81}$$

Hence

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu}. \tag{4.82}$$

Proof. The complete graph K_n is $(n-1)$-regular. Hence

$$P = K^{-1} A = \frac{1}{n-1} A, \tag{4.83}$$

and therefore

$$L_{\mathrm{rw}} = I - P = I - \frac{1}{n-1} A. \tag{4.84}$$

The adjacency matrix of K_n has eigenvalues

$$n-1 \tag{4.85}$$

with eigenvector $\mathbf{1}$, and

$$-1 \tag{4.86}$$

with multiplicity $n-1$, corresponding to all vectors orthogonal to $\mathbf{1}$. Therefore, the eigenvalues of L_{rw} are

$$1 - \frac{n-1}{n-1} = 0 \tag{4.87}$$

for the consensus mode, and

$$1 - \frac{-1}{n-1} = \frac{n}{n-1} \tag{4.88}$$

with multiplicity $n-1$.

Thus,

$$\lambda_1 = 0, \qquad \lambda = \frac{n}{n-1} \quad \text{with multiplicity } n-1. \tag{4.89}$$

Since

$$\lambda_2 = \frac{n}{n-1} \sim 1, \tag{4.90}$$

the mixing-time estimate

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu\lambda_2} \tag{4.91}$$

gives

$$\tau_{\mathrm{mix}} \sim \frac{1}{\mu}. \tag{4.92}$$

□

Complete graphs suppress all non-consensus modes almost uniformly, leading to extremely rapid global mixing.

4.8. **Structure–Dynamics Relations.** The previous examples illustrate a general qualitative ordering:

(4.93) paths/cycles (slow mixing) $\ll$ stars and complete graphs (fast mixing).

Thus, the spectral gap provides a principled predictor of how rapidly mutation redistributes evolutionary mass across the graph. In the replicator–mutator dynamics, the nonlinear selection term generates state-dependent forcing in graph space, while the mutation operator filters this forcing through the Laplacian eigenmodes. Consequently, graph eigenvectors determine the natural spatial patterns of the dynamics, whereas Laplacian eigenvalues determine their characteristic attenuation time scales under mutation.

5. Equilibrium Analysis

The main purpose of this section is to relate the local convergence rate of the nonlocal replicator–mutator system to the spectral structure of the nonlocal random-walk Laplacian. We keep the mathematical formulation explicit, but we avoid unnecessary repetitions in the derivations.

Throughout this section we assume that the underlying graph is connected, undirected, and simple.

For a fixed Mellin exponent $s > 0$, define the nonlocal weighted adjacency matrix A_s, which is symmetric and has strictly positive off-diagonal entries. Let

$$K_s = \mathrm{diag}(A_s\vec{1}), \qquad P_s = K_s^{-1}A_s, \qquad L_s = I - P_s. \tag{5.1}$$

Here P_s is row-stochastic and $L_s\vec{1} = \vec{0}$.

The nonlocal replicator–mutator equation is

$$\dot{\vec{x}} = (I - \mu L_s^T)\big((A_s\vec{x}) \odot \vec{x}\big) - (\vec{x}^T A_s \vec{x})\vec{x}, \qquad \vec{x} \in \Delta^{n-1}, \tag{5.2}$$

where

$$\Delta^{n-1} = \left\{ \vec{x} \in \mathbb{R}^n : x_i \geq 0, \sum_{i=1}^{n} x_i = 1 \right\}. \tag{5.3}$$

Lemma 42. *Let G be connected and undirected, and let A_s is nonlocal weighted adjacency matrix. Then the random-walk Laplacian $L_s = I - K_s^{-1}A_s$ is diagonalizable over $\mathbb{R}$. Moreover, its spectrum is real and nonnegative.*

Proof. Since $A_s = A_s^T$ and K_s is positive diagonal, L_s is similar to the symmetric normalized Laplacian

$$L_s^{\mathrm{sym}} = I - K_s^{-1/2} A_s K_s^{-1/2}. \tag{5.4}$$

Indeed,

$$L_s = K_s^{-1/2} L_s^{\mathrm{sym}} K_s^{1/2}. \tag{5.5}$$

The matrix L_s^{sym} is real symmetric and positive semidefinite. Therefore it is diagonalizable over $\mathbb{R}$ with nonnegative real eigenvalues. Since L_s is similar to it, the same spectral conclusions hold for L_s. □

5.1. **Interior equilibria.**

Definition 43. A point $\vec{x}^* \in \Delta^{n-1}$ is an interior equilibrium of (5.2) if $x_i^* > 0$ for every i and

$$(I - \mu L_s^T)\big((A_s\vec{x}^*) \odot \vec{x}^*\big) - (\vec{x}^{*T} A_s \vec{x}^*)\vec{x}^* = \vec{0}. \tag{5.6}$$

Let

$$c = \vec{x}^{*T} A_s \vec{x}^*. \tag{5.7}$$

Then the equilibrium condition can be written as

$$(I - \mu L_s^T)\big((A_s\vec{x}^*) \odot \vec{x}^*\big) = c\vec{x}^*. \tag{5.8}$$

Proposition 44. *For the pure replicator system, obtained from* (5.2) *by setting $\mu = 0$, an interior state $\vec{x}^*$ is an equilibrium if and only if*

$$A_s\vec{x}^* = c\vec{1}, \qquad c = \vec{x}^{*T} A_s \vec{x}^*. \tag{5.9}$$

Proof. For $\mu = 0$, the equilibrium condition is

$$(A_s\vec{x}^*) \odot \vec{x}^* = c\vec{x}^*. \tag{5.10}$$

Since $x_i^* > 0$ for all i, division by x_i^* gives $(A_s\vec{x}^*)_i = c$ for every i. This is exactly $A_s\vec{x}^* = c\vec{1}$. The converse follows immediately by multiplying $A_s\vec{x}^* = c\vec{1}$ componentwise by $\vec{x}^*$. □

Proposition 45. *Let $\mu > 0$. Suppose that an interior state $\vec{x}^*$ satisfies*

$$A_s\vec{x}^* = c\vec{1}, \qquad L_s^T\vec{x}^* = \vec{0}, \tag{5.11}$$

where $c = \vec{x}^{*T} A_s \vec{x}^* > 0$. Then $\vec{x}^*$ is an interior equilibrium of the nonlocal replicator–mutator system (5.2).

Proof. The first condition gives

$$(A_s\vec{x}^*) \odot \vec{x}^* = c\vec{x}^*. \tag{5.12}$$

Therefore

$$(I - \mu L_s^T)\big((A_s\vec{x}^*) \odot \vec{x}^*\big) = (I - \mu L_s^T)c\vec{x}^* = c\vec{x}^*, \tag{5.13}$$

where the last equality uses $L_s^T\vec{x}^* = \vec{0}$. Hence (5.8) holds. □

Remark 46. The two relations in (5.11) are not equivalent in general. The condition $A_s\vec{x}^* = c\vec{1}$ is the uniform-fitness condition familiar from the pure replicator equation. The condition $L_s^T\vec{x}^* = \vec{0}$ says that $\vec{x}^*$ is the stationary distribution of the random walk generated by P_s. For symmetric A_s, this stationary distribution is proportional to the weighted degree vector $A_s\vec{1}$. Thus both conditions hold simultaneously only for a special class of graphs or kernels. In the stability analysis below, we restrict attention to this class of equilibria.

5.2. **Linearization around an interior equilibrium.** Let $\vec{x}^*$ be an interior equilibrium satisfying (5.11). Consider a perturbation

$$\vec{x}(t) = \vec{x}^* + \delta\vec{x}(t), \qquad \vec{1}^T\delta\vec{x}(t) = 0. \tag{5.14}$$

Thus perturbations belong to the tangent space

$$T\Delta^{n-1} = \{\vec{v} \in \mathbb{R}^n : \vec{1}^T\vec{v} = 0\}. \tag{5.15}$$

Let

$$D_* = \mathrm{diag}(\vec{x}^*). \tag{5.16}$$

Lemma 47. *At an interior equilibrium satisfying $A_s\vec{x}^* = c\vec{1}$, the pure replicator term satisfies*

$$(A_s\vec{x}) \odot \vec{x} - (\vec{x}^T A_s\vec{x})\vec{x} = D_*A_s\delta\vec{x} + O(\|\delta\vec{x}\|^2). \tag{5.17}$$

Proof. Since $A_s\vec{x}^* = c\vec{1}$,

$$A_s\vec{x} = c\vec{1} + A_s\delta\vec{x}. \tag{5.18}$$

Therefore

$$(A_s\vec{x}) \odot \vec{x} = c\vec{x}^* + c\delta\vec{x} + D_*A_s\delta\vec{x} + O(\|\delta\vec{x}\|^2). \tag{5.19}$$

Moreover, using the symmetry of A_s and the tangent-space condition,

$$\vec{x}^{*T}A_s\delta\vec{x} = (A_s\vec{x}^*)^T\delta\vec{x} = c\vec{1}^T\delta\vec{x} = 0. \tag{5.20}$$

Thus

$$\vec{x}^TA_s\vec{x} = c + O(\|\delta\vec{x}\|^2), \tag{5.21}$$

and consequently

$$(\vec{x}^TA_s\vec{x})\vec{x} = c\vec{x}^* + c\delta\vec{x} + O(\|\delta\vec{x}\|^2). \tag{5.22}$$

Subtracting the last expression from the expansion of $(A_s\vec{x}) \odot \vec{x}$ gives (5.17). □

Lemma 48. *If, in addition, $L_s^T\vec{x}^* = \vec{0}$, then*

$$\mu L_s^T\big((A_s\vec{x}) \odot \vec{x}\big) = \mu L_s^T(cI + D_*A_s)\delta\vec{x} + O(\mu\|\delta\vec{x}\|^2). \tag{5.23}$$

Proof. From the expansion in the proof of Lemma 47,

$$(A_s\vec{x})\odot\vec{x} = c\vec{x}^* + (cI + D_*A_s)\delta\vec{x} + O(\|\delta\vec{x}\|^2). \tag{5.24}$$

Applying μL_s^T and using $L_s^T\vec{x}^* = \vec{0}$ gives the result. □

Proposition 49. *The linearization of* (5.2) *at an interior equilibrium satisfying* (5.11) *is*

$$\delta\dot{\vec{x}} = J_{\text{eff}}\delta\vec{x} + O(\|\delta\vec{x}\|^2), \tag{5.25}$$

where

$$J_{\text{eff}} = J_0 - \mu L_s^T B, \qquad J_0 = D_*A_s, \qquad B = cI + D_*A_s. \tag{5.26}$$

Proof. Combine Lemmas 47 and 48 in the decomposition

$$\dot{\vec{x}} = \big((A_s\vec{x})\odot\vec{x} - (\vec{x}^TA_s\vec{x})\vec{x}\big) - \mu L_s^T\big((A_s\vec{x})\odot\vec{x}\big). \tag{5.27}$$

□

5.3. **Spectral structure of the pure replicator Jacobian.**

Lemma 50. *The matrix $J_0 = D_*A_s$ is similar to the real symmetric matrix*

$$S_0 = D_*^{1/2}WD_*^{1/2}. \tag{5.28}$$

Consequently, J_0 is diagonalizable and has real eigenvalues.

Proof. Since D_* is positive diagonal and $A_s = A_s^T$, the matrix S_0 is real symmetric. Moreover,

$$J_0 = D_*^{1/2}S_0D_*^{-1/2}. \tag{5.29}$$

Thus J_0 is similar to a real symmetric matrix. □

Let

$$J_0\vec{u}_k = \alpha_k\vec{u}_k, \qquad \vec{v}_k^TJ_0 = \alpha_k\vec{v}_k^T, \qquad \vec{v}_k^T\vec{u}_\ell = \delta_{k\ell}. \tag{5.30}$$

If $\vec{w}_k$ is an orthonormal eigenvector of S_0, then

$$\vec{u}_k = D_*^{1/2}\vec{w}_k, \qquad \vec{v}_k = D_*^{-1/2}\vec{w}_k. \tag{5.31}$$

Remark 51. Because the dynamics preserve the simplex, local stability is determined by the spectrum of the Jacobian restricted to $T\Delta^{n-1}$. In particular, the relevant dominant eigenvalue is the one with largest real part among tangent perturbations. Since J_0 is similar to a symmetric matrix, the eigenvalues of J_0 are real. We denote the dominant tangent eigenvalue by α_1.

5.4. **First-order perturbation by mutation.** We now assume that α_1 is a simple eigenvalue of J_0 restricted to $T\Delta^{n-1}$. This assumption avoids degeneracies and allows the use of standard first-order perturbation theory.

Proposition 52. *Let α_1 be a simple dominant eigenvalue of J_0 on the tangent space, with corresponding right and left eigenvectors $\vec{u}_1$ and $\vec{v}_1$ normalized by $\vec{v}_1^T\vec{u}_1 = 1$. Then the corresponding eigenvalue of J_{eff} satisfies*

$$\widetilde{\lambda}_1 = \alpha_1 - \mu(\alpha_1 + c)\mathcal{R}_1 + O(\mu^2), \tag{5.32}$$

where

$$\mathcal{R}_1 = \vec{v}_1^TL_s^T\vec{u}_1. \tag{5.33}$$

Proof. From Proposition 49,

$$J_{\text{eff}} = J_0 + \mu E, \qquad E = -L_s^T B. \tag{5.34}$$

First-order perturbation theory for a simple eigenvalue gives

$$\widetilde{\lambda}_1 = \alpha_1 + \mu \vec{v}_1^T E \vec{u}_1 + O(\mu^2). \tag{5.35}$$

Since $B = cI + J_0$, we have

$$B\vec{u}_1 = (c + \alpha_1)\vec{u}_1. \tag{5.36}$$

Therefore

$$\vec{v}_1^T E \vec{u}_1 = -\vec{v}_1^T L_s^T B \vec{u}_1 = -(c + \alpha_1)\vec{v}_1^T L_s^T \vec{u}_1, \tag{5.37}$$

which proves (5.32). □

5.5. **Convergence time.**

Definition 53. For a linearly stable equilibrium $\vec{x}^*$, the ε-convergence time is

$$\tau_\varepsilon = \inf\{t > 0 : \|\vec{x}(t) - \vec{x}^*\| < \varepsilon\}. \tag{5.38}$$

Proposition 54. *Assume that the dominant eigenvalue $\widetilde{\lambda}_1$ of J_{eff} is real, simple, and negative. Then, for generic initial perturbations with nonzero projection onto the corresponding eigenmode,*

$$\tau_\varepsilon = \frac{1}{|\widetilde{\lambda}_1|} \log \frac{1}{\varepsilon} + O(1), \qquad \varepsilon \to 0. \tag{5.39}$$

Using (5.32)*, this becomes, to first order in μ,*

$$\tau_\varepsilon = \frac{1}{\mu(\alpha_1 + c)\mathcal{R}_1 - \alpha_1} \log \frac{1}{\varepsilon} + O(1), \qquad \varepsilon \to 0, \tag{5.40}$$

provided the denominator is positive.

Proof. For the linearized system,

$$\delta\vec{x}(t) = e^{J_{\text{eff}} t}\delta\vec{x}(0). \tag{5.41}$$

If $\widetilde{\lambda}_1$ is the dominant stable eigenvalue and the initial perturbation has nonzero projection onto its eigenvector, then

$$\|\delta\vec{x}(t)\| \sim Ce^{\widetilde{\lambda}_1 t} = Ce^{-|\widetilde{\lambda}_1|t}, \tag{5.42}$$

for some $C > 0$. Solving $Ce^{-|\widetilde{\lambda}_1|\tau_\varepsilon} = \varepsilon$ gives (5.39). Substituting the first-order expansion (5.32) gives (5.40). □

Theorem 55. *Let G be a connected undirected graph, and let A_s be the nonlocal Mellin weighted adjacency matrix. Suppose that there exists an interior equilibrium $\vec{x}^*$ satisfying*

$$A_s\vec{x}^* = c\vec{1}, \qquad L_s^T\vec{x}^* = \vec{0}, \qquad c = \vec{x}^{*T} A_s \vec{x}^* > 0. \tag{5.43}$$

Let $D_* = \operatorname{diag}(\vec{x}^*)$ and $J_0 = D_* A_s$. Assume that the dominant eigenvalue α_1 of J_0 on $T\Delta^{n-1}$ is simple, with right and left eigenvectors $\vec{u}_1$ and $\vec{v}_1$ normalized by $\vec{v}_1^T\vec{u}_1 = 1$. Then, to first order in μ,

$$\widetilde{\lambda}_1 = \alpha_1 - \mu(\alpha_1 + c)\mathcal{R}_1 + O(\mu^2), \qquad \mathcal{R}_1 = \vec{v}_1^T L_s^T \vec{u}_1. \tag{5.44}$$

The equilibrium is linearly stable to first order whenever

$$\mu(\alpha_1 + c)\mathcal{R}_1 > \alpha_1. \tag{5.45}$$

In that case,

$$\tau_\varepsilon = \frac{1}{\mu(\alpha_1 + c)\mathcal{R}_1 - \alpha_1} \log \frac{1}{\varepsilon} + O(1), \qquad \varepsilon \to 0. \tag{5.46}$$

Proof. The eigenvalue formula follows from Proposition 52. The stability condition is the requirement that the dominant eigenvalue be negative to first order in μ. The convergence-time estimate follows from Proposition 54. □

Remark 56. The quantity $\mathcal{R}_1 = \vec{v}_1^T L_s^T \vec{u}_1$ measures how strongly the dominant selection mode is affected by nonlocal mutation. Thus the convergence rate depends not only on the spectral gap of L_s, but also on the alignment between the leading mode of the replicator Jacobian and the nonlocal random-walk Laplacian. A direct bound purely in terms of the ordinary spectral gap of L_s requires additional assumptions, such as compatibility between the equilibrium distribution and the random-walk stationary measure.

6. Computational Results

In this section, we present a series of numerical experiments designed to investigate the impact of long-range interactions arising from observational learning. In particular, we analyze how these nonlocal interactions promote globalization in the system. Recall that globalization corresponds to the regime in which no single behavior dominates and the effective diversity approaches $n_e \approx n$.

For a given network, we compute the asymptotic state of the generalized replicator–mutator dynamics and evaluate the corresponding effective diversity

$$n_e = \left(\sum_{i=1}^{n} x_i^2 \right)^{-1}. \tag{6.1}$$

The quantity n_e is then used to characterize the evolutionary phases reached at sufficiently long times (ideally $t \to \infty$).

To examine the role of long-range interactions, we consider several classes of networks. For each network, the simulations are performed for different values of the mutation parameter μ and the Mellin exponent s. This allows us to compute the diversity landscape $n_e = n_e(\mu, s)$.

6.1. Watts–Strogatz Networks (WS).

The Watts–Strogatz model [39] introduces a rewiring probability p, which randomly rewires the edges of an initial ring lattice. Figure 6.1 illustrates the relationship between the mutation rate μ and the effective diversity n_e for different values of the Mellin exponent s.

According to the theoretical analysis developed in the previous sections, smaller values of s increase the contribution of the long-range terms in the transformed d-path random-walk Laplacian,

$$L_\star = I - K_\star^{-1} A_\star, \tag{6.2}$$

thereby strengthening long-range interactions. As a consequence, the system requires smaller mutation rates μ to achieve large diversity values, eventually reaching the globalization regime. In contrast, larger values of s suppress long-range interactions. For sufficiently large s, observational learning becomes negligible and the system effectively reduces to the original local replicator–mutator model.

Figure 6.1 shows the diversity–mutation curves for different values of the Mellin exponent s. In particular, large values of s (e.g., $s = 10$ or $s = 20$) closely reproduce

the behavior of the original replicator–mutator dynamics, in agreement with the theoretical predictions.

In these simulations, we consider networks with $n = 50$ nodes, where each node is initially connected to $k = 10$ nearest neighbors. The reward parameter is fixed at $a = 0.3$.

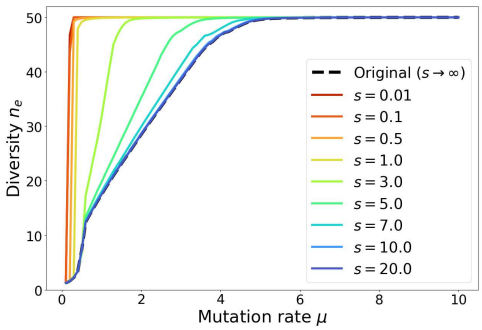

FIGURE 6.1. Diversity–mutation curves corresponding to different values of the Mellin exponent s on a Watts–Strogatz network. The remaining parameters are $n = 50$, $k = 10$, $a = 0.3$, and $p = 0.1$.

To analyze these results in greater detail, we next consider the contour plot shown in Figure 6.2, where the mutation rate μ and the Mellin exponent s are plotted against the resulting diversity n_e. Subfigure 6.2(a), corresponding to rewiring probability $p = 0.1$, clearly illustrates the effect of long-range interactions. Smaller values of s enhance observational learning and therefore reduce the mutation rate required to reach large diversity values and globalization.

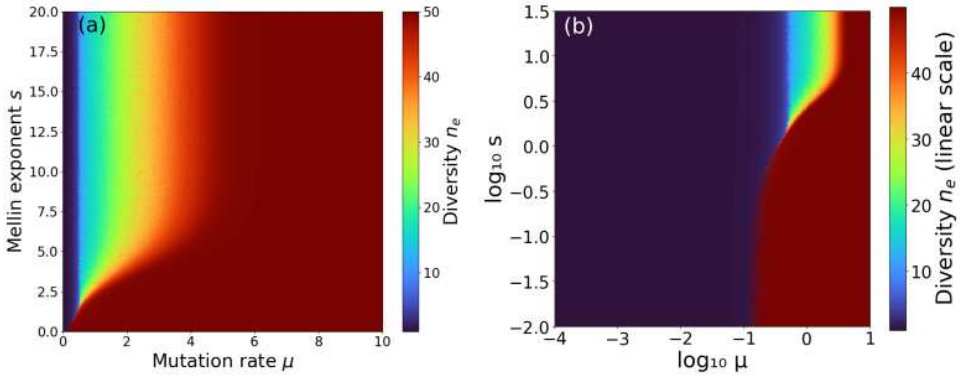

FIGURE 6.2. Contour plots for the Watts–Strogatz network, where the mutation rate μ and the Mellin exponent s are plotted against the effective diversity n_e. Subfigure (a) is shown on a linear scale, whereas subfigure (b) is shown on a logarithmic scale. Both correspond to rewiring probability $p = 0.1$. The common simulation parameters are $n = 50$, $k = 10$, and $a = 0.3$.

To investigate the regime of small Mellin exponents more carefully, we also consider logarithmic scaling for both μ and s, while retaining a linear scale for the

diversity n_e in order to clearly distinguish the different evolutionary phases. Subfigure 6.2(b) corresponds to the logarithmic representation of subfigure 6.2(a).

One particularly interesting feature emerges in the regime of very small Mellin exponents. Once the mutation rate exceeds a small threshold, the system undergoes an abrupt transition in diversity n_e. This behavior reveals a sharp crossover, consistent with a discontinuous transition, from behavioral flocking ($n_e \approx 1$) directly to complete collapse ($n_e \approx n$). In other words, when long-range interactions are sufficiently strong, even a very small mutation rate is capable of driving the system from an ordered flocking state to maximal diversity. We refer to this phenomenon as a *nonlocality-induced collapse*.

6.2. **First-order dynamical phase transition.** Figure 6.2 indicates a sharp transition from behavioural flocking to complete collapse, and hence the occurrence of a *nonlocality-induced collapse* phenomenon. However, a sharp crossover by itself is not sufficient to identify the transition as a first-order phase transition. Therefore, we perform a systematic three-step numerical verification: (i) system-size dependence, (ii) hysteresis under forward and backward mutation-rate sweeps, and (iii) initial-condition dependence and quantitative bistability.

Here we discuss these analyses briefly; the detailed calculations and explanations are provided in Section S2 of the Supplementary Information.

6.2.1. *System-size analysis.* In the system-size analysis, we investigate whether the width of the transition decreases, or equivalently whether the maximum slope increases, as the number of behaviours/nodes is increased. This provides a finite-size analogue of asking whether the observed discontinuous response persists and sharpens with increasing system size.

As the diversity n_e itself depends on n, to compare different system sizes, we first normalize the diversity by defining

$$\eta = \frac{n_e - 1}{n - 1}. \tag{6.3}$$

Then,

$$n_e = 1 \Longrightarrow \eta = 0, \tag{6.4}$$

$$n_e = n \Longrightarrow \eta = 1. \tag{6.5}$$

Consequently,

$$\eta \approx 0 \tag{6.6}$$

represents behavioural flocking, while

$$\eta \approx 1 \tag{6.7}$$

represents complete collapse.

To perform the numerical experiment, we consider a Watts–Strogatz network with the fixed common parameters ($k = 10$, $p = 0.1$, $a = 0.3$, and $s = 10^{-2}$). We also consider one fixed Dirichlet realization for each system size, reused across the entire μ scan for that system size. Here, the initial states are sampled from the symmetric Dirichlet distribution,

$$\mathbf{x}_0 \sim \text{Dirichlet}(1, \ldots, 1), \tag{6.8}$$

which is uniform over the simplex.

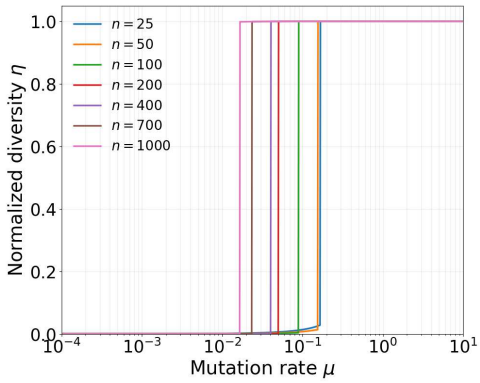

FIGURE 6.3. System-size analysis for the WS network (`seed=123`) with $k = 10$, $a = 0.3$, $p = 0.1$, and $s = 10^{-2}$ under fixed initial conditions.

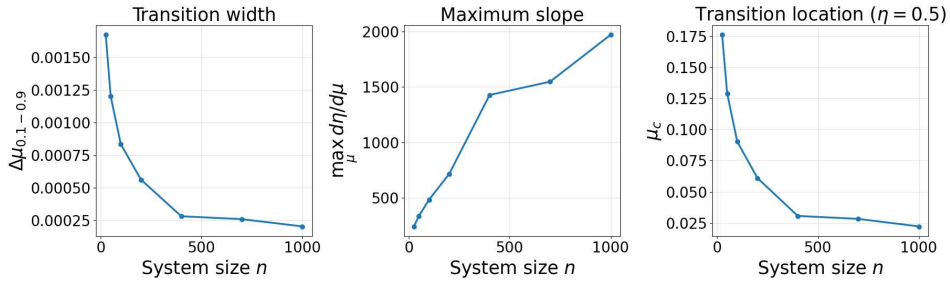

FIGURE 6.4. Quantification of the system-size sharpening for the WS network with $k = 10$, $a = 0.3$, $p = 0.1$, and $s = 10^{-2}$, under fixed initial conditions.

For a sufficiently small value of the Mellin exponent s, namely $s = 10^{-2}$, we compute $\eta(\mu)$ for

$$n = 25,\ 50,\ 100,\ 200,\ 400,\ 700,\ 1000. \tag{6.9}$$

The outcome is shown in figure 6.3, and we observe that the response approaches a step-like transition between

$$\eta \approx 0 \tag{6.10}$$

and

$$\eta \approx 1. \tag{6.11}$$

To quantify the system-size sharpening, we define the transition width as

$$\Delta\mu_{0.1-0.9} = \mu(\eta = 0.9) - \mu(\eta = 0.1), \tag{6.12}$$

which measures the interval over which the normalized diversity changes from 0.1 to 0.9. Second, we calculate the maximum slope

$$\max_{\mu} \frac{d\eta}{d\mu} \tag{6.13}$$

as a measure of the steepness of the response. Finally, we define an operational transition location μ_c by

$$\eta(\mu_c) = 0.5. \tag{6.14}$$

The Fig. 6.4 demonstrates that the transition width decreases systematically with increasing system size n, and equivalently, the maximum slope increases. Thus, the diversity response becomes progressively sharper as the network size increases. We also performed similar experiments for different sufficiently small values of s, and many realizations of different networks, as explained in the Supplementary Information. The observed trend is robust, and these results provide evidence of finite-size sharpening of the transition.

6.2.2. *Hysteresis analysis.* Here we explore whether, at the same value of μ, the state obtained by increasing μ differs from the state obtained by decreasing μ. If, over a finite interval of μ, one observes

$$\eta^{\uparrow}(\mu) \neq \eta^{\downarrow}(\mu), \tag{6.15}$$

this indicates hysteresis. Here, $\eta^{\uparrow}(\mu)$ denotes the response obtained while increasing μ, whereas $\eta^{\downarrow}(\mu)$ denotes the response obtained while decreasing μ.

We define the hysteresis width by

$$\Delta\mu_{\text{hyst}} = \mu^{\uparrow} - \mu^{\downarrow}, \tag{6.16}$$

where $\mu^{\uparrow}$ and $\mu^{\downarrow}$ denote the transition points obtained from the forward and backward sweeps, respectively.

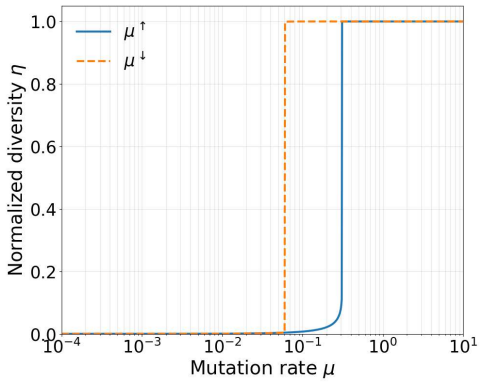

FIGURE 6.5. Hysteresis analysis for the WS network (`seed=123`) with $k = 10$, $a = 0.3$, $p = 0.1$, and $s = 10^{-2}$.

For numerical investigation, we consider a Watts–Strogatz network with the fixed common parameters ($k = 10$, $p = 0.1$, $a = 0.3$, and $s = 10^{-2}$). We visualize the transition in figure 6.5, and after quantifying the simulation results, we obtain

$$\mu^{\uparrow} \approx 0.309516, \qquad \mu^{\downarrow} \approx 0.059604. \tag{6.17}$$

Consequently,

$$\Delta\mu_{\text{hyst}} \approx 0.249912. \tag{6.18}$$

Note that he transition points $\mu^{\uparrow}$ and $\mu^{\downarrow}$ are determined from the first crossing of the forward and backward $\eta(\mu)$ curves, respectively, with the threshold $\eta = 0.5$;

the corresponding μ values are obtained by linear interpolation between the two neighboring points of the continuation grid.

The persistence of a finite separation,

$$\mu^{\downarrow} < \mu^{\uparrow}, \tag{6.19}$$

under refinement of the continuation grid is the central observation. We performed similar experiments for different sufficiently small values of s, and many realizations of different networks, as explained in the Supplementary Information. The behaviour is similar across the tested cases, and these observations provide numerical evidence for the existence of hysteresis in the system.

6.2.3. *Bistability analysis.* The main question we address is: for exactly the same (n, s, μ) and the same network realization, can different initial conditions converge to two distinct terminal states?

In our case, the candidate terminal states are (a) $\eta_{\text{low}} \approx 0$, corresponding to behavioural flocking, and (b) $\eta_{\text{high}} \approx 1$, corresponding to complete collapse.

Thus, the direct bistability test asks whether, for exactly the same (n, s, μ), two different initial conditions can satisfy

$$\mathbf{x}_0^{(1)} \longrightarrow \eta_{\text{low}}, \tag{6.20}$$

while

$$\mathbf{x}_0^{(2)} \longrightarrow \eta_{\text{high}}. \tag{6.21}$$

Here, $\mathbf{x}_0^{(1)}$ and $\mathbf{x}_0^{(2)}$ denote two distinct initial conditions that converge to the low- and high-diversity terminal states, respectively.

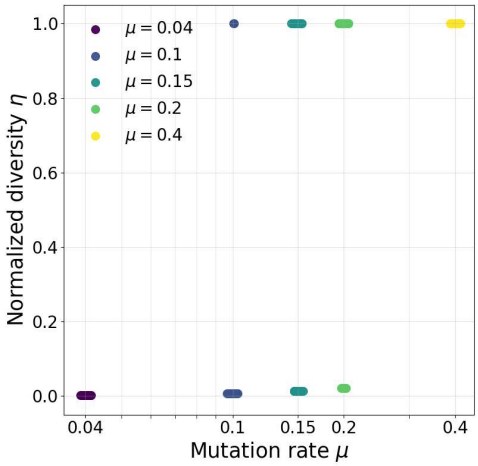

FIGURE 6.6. Bistability analysis for the WS network (`seed=123`) with $k = 10$, $a = 0.3$, $p = 0.1$, and $s = 10^{-2}$.

For the numerical experiments, we consider a Watts–Strogatz network with the fixed common parameters ($k = 10$, $p = 0.1$, $a = 0.3$, and $s = 10^{-2}$).

The tested mutation rates are

$$\mu = 0.04,\ 0.10,\ 0.15,\ 0.20,\ 0.40. \tag{6.22}$$

For every fixed μ, 100 independent Dirichlet initial conditions are generated:

$$\mathbf{x}_0^{(r)} \sim \text{Dirichlet}(1,\dots,1), \qquad r = 1,\dots,100. \tag{6.23}$$

We impose a criterion by defining well-separated thresholds:

$$\begin{cases} \eta < 0.1, & \text{low-diversity branch,} \\ \eta > 0.9, & \text{high-diversity branch,} \\ 0.1 \le \eta \le 0.9, & \text{intermediate state.} \end{cases} \tag{6.24}$$

Thus, a trajectory is counted as belonging to a clearly identified branch only when its terminal diversity is sufficiently far from the middle of the interval.

All trajectories were integrated until the terminal convergence criterion

$$\|\dot{\mathbf{x}}(t_f)\|_2 < 10^{-6} \tag{6.25}$$

was satisfied, with a maximum integration time of $t_{\max} = 5000$.

The figure 6.6 shows the distribution of the final normalized diversity for the different values of the mutation rate μ. The terminal states can split into distinct branches for suitable values of the mutation rate μ. Each trajectory corresponds to a different initial condition, while the colour identifies the mutation rate associated with the corresponding set of initial conditions.

The quantification from the simulation, using the above criterion, shows that the number of trajectories classified as intermediate is zero for every tested mutation rate:

$$M_{\text{intermediate}} = 0. \tag{6.26}$$

Furthermore,

$$M_{\text{converged}} = M_{\text{low}} + M_{\text{high}} = 100 \tag{6.27}$$

for every tested μ. Here, M_{low} corresponds to the low-diversity branch, and M_{high} corresponds to the high-diversity branch.

When we consider mutation rates $\mu = 0.04$ and $\mu = 0.40$, which are outside the hysteresis region (recall that for $s = 10^{-2}$, the hysteresis analysis gave $\mu^{\uparrow} \approx 0.309516$ and $\mu^{\downarrow} \approx 0.059604$), the initial conditions select only one of the two observed branches.

For instance, at $\mu = 0.04$, all tested initial conditions converge to $\eta \approx 0$. Thus, the low-diversity branch is selected for all tested initial conditions at this parameter value. At $\mu = 0.40$, all tested initial conditions converge to $\eta \approx 1$, i.e., the high-diversity branch is selected for all tested initial conditions at this parameter value.

For mutation rates inside the hysteresis region, the initial conditions can split into two distinct branches. The clearest coexistence point is $\mu = 0.15$. At this parameter value,

$$M_{\text{low}} = 44, \qquad M_{\text{high}} = 56. \tag{6.28}$$

We also obtain the separation between the two branch means,

$$\langle \eta \rangle_{\text{high}} - \langle \eta \rangle_{\text{low}} = 0.999980 - 0.012851 \approx 0.987129, \tag{6.29}$$

which is much larger than either within-branch standard deviation. This strong separation indicates that the two populations correspond to distinct terminal states rather than a single broad distribution of slowly relaxing trajectories.

The transition in the basin fractions is also visible across the tested parameter range. At $\mu = 0.10$, the low-diversity branch contains $M_{\text{low}} = 98$ of the 100 initial conditions, whereas at $\mu = 0.20$ it contains only $M_{\text{low}} = 5$. At $\mu = 0.15$, both branches are substantially populated. This behaviour is consistent with the bistable region identified from the hysteresis analysis.

We performed similar experiments for different sufficiently small values of s and many realizations of different networks, and the behaviour is robust, as explained in the Supplementary Information.

6.2.4. *Conclusion.* For sufficiently strong nonlocality, the system undergoes an abrupt transition between a low-diversity state characterized by behavioural flocking and a high-diversity state characterized by an almost homogeneous distribution over the available behaviours, referred to as complete collapse. The transition becomes progressively sharper with increasing system size. Forward and backward continuation in the mutation parameter produces a finite hysteresis interval, while independent initial conditions at identical parameter values converge to either the low- or high-diversity branch. Taken together, these observations provide strong numerical evidence for a first-order dynamical phase transition in individual network realizations. As discussed in the Supplementary Information, ensemble averaging rounds the transition because the transition location and the basins of attraction can vary across network realizations.

6.3. **Erdős–Rényi Networks (ER).** We now consider Erdős–Rényi random graphs [7] and examine the influence of long-range interactions in this class of networks. In the Erdős–Rényi model $G(n, p)$, edges are independently added between pairs of nodes with probability p. Here we denote this probability by p_{edge}, in order to distinguish it from the rewiring probability used in the Watts–Strogatz model.

The ER networks display behavior qualitatively similar to that observed for the Watts–Strogatz networks. In particular, smaller values of the Mellin exponent s strengthen long-range interactions and therefore reduce the mutation rate required to achieve large diversity values (see Figure 6.7).

Moreover, the nonlocality-induced collapse phenomenon is again clearly visible, particularly in the logarithmic representation shown in subfigure 6.7(b). This transition is likewise associated with a first-order dynamical phase transition, as discussed in the Supplementary Information. In these simulations, we use $n = 50$, $a = 0.3$, and an edge probability of $p_{\text{edge}} = 0.118$.

6.4. **Barabási–Albert Networks (BA).** We next investigate the influence of long-range interactions on Barabási–Albert (BA) networks [3]. The BA model generates a growing network in which new nodes preferentially attach to highly connected nodes, thereby producing hubs and a power-law degree distribution. Here, n denotes the number of vertices and m is the number of edges added by each new node when it joins the network (attachment parameter).

The BA networks exhibit behavior qualitatively similar to that observed in the Watts–Strogatz and Erdős–Rényi models. In particular, the nonlocality-induced collapse phenomenon and corresponding first-order dynamical phase transition again appears clearly in subfigure 6.8(b). Subfigure 6.8(a) is shown on a linear scale, whereas subfigure 6.8(b) uses logarithmic scaling. The simulations correspond to parameter $m = 3$, while the remaining parameters are fixed at $n = 50$ and $a = 0.3$.

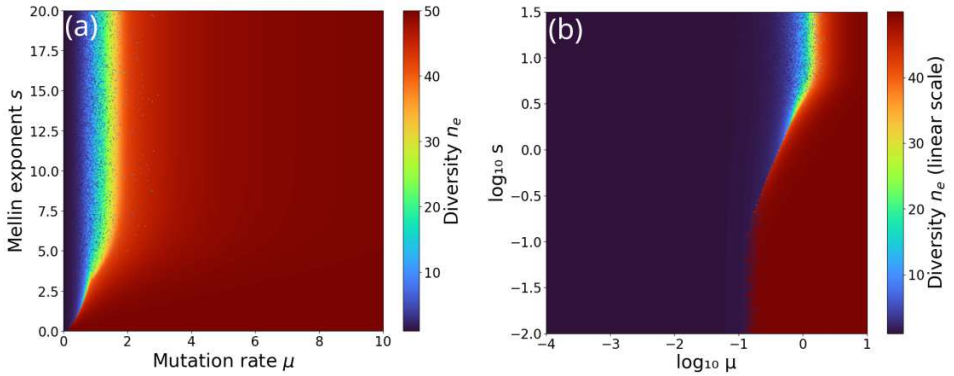

FIGURE 6.7. Contour plots for the Erdős–Rényi network, where the mutation rate μ and Mellin exponent s are plotted against the effective diversity n_e. Subfigure (a) is shown on a linear scale, whereas subfigure (b) is shown on a logarithmic scale. The common simulation parameters are $n = 50$, $a = 0.3$, and edge probability $p_{\text{edge}} = 0.118$.

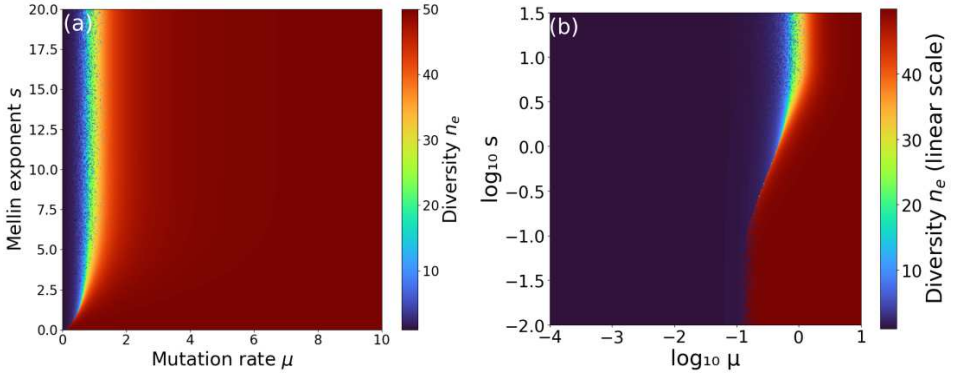

FIGURE 6.8. Contour plots for the Barabási–Albert network, where the mutation rate μ and Mellin exponent s are plotted against the effective diversity n_e. Subfigure (a) is shown on a linear scale, whereas subfigure (b) is shown on a logarithmic scale. The common simulation parameters are $n = 50$, $a = 0.3$, and $m = 3$.

The Erdős–Rényi (ER) network is a classical random graph model, where the edge connection probability p controls the network structure, ranging from an isolated network ($p = 0$) to a complete graph ($p = 1$). In contrast, the Barabási–Albert (BA) network is a classical scale-free network generated through preferential attachment, leading to a power-law degree distribution and the emergence of highly connected hub nodes. To facilitate a fair comparison between these two network topologies, we considered networks with the same number of nodes n and approximately the same average degree $\langle k \rangle$, thereby ensuring a comparable network density. Comparing Figs. 6.7 and 6.8, we observe that the phase transition occurs

more rapidly in the BA network than in the ER network. Specifically, for the same parameter space (μ, s), the region corresponding to complete collapse is noticeably larger in the BA network than in the ER network, suggesting that the generalized RM dynamics is more sensitive to nonlocal interactions in scale-free networks than in random networks.

Note that the numerical methods and computational details are discussed in Section S1 of the Supplementary Information.

6.5. **General Remarks on Artificial Networks.** The numerical results obtained for the Watts–Strogatz, Erdős–Rényi, and Barabási–Albert networks reveal a remarkably robust qualitative behavior under the introduction of long-range interactions. Despite the substantial structural differences among these network families, all three models exhibit the same fundamental phenomenology.

First, decreasing the Mellin exponent s systematically enhances the strength of nonlocal interactions. As predicted by the theoretical analysis, this amplification of long-range observational learning reduces the mutation rate μ required to achieve large diversity values n_e. In all network classes studied here, sufficiently strong nonlocality drives the system toward the globalization regime, characterized by $n_e \approx n$.

Second, the numerical simulations consistently show that for large values of s, the dynamics converge to those of the original local replicator–mutator model. This confirms that the Mellin-transformed interaction operator provides a continuous interpolation between purely local and strongly nonlocal evolutionary dynamics.

A particularly important result emerging across all artificial networks is the appearance of an abrupt transition from behavioral flocking $n_e \approx 1$ to complete collapse $n_e \approx n$ in the regime of small Mellin exponents. This transition occurs once the mutation parameter exceeds a relatively small threshold value and is especially evident in the logarithmic contour plots. The sharpness of the transition, together with the observed system-size dependence, hysteresis, and bistability, demonstrates that the nonlocality-induced collapse is associated with a first-order dynamical phase transition.

From a dynamical perspective, these results indicate that sufficiently strong long-range observational learning fundamentally alters the balance between selection and mutation. In the local regime, large mutation rates are required to destabilize coherent collective behavior. However, in the strongly nonlocal regime, even weak mutation becomes capable of producing rapid diversification across the entire behavioral space.

Another noteworthy observation is the robustness of the nonlocality-induced collapse phenomenon with respect to network topology. The effect is present in networks with highly heterogeneous degree distributions (Barabási–Albert), in homogeneous random networks (Erdős–Rényi), and in small-world structures (Watts–Strogatz). This suggests that the phenomenon is not tied to a specific topological feature, but instead emerges as a generic consequence of sufficiently strong distance-dependent interactions.

Overall, the computational results strongly support the central hypothesis of this work: long-range observational learning acts as an effective mechanism for accelerating homogenization and globalization in evolutionary dynamics on networks. The Mellin-transformed nonlocal interaction operator therefore provides a natural and

flexible framework for studying the interplay between network topology, mutation, and collective behavioral diversity.

7. Dynamical Response Distance and Average Normalized Dynamical Response

The main motivation of this section is to compare the dynamical responses of different graphs. In other words, when the generalised replicator–mutator dynamics are implemented on different graphs, the systems may respond differently due to their different topological structures. Here, we explore a range of artificial and real-world networks to compare their dynamical responses and investigate which graphs are dynamically close and which are dynamically far apart.

To facilitate comparison between networks of different sizes, we normalize the diversity by the number of nodes. This removes the direct dependence of the diversity measure on network size and places the dynamical response on a common dimensionless scale. In other words, we put the dynamical observable on a common dimensionless scale and evaluate all networks under the same dynamical parameter domain. So we define the dynamical response surface of a graph G by normalizing the diversity as

$$r_G(\mu, s) = \frac{n_e^G(\mu, s)}{n_G}, \tag{7.1}$$

where n_G is the number of nodes of graph G, μ is the mutation parameter and s is the Mellin parameter, which controls the strength of nonlocal interactions. Since $1 \leq n_e^G \leq n_G$, the normalized diversity satisfies

$$\frac{1}{n_G} \leq r_G(\mu, s) \leq 1. \tag{7.2}$$

Definition 57. If G_1 and G_2 are two graphs, we define the *Dynamical Response Distance* (DRD), associated with the generalised replicator–mutator dynamics, as

$$D(G_1, G_2) = \left[\frac{1}{|\Omega|} \int_\Omega \left(r_{G_1}(\mu, s) - r_{G_2}(\mu, s) \right)^2 \, d\mu \, ds \right]^{\frac{1}{2}}, \tag{7.3}$$

where Ω denotes the common parameter domain in (μ, s) and $|\Omega|$ denotes its area. Thus, $D(G_1, G_2)$ measures the root-mean-square difference between the dynamical response surfaces of the two graphs over the considered parameter domain.

If $D(G_1, G_2) \approx 0$, then the two graphs exhibit similar dynamical responses over the considered parameter domain. In contrast, a large value of $D(G_1, G_2)$ indicates that their collective dynamics differ strongly. Since the response surfaces are compared over the same parameter domain, the DRD provides a direct measure of dynamical similarity within the generalised replicator–mutator framework. Mathematically, D is the normalized L^2 distance between response surfaces and therefore satisfies nonnegativity, symmetry, and the triangle inequality. It is a pseudometric on graphs because nonisomorphic graphs may have identical response surfaces on Ω; it becomes a metric on the equivalence classes defined by equality of those surfaces.

For numerical calculations, the above expression is approximated using a two-dimensional composite trapezoidal quadrature. Let $s_1, \ldots, s_{M_s}$ and $\mu_1, \ldots, \mu_{M_\mu}$ denote the grid points, and let $w_i^{(s)}$ and $w_j^{(\mu)}$ denote the corresponding one-dimensional

trapezoidal weights. Defining

$$W_{ij} = w_i^{(s)} w_j^{(\mu)}, \tag{7.4}$$

the numerical DRD is evaluated as

$$D(G_1, G_2) = \left[\frac{\sum_{i=1}^{M_s} \sum_{j=1}^{M_\mu} W_{ij} \left[r_{G_1}(s_i, \mu_j) - r_{G_2}(s_i, \mu_j) \right]^2}{\sum_{i=1}^{M_s} \sum_{j=1}^{M_\mu} W_{ij}} \right]^{\frac{1}{2}}. \tag{7.5}$$

For a meaningful comparison, the same parameter domain and the same scaling of the (μ, s) response surface should be considered for all networks. In particular, in the numerical studies, we use the same values of M_μ and M_s for all networks.

We also define the *average normalized dynamical response* of a graph as

$$\chi_G = \frac{1}{|\Omega|} \int_\Omega r_G(\mu, s), d\mu, ds. \tag{7.6}$$

Here, χ_G is the mean normalized diversity over the parameter domain. It measures the average level of diversity maintained by graph G over the considered ranges of (μ, s).

For numerical calculations, χ_G is evaluated using the same two-dimensional composite trapezoidal quadrature,

$$\chi_G = \frac{\sum_{i=1}^{M_s} \sum_{j=1}^{M_\mu} W_{ij} r_G(s_i, \mu_j)}{\sum_{i=1}^{M_s} \sum_{j=1}^{M_\mu} W_{ij}}. \tag{7.7}$$

The two quantities $D(G_1, G_2)$ and χ_G provide complementary information about the dynamical behaviour of a network. While χ_G characterizes the average response of an individual graph over the parameter domain, $D(G_1, G_2)$ quantifies the difference between the complete dynamical response surfaces of two graphs.

7.1. **Numerical Studies.** We first consider several well-known artificial networks, including (a) the path graph, (b) the cycle graph, (c) the star graph, (d) the complete graph, (e) the Watts–Strogatz network, (f) the Erdős–Rényi network, and (g) the Barabási–Albert network. By implementing the generalised replicator–mutator dynamics and collecting the diversity values over the (μ, s) response surface, we calculate the average normalized dynamical response χ_G for each network. According to the ascending values of χ_G, we order the networks and denote them by $G_1, G_2, G_3, \ldots$. We then calculate the Dynamical Response Distance for every pair of networks.

Figure 7.1 shows the Dynamical Response Distance matrix for the artificial networks. For numerical purposes, we consider $n = 50$ nodes for all the networks, including the path graph, cycle graph, star graph, complete graph, Watts–Strogatz network (initial nearest-neighbour parameter $k = 10$ and rewiring probability $p = 0.1$), Erdős–Rényi network (edge probability $p_{\text{edge}} = 0.1$), and Barabási–Albert network (attachment parameter $m = 5$). After calculating the average normalized

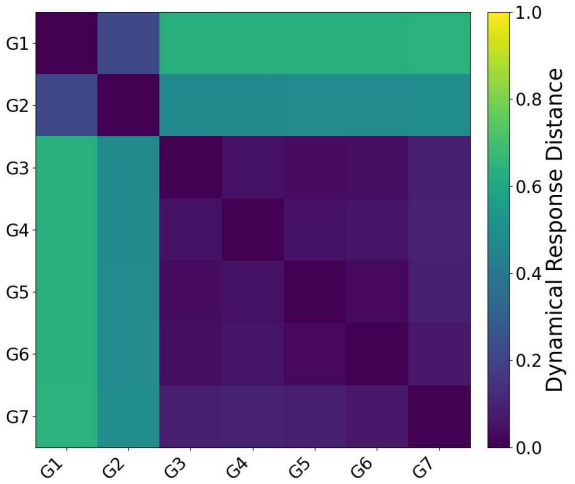

FIGURE 7.1. Dynamical Response Distance matrix for artificial networks. The sequence of graphs is arranged according to the ascending values of χ_G.

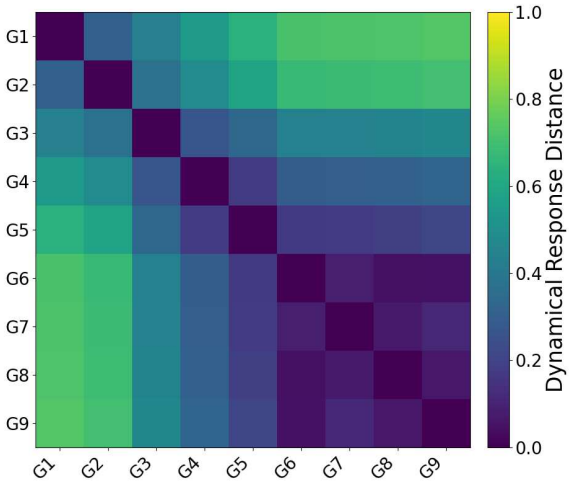

FIGURE 7.2. Dynamical Response Distance matrix for social networks.

dynamical response χ_G, we calculate the DRD for all possible pairs of graphs and represent the results as a matrix in Figure 7.1.

According to the values of χ_G, we assign G_1 to the path graph, G_2 to the cycle graph, G_3 to the Watts–Strogatz network, G_4 to the star graph, G_5 to the Erdős–Rényi network, G_6 to the Barabási–Albert network, and G_7 to the complete graph. The colour intensity in the heat map represents the Dynamical Response Distance between a pair of networks. Consequently, networks with smaller DRD exhibit more similar dynamical responses, whereas larger DRD indicate stronger differences in their collective dynamics.

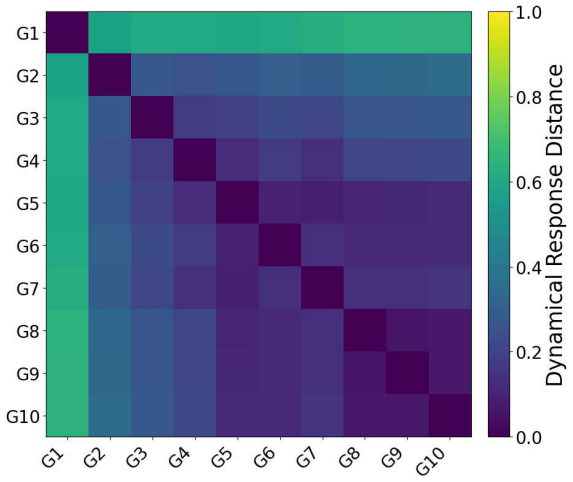

FIGURE 7.3. Dynamical Response Distance matrix for biological networks.

We perform the same analysis for collections of social and biological networks. Figures 7.2 and 7.3 show the corresponding Dynamical Response Distance matrices. The identity, size, edge interpretation, and source information available for the empirical networks are given in Appendix A. For the social networks 7.2, according to the values of χ_G, we assign G_1 to the *Drugs*, G_2 to the *Colorado Springs*, G_3 to the *Heterosexual*, G_4 to the *Twitter Russian Troll*, G_5 to the *Jazz*, G_6 to the *PolBlogs*, G_7 to the *Prison*, G_8 to the *Sawmill*, and G_9 to the *Corporate People* network. In a similar fashion, for the biological networks, we assign G_1 to the *PIN E. coli validate-edges*, G_2 to the *Trans E. coli*, G_3 to the *Transc Yeast*, G_4 to the *PIN B. subtilis*, G_5 to the *PIN E. coli validated-main*, G_6 to the *Trans Urchin*, G_7 to the *PIN A. fulgidus*, G_8 to the *PIN Human*, G_9 to the *PIN Malaria*, and G_{10} to the *KSHV* network.

The above analysis of the dynamical responses of artificial and real-world networks allows us to identify which networks are dynamically close and which differ strongly. This part of the study also indicates how similarly or differently different network structures respond when the generalised replicator–mutator dynamics are implemented.

Note that the numerical methods and computational details for constructing the dynamical response distance matrices are discussed in Section S3 of the Supplementary Information.

8. Discussion and Conclusion

We have generalised the replicator–mutator dynamics by incorporating long-range interactions. We introduced the d-path Laplacian to generalize the interaction structure, and the transformed d-path Laplacian captures interactions across different hop distances. We reconstructed and generalised the Olfati-Saber replicator–mutator model in matrix form by implementing the transformed random-walk d-path Laplacian. The nonlocal adjacency matrix and the corresponding degree matrix of the transformed d-path Laplacian are constructed by considering coefficients

obtained through Mellin or Laplace transformations. These transformations provide natural choices for modelling nonlocal interactions, since, in many systems, the strength of interactions decays with increasing distance. In general, different types of transformations can be considered to construct the nonlocal adjacency matrix; however, we mainly restrict our analysis to the Mellin transformation, as the term d^{-s} produces a power-law decay, which is widely observed in natural systems. Moreover, Mellin-transformed d-path Laplacians have been used to generate superdiffusive transport regimes on infinite lattices and complex networks.

The spectral analysis provides a complementary interpretation of the replicator–mutator dynamics in terms of the underlying network structure. Since the mutation operator Q is directly related to the random-walk Laplacian L_{rw}, the Laplacian eigenmodes provide a natural basis for describing the propagation and attenuation of heterogeneous behavioural patterns. In particular, the spectral gap λ_2 determines the slowest non-trivial mixing mode and therefore provides a characteristic time scale $\tau_{\mathrm{mix}} \sim 1/(\mu\lambda_2)$ for mutation-induced homogenization. This interpretation is consistent with the contrasting behaviour of different graph structures: path and cycle graphs, having small spectral gaps, support long-lived heterogeneous patterns, whereas star and complete graphs, having relatively large spectral gaps, promote rapid mixing. These results establish a connection between network topology, spectral properties, and the characteristic dynamical response of the replicator–mutator system.

We analysed the equilibrium of the generalised replicator–mutator dynamics under a set of assumptions and approximations. We then related the local convergence rate of the nonlocal replicator–mutator system to the spectral structure of the nonlocal random-walk Laplacian. Initially, we proved that the transformed d-path Laplacian is diagonalizable. Considering the existence of an interior equilibrium, we linearized the nonlocal replicator–mutator equation and obtained an effective Jacobian J_{eff}, which contains two parts: the pure replicator part J_0 and the pure mutation part $\mu L_s^T B$. For simplicity and mathematical tractability, we assumed that the mutation parameter μ is small and, using first-order perturbation theory, obtained a formula for the dominant eigenvalue of J_{eff} in terms of the spectral structure of the nonlocal random-walk Laplacian. Moreover, we determined the ε-convergence time for generic initial perturbations and formulated it in terms of the dominant eigenvalue of J_{eff}.

The impact of nonlocality is illustrated through the computational outcomes. In the original Olfati-Saber replicator–mutator model, according to the increase of the mutation parameter μ, four possible emerging behaviours can be observed: (a) behavioural flocking, (b) cohesion, (c) collapse, and (d) complete collapse. In our nonlocal replicator–mutator model, we control the strength of nonlocality through the Mellin parameter s. Ideally, $s \in (0, \infty)$, and computationally, as $s \to 0$, we obtain a very strong nonlocality effect, whereas for large values of s, the contribution of long-range interactions becomes negligible and the model approaches the original replicator–mutator behaviour dominated by local interactions. Thus, the nonlocality-controlling parameter s captures a broad spectrum of interaction strengths, from strong nonlocal interactions to predominantly local interactions. The dynamics are implemented on different network topologies, including Watts–Strogatz, Erdős–Rényi, and Barabási–Albert networks, as well as real-world networks.

When the strength of nonlocality is sufficiently strong ($s \to 0$), we observe a common phenomenon across the considered network topologies, namely, nonlocality-induced collapse. Once the mutation rate exceeds a small threshold, the system undergoes an abrupt transition in diversity n_e, resulting in a direct transition from behavioural flocking ($n_e \approx 1$) to complete collapse ($n_e \approx n$). Through numerical analyses of system-size dependence, hysteresis, and bistability, we demonstrate that the nonlocality-induced collapse phenomenon is associated with a first-order dynamical phase transition. This phenomenon is also observed in several real-world networks, suggesting that, despite the substantial structural differences among the considered network families, they exhibit the same fundamental phenomenology under sufficiently strong long-range interactions. These results suggest that long-range observational learning can act as an effective mechanism for accelerating homogenization and globalization.

We also explored the behaviour of the dynamics on different topologies by defining the Dynamical Response Distance and the average normalized dynamical response. The topological structure plays an important role in determining the dynamical response of different networks. Here, we investigated a range of artificial and real-world networks to examine which networks are dynamically similar and which differ strongly in their dynamical responses. To facilitate comparisons between networks of different sizes, we first normalized the diversity by the number of nodes. In the numerical studies, we calculated the average normalized dynamical response and ordered the networks according to its value. This scalar ordering complements the Dynamical Response Distance matrix, which compares the complete response surfaces.

The overall impact of nonlocality on replicator–mutator dynamics has been examined through the development of the generalised replicator–mutator model, the spectral interpretation of the dynamics, equilibrium analysis, the study of dynamical responses, and extensive numerical investigations. The incorporation of nonlocal interactions provides a framework that is potentially relevant to a broader range of real-world systems in which interactions are not restricted to immediate neighbours. In particular, the interplay between mutation and long-range interactions suggests a possible mechanism through which behavioural homogenization and globalization can emerge in systems with nonlocal observational learning. Such a framework may have potential applications in the study of cultural and behavioural diffusion, social influence, information spreading, and other systems in which local interactions coexist with long-range connections.

Several directions remain open for future investigation. A natural extension is to study the analytical conditions under which the nonlocality-induced collapse occurs, including the dependence of the transition threshold on the mutation parameter μ, the Mellin exponent s, and the spectral properties of the underlying network. It would also be interesting to investigate the dependence of the phenomenon on network size, degree heterogeneity, community structure, and temporal or adaptive network structures. Furthermore, extending the present framework to other forms of distance-dependent interaction kernels beyond the power-law weighting generated by the Mellin transformation may provide a broader class of nonlocal replicator–mutator models. Finally, comparison with empirical temporal and multilayer networks could help determine whether the observed nonlocality-induced

collapse and the associated acceleration of homogenization and globalization are relevant to specific real-world dynamical systems.

Appendix A. Empirical networks used in the numerical analysis

The social-network analysis comprises nine networks. *PolBlogs* is a hyperlink network of political blogs associated with the 2004 U.S. presidential election; the analyzed graph contains 1,224 vertices and 16,715 edges [1]. *Corporate People* is a network of corporate affiliations among people and contains 1,586 vertices and 11,540 edges. *Sawmill* is a communication network within a small enterprise and contains 36 vertices and 62 edges. *Heterosexual* is a contact network of heterosexual individuals recorded during a six-month period and contains 82 vertices and 83 edges. *Prison* is a social network among prison inmates and contains 67 vertices and 142 edges. *Jazz* is a collaboration network among jazz artists and contains 1,265 vertices and 38,356 edges. *Drugs* represents injecting drug users connected when they had shared a needle during the preceding six months; it contains 616 vertices and 2,012 edges [27]. *Colorado Springs* is a contact network of individuals at risk of HIV infection during the early epidemic in Colorado Springs and contains 324 vertices and 347 edges. Finally, *Twitter Russian Troll* is an online interaction network associated with accounts identified as participating in Russian information operations; it contains 1,149 vertices and 2,587 edges. The dataset names and graph sizes follow the source files used in the computations; general repository information for these graph collections is available in Ref. [32].

The biological-network analysis comprises ten molecular networks. *PIN E. coli validate-edges* and *PIN E. coli validated-main* are two validated protein–protein interaction networks for *Escherichia coli*; they contain, respectively, 270 vertices and 716 edges, and 230 vertices and 695 edges. *KSHV* is the protein–protein interaction network of Kaposi's sarcoma-associated herpesvirus and contains 50 vertices and 114 edges. *PIN Malaria* describes protein interactions associated with the malaria parasite and contains 229 vertices and 604 edges. *PIN B. subtilis* is the protein–protein interaction network of *Bacillus subtilis* and contains 84 vertices and 98 edges. *PIN A. fulgidus* is the protein–protein interaction network of *Archaeoglobus fulgidus* and contains 32 vertices and 36 edges. *Transc Yeast* represents direct transcriptional regulation among genes in *Saccharomyces cerevisiae*; it contains 662 vertices and 1,062 edges [26]. *Trans Urchin* is a transcriptional-regulation network for sea urchin and contains 45 vertices and 80 edges. *PIN Human* is a human protein–protein interaction network and contains 2,783 vertices and 6,007 edges. Finally, *Trans E. coli* is a transcriptional-regulation network for *E. coli* and contains 328 vertices and 912 edges [33]. The nomenclature and sizes are those of the adjacency matrices used in the computations; the network-repository source is cited in Ref. [32].

References

[1] Lada A. Adamic and Natalie Glance. The political blogosphere and the 2004 U.S. election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, pages 36–43, 2005.

[2] Avriel Allen-Perkins, Ana B. Serrano, T. A. de Assis, Juan M. Pastor, and Roberto F. S. Andrade. Superdiffusion on complex networks: The role of shortcuts and long-range interactions. *arXiv preprint arXiv:1802.01200*, 2018.

[3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
[4] Daniele Bianchi. Compatibility, embedding and regularization of non-local random walks on graphs. *arXiv preprint arXiv:2101.00425*, 2021.
[5] Michele Coscia. Generalized euclidean measure to estimate network distances. *Proceedings of the International AAAI Conference on Web and Social Media*, 14:119–129, 2020.
[6] Francisco Diaz-Diaz. Time and space generalized diffusion equation on graphs/networks. *arXiv preprint arXiv:2202.00318*, 2022.
[7] Paul Erdős and Alfréd Rényi. On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3–4):290–297, 1959.
[8] Ernesto Estrada. Path laplacian matrices: introduction and application to the analysis of consensus in networks. *Linear algebra and its applications*, 436(9):3373–3391, 2012.
[9] Ernesto Estrada. d-path laplacians and quantum transport on graphs. *Mathematics*, 8(4):527, 2020.
[10] Ernesto Estrada, Jean-Charles Delvenne, Naomichi Hatano, José L. Mateos, Ralf Metzler, Alexander P. Riascos, and Michael T. Schaub. Random multi-hopper model: Super-fast random walks on graphs. *Journal of Complex Networks*, 6(3):382–403, 2017.
[11] Ernesto Estrada, Guillermo Estrada-Rodriguez, and Heiko Gimperlein. Metaplex networks: Influence of the exo-endo structure of complex systems on diffusion. *SIAM Review*, 62(3):617–645, 2020.
[12] Ernesto Estrada, L. V. Gambuzza, and Mattia Frasca. Long-range interactions and network synchronization. *SIAM Journal on Applied Dynamical Systems*, 17(1):672–693, 2018.
[13] Ernesto Estrada, Eman M. Hameed, Naomichi Hatano, and Matthias Langer. Path laplacian operators and superdiffusive processes on graphs. i. one-dimensional case. *Linear Algebra and its Applications*, 523:307–334, 2017.
[14] Ernesto Estrada, Eman M. Hameed, Matthias Langer, and Agnieszka Puchalska. Path laplacian operators and superdiffusive processes on graphs. ii. two-dimensional lattice. *Linear Algebra and its Applications*, 555:373–397, 2018.
[15] Ernesto Estrada and Michael Sheerin. Random neighborhood graphs as models of fracture networks on rocks: Structural and dynamical analysis. *Applied Mathematics and Computation*, 314:360–379, 2017.
[16] Islam I. Hussein. An individual-based evolutionary dynamics model for networked social behaviors. In *Proceedings of the American Control Conference*, pages 5789–5794, 2009.
[17] Swami Iyer, Javier Reyes, and Timothy Killingback. An application of evolutionary game theory to social dilemmas: The traveler's dilemma and the minimum effort coordination game. *PLOS ONE*, 9(4):e93988, 2014.
[18] Natalia L. Komarova. Replicator–mutator equation, universality property and population dynamics of learning. *Journal of Theoretical Biology*, 230(2):227–239, 2004.
[19] Yu-Ping Kuo and Oana Carja. Evolutionary graph theory on rugged fitness landscapes. *bioRxiv*, 2023.
[20] Yu-Ping Kuo and Oana Carja. Evolutionary graph theory beyond pairwise interactions: Higher-order network motifs shape times to fixation in structured populations. *PLOS Computational Biology*, 20(3):e1011905, 2024.
[21] Yu-Ping Kuo and Oana Carja. Evolutionary graph theory beyond single mutation dynamics: On how network-structured populations cross fitness landscapes. *Genetics*, 227(2):iyae055, 2024.
[22] Yu-Ping Kuo, César Nombela-Arrieta, and Oana Carja. A theory of evolutionary dynamics on any complex spatial structure. *bioRxiv*, 2021.
[23] Naomi Ehrich Leonard. Multi-agent system dynamics: Bifurcation and behavior of animal groups. *Annual Reviews in Control*, 38(2):171–183, 2014.
[24] Linjie Liu, Xiaojie Chen, and Matjaž Perc. Evolutionary dynamics of cooperation in the public goods game with pool exclusion strategies. *Nonlinear Dynamics*, 97(1):749–766, 2019.
[25] Yosef E. Maruvka, Tomer Kalisky, and Nadav M. Shnerb. Nonlocal competition and the speciation transition on random networks. *Physical Review E*, 78(3):031920, 2008.
[26] Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon. Superfamilies of evolved and designed networks. *Science*, 303(5663):1538–1542, 2004.

[27] James Moody. Injecting drug users network data, 2001. Data collected with support from NIH grants DA12831 and HD41877.
[28] Reza Olfati-Saber. Evolutionary dynamics of behavior in social networks. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 4051–4056, 2007.
[29] Karen M. Page and Martin A. Nowak. Unifying evolutionary dynamics. *Journal of Theoretical Biology*, 219(1):93–98, 2002.
[30] David Pais and Naomi Ehrich Leonard. Limit cycles in replicator–mutator network dynamics. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 3922–3927, 2011.
[31] Tomasz Raducha and Maxi San Miguel. Coordination and equilibrium selection in games: The role of local effects. *Research Square*, 2021.
[32] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pages 4292–4293, 2015.
[33] Shai S. Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon. Network motifs in the transcriptional regulation network of Escherichia coli. *Nature Genetics*, 31(1):64–68, 2002.
[34] Matteo Smerlak. Neutral quasispecies evolution and the maximal entropy random walk. *Science Advances*, 7(16):eabb2376, 2021.
[35] Jorge M. Tavares, M. M. Telo da Gama, and Antonio Nunes. Coherence thresholds in models of language change and evolution: The effects of noise, dynamics, and network of interactions. *Physical Review E*, 77(4):046108, 2008.
[36] Peter D. Taylor and Leo B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1):145–156, 1978.
[37] Lluís Torres-Hugas, Jordi Duch, and Sergio Gómez. Indirect influence on network diffusion. *Communications Physics*, 8(1):494, 2025.
[38] Erik van Nimwegen, James P. Crutchfield, and Martijn A. Huynen. Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences*, 96(17):9716–9720, 1999.
[39] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.
[40] Claus O. Wilke. Adaptive evolution on neutral networks. *Bulletin of Mathematical Biology*, 63(4):715–730, 2001.

Institute for Cross-Disciplinary Physics and Complex Systems (IFISC) CSIC-UIB, 07122 Palma de Mallorca, Spain

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.pdf