Multi-strain spreading dynamics under arbitrary transmission kernels

Sagar Kumar
Network Science Institute, Boston, MA, USA and
Center for Health Informatics Program, Boston Children's Hospital, Boston, MA, USA

Moritz Laber
Network Science Institute, Boston, MA, USA and
Complexity Science Hub Vienna, Vienna, Austria

Maimuna S. Majumder
Center for Health Informatics Program, Boston Children's Hospital, Boston, MA, USA and
Department of Pediatrics, Harvard Medical School, Cambridge, MA, USA

Brooke Foucault Welles
Network Science Institute, Northeastern University, Boston, MA, USA and
Department of Communication, Northeastern University, Boston, MA, USA

István Zoltán Kiss*
Network Science Institute, Northeastern University, London, UK and
Department of Mathematics, Northeastern University, Boston, MA 02115, USA

(Dated: June 26, 2026)

Compartmental models of epidemic dynamics have long described the propagation of a single, immutable transmissible state through a population via pairwise contact, and multi-strain generalizations have extended this framework to incorporate mutation, competition, and cross-immunity. Here we study a minimal generalization with no sink states or feedback, in which transmission acts through an arbitrary column-stochastic kernel Q on a finite set of strains, encoding mutation during transmission with no further structural assumptions. We derive the mean-field approximation for the well-mixed regime and show that it admits an exact closed-form solution for any Q, expressible as a single matrix exponential applied to the initial condition. A spectral decomposition of this solution reveals that the location of the long-time attractor and the rate of approach are governed by the eigenstructure of Q. We extend the analysis to structured populations via a pairwise mean-field approximation on regular contact networks, and validate both approximations against stochastic simulations. The framework provides an entry into the analysis of dynamical systems in which mutation and transmission occur on the same time scale, drawing parallels to the propagation of discrete signals through populations under noisy communication.

I. INTRODUCTION

The classical compartmental models of epidemic dynamics, beginning with the Susceptible-Infected (SI) model and its extensions [1, 2], describe the propagation of a single transmissible state through a population via pairwise contact. Multi-strain generalizations have extended this framework in several directions. For example, models have been constructed for competing pathogens with cross-immunity [3, 4], interactions between strains [5], mutation during transmission via specified parametric structures [6], and mutation along an explicit genotype network with strain-dependent immune interactions [7]. Each of these models includes either some additional structure that constrains the kinds of strain-to-strain transitions the model permits, or includes some feedback or sink state to extend the dynamics beyond a simple spreading. While this additional dynamical machinery is essential for quantitative comparison with real disease spread, it can obscure more basic questions about how the long-time composition of a population depends on the rate and structure of mutation during transmission, and limits extensions of these frameworks to spreading processes beyond epidemiology (e.g., the spread of information, beliefs, or innovations through a population, where what is being transmitted is itself subject to distortion).

In this work, we strip all of this back to study a minimal generalization of the Susceptible-Infected model in which transmission occurs through an arbitrary column-stochastic kernel, Q, acting on a finite set of strains, $\mathcal{A}$. Each transmission event from an individual infected with strain j to a susceptible individual produces a new infection in strain k with probability Q_{kj}. As such, the diagonal elements of Q encode faithful transmission, and the off-diagonal elements encode mutation. No structure is imposed on Q beyond column-stochasticity, so the model accommodates arbitrary mutation patterns within a single framework. This combination of generality and analytical tractability distinguishes the present work from

* Contact author: istvan.kiss@nulondon.ac.uk

prior multi-strain treatments, which have either fixed restrictive parametric forms for the mutation channel in order to retain closed-form analysis, or admitted more general mutation structures at the cost of analytical access to the long-time behavior. We refer to this as the "Noisy" SI model, due to the role played by the transition kernel Q, which is formally identical to a noisy channel in information theory [8–10]. This correspondence motivates a second reading of the model as approximating the propagation of a discrete signal through a population in which each transmission introduces fixed-rate distortion. Seen this way, the Noisy SI model can be considered as generalizing existing treatments of rumor and information dynamics [11–14] which have largely treated transmission either as faithful or as a single-state stochastic event. We develop this communicative interpretation in Sec. VI such that the analytical results of the paper serve both as a contribution to the multi-strain epidemic literature and as a foundation for the communicative case study. The principal results of the paper concern the analytical tractability of this model. We derive a mean-field approximation for the well-mixed regime that admits an exact closed-form solution that is expressible as a single matrix exponential applied to the initial condition. Spectral decomposition of this solution reveals that the long-time attractor is governed by the eigenstructure of Q, with the leading eigenvector determining the location of the attractor in the space of strain densities, and the spectral gap controlling the rate at which trajectories converge to this distribution. We also develop a pairwise approximation for structured populations and find that network sparsity reshapes the long-time attractor in a manner qualitatively similar to increased channel noise. Both approximations are validated against stochastic simulations. In the communicative register, these spectral results translate directly to a statement about information loss: network sparsity, initial seed size, and the spectral structure of the channel together determine how rapidly source information is forgotten as the cascade propagates.

The remainder of the paper is organized as follows. Section II introduces the stochastic model and its reaction-kinetic specification. Section III develops the well-mixed mean-field approximation and its exact solution. Section V extends the analysis to networked populations via pairwise closure. Section IV characterizes the attractor through the spectral structure of Q. Section VI develops the communicative reinterpretation and traces how the spectral structure governs information transmission through the population. We conclude with a discussion of regimes of applicability and connections to information-theoretic and cultural-evolution models in Section VII.

II. STOCHASTIC MODEL

The Susceptible-Infected (SI) model is a special case in which the recovery rate of the Susceptible-Infected-Recovered (SIR) model of disease spread [1] is zero. As such, it partitions populations into two states: susceptible and infected. Its dynamics are defined by the single reaction

$$\mathsf{S} + \mathsf{I} \xrightarrow{\beta} 2\mathsf{I}, \tag{1}$$

which describes an individual in the susceptible state coming into contact with an individual in the infected state and becoming infected with probability β.

To allow for noise or mutation in the spreading contagion, we now extend this so that instead of a singular infected state, there is a set of infected states, $\mathcal{A} = \{\mathsf{I}_1, \mathsf{I}_2, ..., \mathsf{I}_{|\mathcal{A}|}\}$. We assume that each transmission from an infected node to a susceptible node occurs over a noisy channel [8], Q, such that

$$\mathsf{S} + \mathsf{I}_j \xrightarrow{\beta Q_{kj}} \mathsf{I}_j + \mathsf{I}_k \tag{2}$$

where Q_{kj} is the probability that a susceptible individual successfully infected by an individual in state I_j will transition into state I_k. In a two-strain setting, the full set of reactions is thus

$$\begin{aligned}
\mathsf{S} + \mathsf{I}_1 &\xrightarrow{\beta Q_{11}} 2\mathsf{I}_1 \\
\mathsf{S} + \mathsf{I}_1 &\xrightarrow{\beta Q_{21}} \mathsf{I}_1 + \mathsf{I}_2 \\
\mathsf{S} + \mathsf{I}_2 &\xrightarrow{\beta Q_{12}} \mathsf{I}_1 + \mathsf{I}_2 \\
\mathsf{S} + \mathsf{I}_2 &\xrightarrow{\beta Q_{22}} 2\mathsf{I}_2.
\end{aligned}$$

III. WELL-MIXED MEAN FIELD

A. Deriving the Well-Mixed Noisy SI Mean Field

Following the procedure outlined in [15], we can define a linear transition function

$$f_{SI_k} = \beta \bar{k} \sum_{j}^{|\mathcal{A}|} Q_{kj} [SI_j],$$

which reflects the contribution from each strain to I_k, the number of individuals in state I_k at that time. Here, we use $\bar{k}$ to refer to the average number of interactions per unit time, and $[SI_j]$ to denote the expected number of edges between susceptible individuals and individuals infected with strain j, which, for a well-mixed population, is simply the product of I_j and the number of susceptible individuals, S. Because we do not allow transitions $\mathsf{I}_j \to \mathsf{I}_k$, there are no outgoing reactions, and this is the only transition function to be considered. Thus, the expected number of individuals in state I_k can be expressed as

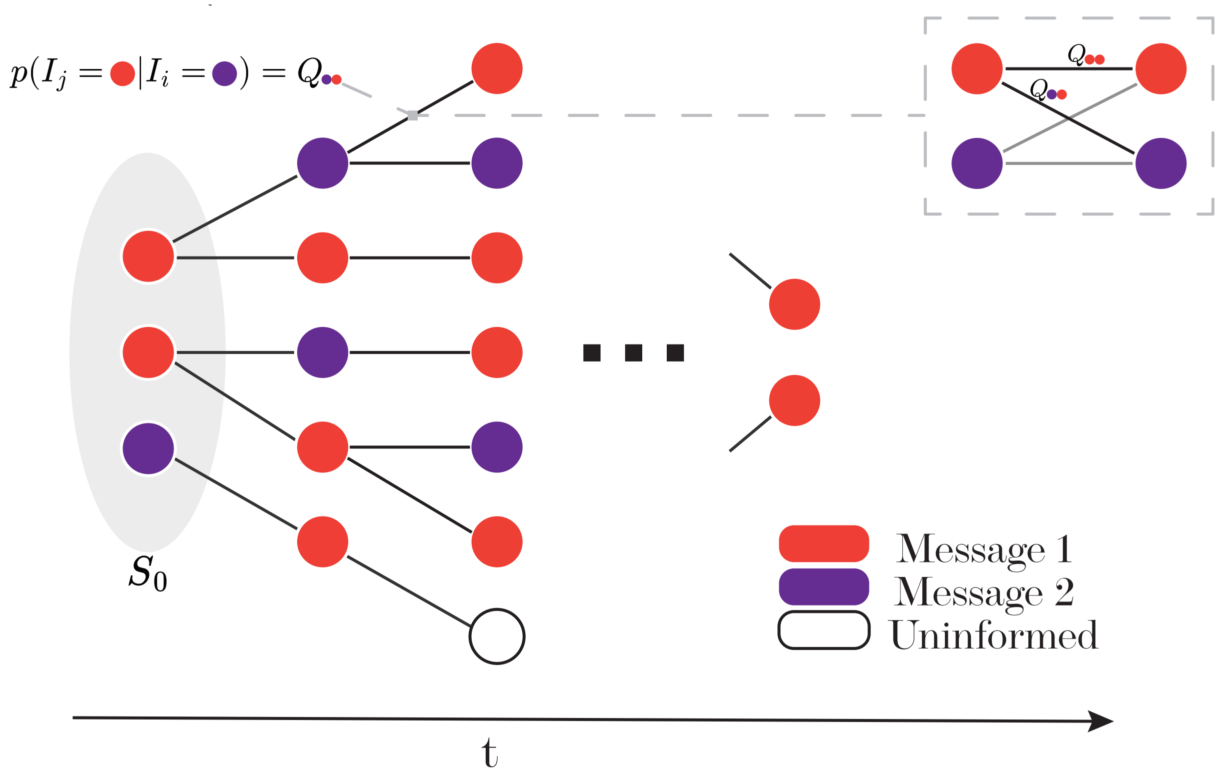

FIG. 1. **Illustration of the Noisy SI model.** Each transmission event (edges going forward in time) is mediated by a noisy channel Q, such that the strain transmitted from an infected individual need not match the strain received by a newly infected individual, as indicated by the colors of the nodes. The probability that a susceptible individual infected by an individual in strain j transitions into strain k is Q_{kj}, the corresponding entry of the column-stochastic kernel Q. The diagonal of Q encodes faithful transmission and the off-diagonal entries encode mutation.

$$\frac{d}{dt}[I_k] = \beta\bar{k}\sum_{j}^{|\mathcal{A}|} Q_{kj}[SI_j].$$

We can express this as a single variable by introducing a state (column) vector $\boldsymbol{x}(t) = ([I_1]/N, [I_2]/N, ..., [I_{|\mathcal{A}|}]/N)$, given a population of size, N. For closed populations,

$$S = N - I_1 - I_2 - ... - I_{|\mathcal{A}|} \tag{3}$$

we can then write the mean-field approximation for noisy susceptible-infected dynamics in a homogeneous population as

$$\frac{d}{dt}\boldsymbol{x}(t) = \beta\bar{k}\left(1 - \sum_{k}^{|\mathcal{A}|}\boldsymbol{x}_k(t)\right)Q\boldsymbol{x}(t). \tag{4}$$

The form of (4) is that of a logistic growth similar to the mean-field approximation for the standard Susceptible-Infected model, with the constants on the left setting a rate, the middle term being the fraction of the population that is still susceptible, and the last term being the fraction of the population in an infected state. Unlike the SI model, however, the well-mixed mean-field approximation of the Noisy SI model introduces Q as "mixing" the infected states.

B. Exact Solution to the Mean-Field Approximation

This mean-field approximation is solvable for any Q by introducing a time-dependent variable

$$y(t) = 1 - \sum_{k=1}^{|\mathcal{A}|}\boldsymbol{x}_k,$$

which translates to the fraction of the population still in the susceptible state. The system can then be written in the vectorial form as

$$\dot{\boldsymbol{x}}(t) = \beta\bar{k}y(t)Q\boldsymbol{x}(t). \tag{5}$$

We first derive a differential equation for y by, in general, assuming that the sum of each column in Q is the same. This is, assuming

$$Q_j = \sum_{k} Q_{kj}$$

does not depend on j. As such, there is a number c, such that $c = Q_j$ for all j. For any column-stochastic matrix, $c = 1$ by definition, but for generality, we will allow it to vary. Now summing the differential equations of $\boldsymbol{x}_k$ for all k yields

$$\begin{aligned} -\dot{y} &= \beta\bar{k}y \sum_{k=1}^{|\mathcal{A}|}\sum_{j=1}^{|\mathcal{A}|} Q_{kj}\boldsymbol{x}_j \\ &= \beta\bar{k}y \sum_{j=1}^{|\mathcal{A}|} \boldsymbol{x}_j \sum_{k=1}^{|\mathcal{A}|} Q_{kj} \\ &= \beta\bar{k}cy(1-y). \end{aligned}$$

This differential equation can be easily solved by separating the variables and decomposing the left side using partial fractions

$$\frac{\dot{y}}{y-1} - \frac{\dot{y}}{y} = \beta\bar{k}c,$$

which can be integrated as

$$y(t) = \frac{y_0 e^{-\beta\bar{k}ct}}{1 - y_0 + y_0 e^{-\beta\bar{k}ct}},$$

where $y_0 = y(0) \in (0,1)$ is the initial condition.

Now we can consider our system (5) as a linear system for the unknown vector $\boldsymbol{x}(t)$ with time-dependent coefficients. However, one can exploit the fact that the time-dependence is in an extremely special form, since each entry of the matrix Q is multiplied by the same scalar $y(t)$. Hence the linear system can be solved in the same way as in the time-independent, i.e. autonomous case using the matrix exponential. Thus let us look for the solution of (5) in the form

$$\boldsymbol{x}(t) = \exp(z(t)Q)\boldsymbol{x}(0).$$

Substituting this form into (5) yields

$$\dot{\boldsymbol{x}}(t) = \dot{z}(t)Q\exp(z(t)Q)\boldsymbol{x}(0) = \dot{z}(t)Q\boldsymbol{x}(t).$$

Therefore we have $\dot{z}(t) = \beta\bar{k}y(t)$ with the initial condition $z(0) = 0$. Hence introducing

$$Y(t) = \int_0^t y(s)ds,$$

the solution of (5) can be written as

$$\boldsymbol{x}(t) = \exp(\beta\bar{k}Y(t)Q)\boldsymbol{x}(0).$$

Integrating the function y yields

$$Y(t) = -\frac{1}{\beta\bar{k}c}\ln\left(1 - y_0 + y_0 e^{-\beta\bar{k}ct}\right).$$

Finally, we can express the limit $\boldsymbol{x}(\infty) = \lim_{t\to\infty}\boldsymbol{x}(t)$ explicitly in terms of the initial condition $\boldsymbol{x}(0)$ as follows

$$\boldsymbol{x}(\infty) = \exp(\beta\bar{k}Y(\infty)Q)\boldsymbol{x}(0), \tag{6}$$

where

$$Y(\infty) = -\frac{1}{\beta\bar{k}c}\ln(1-y_0), \tag{7}$$

with $1 - y_0 = \sum \boldsymbol{x}_i(0)$.

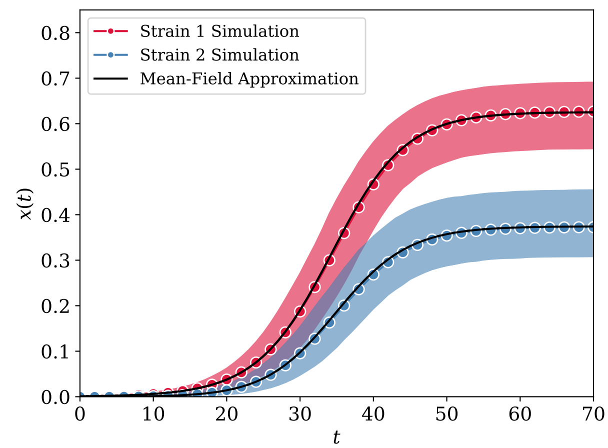

FIG. 2. **Validation of the well-mixed mean-field approximation for the Noisy SI model with binary symmetric channel at $\epsilon = 0.01$.** Dotted curves: stochastic simulation averages over 1000 realizations ($N = 10^4, \beta = 0.05, \bar{k} = 4, \boldsymbol{x}(0) = (10/N, 0)$), shaded regions indicate 95% confidence intervals. Solid black curves: numerical integration of (4). Agreement is excellent across the full trajectory.

C. Agreement with Simulation

Figure 2 demonstrates the agreement between the mean-field approximation of Eq. (4) and stochastic simulations of the underlying reaction system (2) in a well-mixed population.

Defining the transmission kernel as the binary symmetric channel,

$$Q_{\mathrm{BS}} = \begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix},$$

with $\epsilon = 0.01$, the mean-field trajectories of both strains lie within the 2σ confidence band of the simulation averages across the full dynamical range. The endpoint of each strain matches the prediction of Eq. (6) to within sampling noise, providing a direct check that the exact solution derived above describes the long-time behavior of the stochastic process.

IV. SPECTRAL DECOMPOSITION OF THE ATTRACTOR

The exact solution of the well-mixed dynamics reduces the long-time behavior to a single matrix exponential governed by Q. In this section we exploit this structure to characterize the attractor $\boldsymbol{x}(\infty)$ and the manner in which it is approached, and show that both are determined by the spectrum of Q.

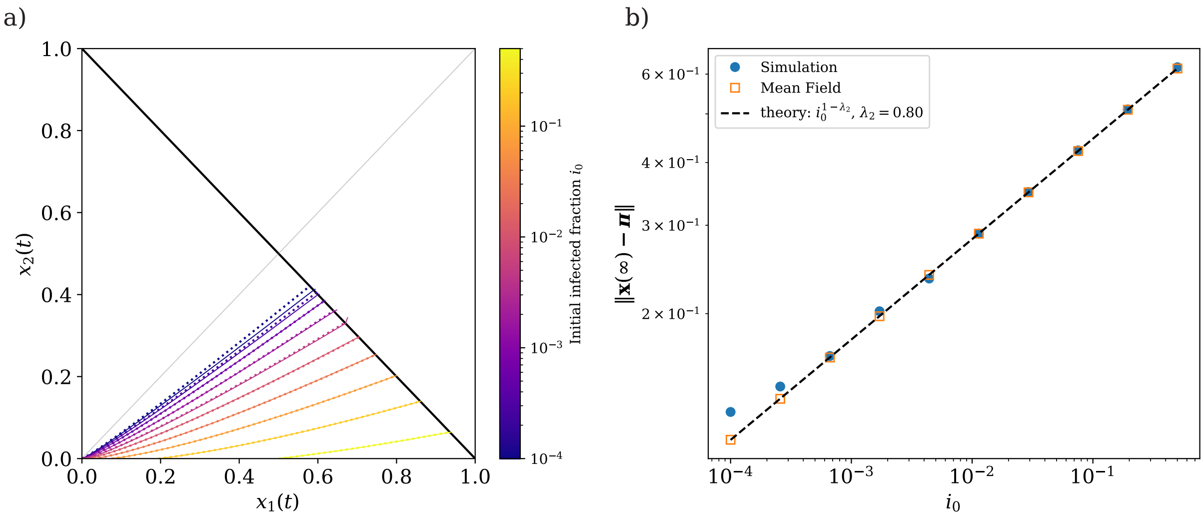

FIG. 3. **Mean-field and simulation agree on how i_0 and Q shape dynamics. a)** Trajectories in the (i_1, i_2)-plane for a fixed, column-stochastic Q_{BS} with $\epsilon = 0.1$. Trajectories are seeded at $(i_0, 0)$ (i.e. all initial mass in strain 1) for i_0 varying logarithmically over a chosen range, and terminate on the diagonal $i_1 + i_2 = 1$ as required by the SI constraint. **2** Solid curves show stochastic simulation averages over 1000 realizations; dashed curves show integrations of the mean-field system. The dominant right eigenvector $\boldsymbol{\pi}$ is marked on the diagonal. Terminal points slide monotonically from the initial-strain vertex toward $\boldsymbol{\pi}$ as i_0 decreases, in quantitative agreement with Eq. (8).

A. Eigenvalue Expansion of the Attractor

To understand how the Q governs the long-time dynamics of the system, we begin by considering all Q that are irreducible, aperiodic, and diagonalizable. We henceforth only consider such Q, except where explicitly stated. Column-stochasticity entails $\lambda_1 = 1$ is an eigenvalue of Q with left eigenvector $\boldsymbol{v}_1 = \mathbf{1}$ and right eigenvector $\boldsymbol{u}_1 = \boldsymbol{\pi}$. Thus, if interpreted as the transition matrix of a Markov chain, $\boldsymbol{\pi}$ is the stationary distribution of Q. To illuminate the relationship between the Noisy SI model and linear transmissions characterized by a random walk on Q, we hereafter refer to $\boldsymbol{\pi}$ as the stationary distribution. We also normalize so that $\mathbf{1}^\top \boldsymbol{\pi} = 1$. All remaining eigenvalues satisfy $|\lambda_m| < 1$, thus the spectral gap, $1 - |\lambda_2|$, is strictly positive.

Using the biorthogonal decomposition

$$Q = \sum_m \lambda_m \, \boldsymbol{u}_m \boldsymbol{v}_m^\top$$

with $\boldsymbol{v}_j^\top \boldsymbol{u}_m = \delta_{jm}$, we find

$$\exp\bigl(\beta \bar{k} c Y(\infty)\, Q\bigr) = \sum_m e^{\beta \bar{k} c Y(\infty) \lambda_m}\, \boldsymbol{u}_m \boldsymbol{v}_m^\top.$$

Writing $\boldsymbol{x}(0) = i_0 \tilde{\boldsymbol{x}}_0$, where $\tilde{\boldsymbol{x}}_0$ is the probability vector specifying the composition of the initial seed across infected strains and $i_0 = I/N$ at time $t = 0$, the $m = 1$ term collapses exactly to $\boldsymbol{\pi}$, as $\boldsymbol{v}_1^\top \boldsymbol{x}(0) = i_0$ and $e^{\beta \bar{k} c Y(\infty)} = 1/i_0$. The remaining terms form a spectral expansion in the sub-leading right eigenvectors of Q,

$$\boldsymbol{x}(\infty) = \boldsymbol{\pi} + \sum_{m \geq 2} i_0^{1-\lambda_m} \, (\boldsymbol{v}_m^\top \tilde{\boldsymbol{x}}_0)\, \boldsymbol{u}_m. \tag{8}$$

Equation (8) is the principal structural result of this section. The attractor decomposes into a universal component $\boldsymbol{\pi}$—the stationary state of a random walk with transition matrix Q, independent of the initial composition–and perturbative corrections along each sub-leading eigendirection, suppressed by the factor $i_0^{1-\lambda_m}$.

B. Geometric Interpretation

For each value of i_0, Eq. (8) defines a simplex of possible attractor states inside the larger $(|\mathcal{A}| - 1)$-simplex of strain compositions. These simplices are nested such as that, as i_0 decreases, the corresponding simplex contracts toward $bm\pi$. For defective Q, the image may collapse to a lower-dimensional face of the simplex, but we restrict attention to diagonalizable Q. In these settings, two limits are immediate. As $i_0 \to 0$, every correction vanishes and the image of the initial simplex contracts to the single point $\boldsymbol{\pi}$, meaning that rare seeding erases all memory of the initial composition. On the other hand, somewhat trivially, as $i_0 \to 1$, then $Y(\infty) \to 0$ and the matrix exponential approaches the identity, so $\boldsymbol{x}(\infty) = \boldsymbol{x}(0)$ and no spreading occurs. Thus, the size of the seed tunes the extent to which the spreading dynamics can act. Dense seeding leaves the population frozen near its initial composition while sparse seeding allows for more transmission events and thus more mixing, driving the system to a steady-state which forgets the initial condition and is determined only by Q. This can be seen in Figure 3a, which shows how trajectories fixed as $i_{-} = (i_1, 0)$ vary with i_i for a fixed Q.

For intermediate i_0, the map $\tilde{\boldsymbol{x}}_0 \mapsto \boldsymbol{x}(\infty)$ is affine, and the image of the initial simplex is a contracted copy nested around $\boldsymbol{\pi}$. The contraction is generically anisotropic so that along the eigendirection $\boldsymbol{u}_m$, the image is compressed by the factor $i_0^{1-\lambda_m}$. Modes with λ_m close to 1 retain visible memory of the initial condition, while modes with small λ_m decouple rapidly. The attractor family thus traces a nested sequence of simplices whose shape is fixed by the spectrum of Q and whose overall scale is set by i_0. Altogether, it is therefore i_0 and the spectrum of Q which control the long-time behavior of the system. In the following sections, we treat these two dependencies in turn.

a. Dependence on i_0 As shown in Figure 3, for a fixed initial distribution $\tilde{\boldsymbol{x}}$, the displacement of the attractor from $\boldsymbol{\pi}$ shrinks as i_0 decreases, at a rate set by the spectral gap of Q. Concretely, the correction in Eq. (8) is driven by $i_0^{1-\lambda_2}$, as higher-mode contributions decay strictly faster in the limit $i_0 \to 0$. This renders the asymptotic scaling

$$\|\boldsymbol{x}(\infty) - \boldsymbol{\pi}\| \sim i_0^{1-\mathrm{Re}\,\lambda_2} \qquad (i_0 \to 0), \tag{9}$$

wherever $\tilde{\boldsymbol{x}}_0 \neq \boldsymbol{\pi}$. As shown in Figure 3b, this means that the displacement of the steady state of the dynamics from the stationary distribution of Q mapped against i_0 has slope $1 - \mathrm{Re}\,\lambda_2$, allowing the spectral gap of the transmission kernel to be inferred from endpoint statistics of simulated or observed spreading dynamics without directly diagonalizing Q.

b. Dependence on the structure of Q Equation (8) implies that the entire attractor family is fixed by the spectrum of Q. Parameterizing Q as

$$Q = \begin{pmatrix} 1-\epsilon_1 & \epsilon_2 \\ \epsilon_1 & 1-\epsilon_2 \end{pmatrix}$$

by its noise rates, ϵ_1, ϵ_2, we can observe how the attractor geometry is systematically deformed. Solving $Q\boldsymbol{\pi} = \boldsymbol{\pi}$, we get the form

$$\boldsymbol{\pi} = \frac{1}{\epsilon_1 + \epsilon_2} \begin{pmatrix} \epsilon_2 \\ \epsilon_1 \end{pmatrix} \tag{10}$$

showing explicitly that the ratio of strain abundances at $\boldsymbol{\pi}$ is set by the inverse ratio of mutation rates. We can use this expression to consider two limits. First, the noiseless channel $Q = I$ is degenerate—every eigenvalue is 1, the contraction factor is $(i_0)^0 = 1$, and the attractor simplex coincides with the initial simplex. This, therefore, recovers the classical multi-strain SI model in which no mixing between strains occurs. Meanwhile, the maximally noisy channel in which $\epsilon_1 = \epsilon_2 = ... = \epsilon_{|\mathcal{A}|} = 1/|\mathcal{A}|$ has $\lambda_m = 0$ for $m \geq 2$. Thus, the contraction factor reduces to i_0, and the attractor collapses almost immediately onto $\boldsymbol{\pi}$. As shown in Figure 4, the stationary state for all symmetric Q is the uniform distribution in $|\mathcal{A}|$ dimensions and as ϵ increases, we see the steady-state of the dynamics approach this value. Meanwhile, when $\epsilon_j \neq \epsilon_k$ for any $j, k \in \mathcal{A}$, the stationary distribution shifts as (10) and the dynamics are pulled towards this distribution proportionally with λ_2. In summary, then, we might think of the two structural roles of the channel matrix, Q, as being cleanly separated. Its eigenvector, $\boldsymbol{\pi}$, sets where the dynamics go, while its spectral gap sets how forcefully they are pulled there.

V. PAIRWISE MEAN FIELD

Relaxing the assumption of a fully connected population to instead consider structured populations, the dynamics of the model can be approximated using a pairwise approximation [15]. This approach considers the expected number of individuals infected as a function of the expected number of edge-linked *pairs* of type SI_k (one susceptible and one k-infected individual which share an edge) at a given moment in time. These models are closed by assuming an approximation for the number of triples of type $\mathsf{I}_j\mathsf{SI}_k$ and SSI_j based on the population topology.

In this approximation, the expected change in the number of nodes in state I_j at any given time is given as

$$\frac{d}{dt}[I_j] = \sum_{k \in \mathcal{A}} \beta Q_{jk} [SI_k] \tag{11}$$

where $[SI_k]$ is the expected number of edges with a susceptible node at one end and a node in state I_k at the other.

Ascertaining the dynamics of edge states becomes more complex. To begin, we first recognize that the expected influx of SI_j edges at each time step comes from changes in triples (i.e. triangles and cherries). An SI_j edge is thus formed when a successful $j \to j$ infection event occurs in an SSI_j triple, or a $k \to j$ mutation occurs in an SSI_k triple.

We decompose these parts as f^{in}, $f^{out,2}$, $f^{\mathrm{out},3}$ so that

$$\frac{d}{dt}[SI_j] = f^{\mathrm{in}} + f^{\mathrm{out},2} + f^{\mathrm{out},3} \tag{12}$$

We can thus capture the expected influx of SI_j edges as

$$f^{\mathrm{in}} = \sum_{k}^{\mathcal{A}} \beta Q_{jk} [SSI_k], \tag{13}$$

the outflux from infections occurring in SI_j edges as

$$f^{\mathrm{out},2} = -\sum_{k}^{\mathcal{A}} \beta Q_{kj} [SI_j], \tag{14}$$

and the outflux from infections occurring in $I_k SI_j$ edges as

$$f^{\mathrm{out},3} = -\sum_{k}^{\mathcal{A}} \left(\sum_{l}^{\mathcal{A}} \beta Q_{lk} \right) [I_k SI_j]. \tag{15}$$

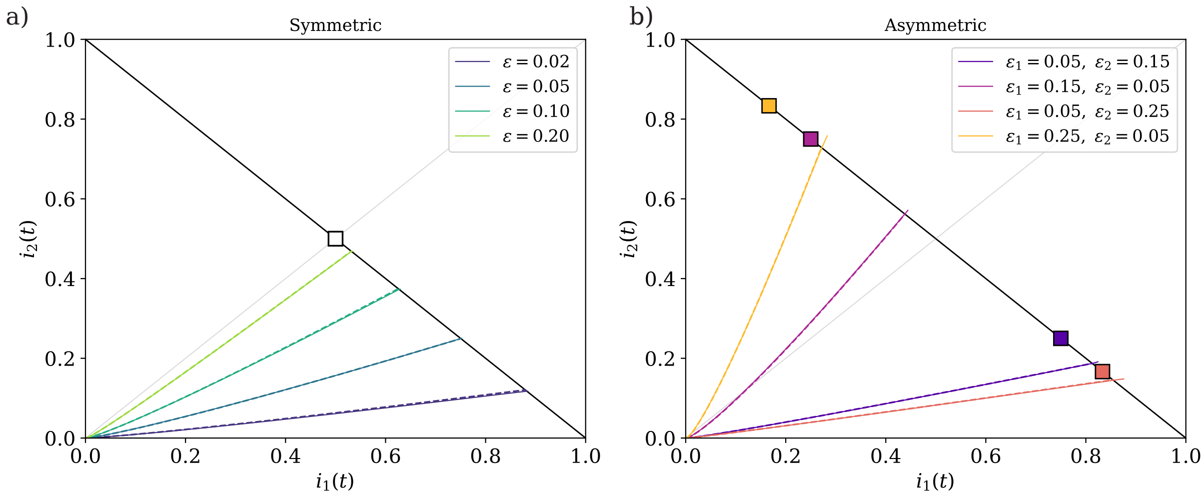

FIG. 4. **Channel asymmetry separately controls the location and the contraction rate of the attractor in the well-mixed two-strain model.** Solid curves show the stochastic simulation averages over 1000 realizations. Dashed curves show the numerical integration of Eq. (5). All trajectories are seeded with $i_0 = 10^{-3}$ all on I_1 and all other parameters are as described in Figure 2. **a)** Trajectories for symmetric Q. Whenever Q is symmetric, the stationary distribution $\boldsymbol{\pi}$ (shown as a white square) is the uniform distribution in $|\mathcal{A}|$ dimensions. **b)** Trajectories for asymmetric Q with their corresponding stationary distributions (shown as colored stars).

In f^{in} for strain j we are counting all the SSI_k triples in which the infected node infects an S to produce an I_j. In $f^{out,2}$ for strain j we are counting all the SI_j edges lost due to infection of the S node by the I_j node to any strain (including j). In $f^{\text{out},3}$ we count all the SI_j edges that are part of $\mathsf{I}_k\mathsf{SI}_j$ triples ($\forall k \in \mathcal{A}$ including j) which are converted to $\mathsf{I}_k\mathsf{I}_\ell\mathsf{I}_j$ ($\forall \ell \in \mathcal{A}$) via infection from a node "outside" of the SI_j edge of interest.

Finally, we must track the number of edges in which nodes at both ends are still susceptible. Because the model has only infections and no recovery, these edges can only be destroyed. Specifically, they are destroyed with any infection event. As such, we have the expected number of SS edges changes as

$$\frac{d}{dt}[SS] = -2\sum_{k}^{\mathcal{A}}\sum_{l}^{\mathcal{A}} \beta Q_{lk}[SSI_k] \tag{16}$$

which depicts the loss of an SS edge occurring when the infected node in an SSI_k triple infects either of the two susceptible nodes (giving us the factor of two in front).

As f^{in}, $f^{\text{out},3}$, and SS all involve third moments, we employ the following moment closure approximation from [15] for arbitrary compartments A, B:

$$[ASB] \approx \kappa \frac{[AS][SB]}{[S]}, \tag{17}$$

where, from [16], $\kappa = \frac{\bar{k}-1}{\bar{k}}$ for $\bar{k}$-regular graphs while $\kappa = 1$ for Poisson-distributed graphs. This approximation makes the simplifying assumption that susceptible and infected nodes are randomly distributed across the network.

Putting together (12)-(17), we obtain

$$\frac{d}{dt}[SI_j] = \frac{(\bar{k}-1)}{\bar{k}[S]}\left([SS]\sum_{k\in\mathcal{A}}\beta Q_{jk}[SI_k] - \sum_{k\in\mathcal{A}}(\sum_{\ell\in\mathcal{A}}\beta Q_{\ell k}[SI_j][SI_k]\right) - \sum_{k\in\mathcal{A}}\beta Q_{kj}[SI_j] \tag{18}$$

$$\frac{d}{dt}[SS] = \frac{2[SS](1-\bar{k})}{\bar{k}[S]}\sum_{k\in\mathcal{A}}\sum_{\ell\in\mathcal{A}}\beta Q_{\ell k}[SI_k], \tag{19}$$

which, alongside (3) and (11), constitute the pairwise approximation of a $\bar{k}$-regular graph.

a. Pairwise dynamics on a regular graph Figure 5 shows the agreement between stochastic simulations of the Noisy SI process on a $\bar{k}$-regular graph and numeri-

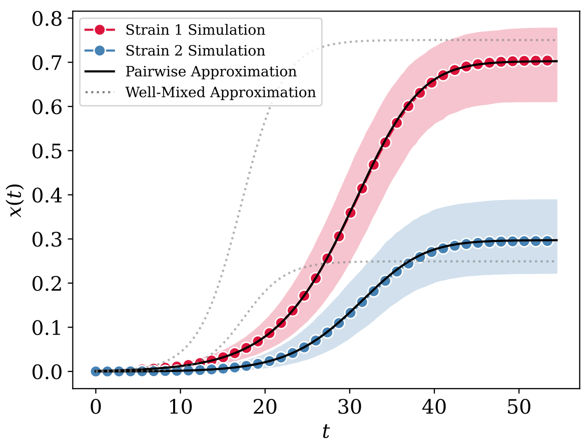

FIG. 5. **Validation of the pairwise mean-field approximation on a $\bar{k}$-regular graph.** Dotted curves show stochastic simulation averages with 5th–95th percentile confidence intervals over 1000 realizations on a random $\bar{k}$-regular graph ($N = 10^4$, $\bar{k} = 4$, $\beta = 0.05$, $\epsilon = 0.05$, $i_0 = 10/N$ all in strain 1). Black curves show numerical integration of the pairwise system of Sec. V. Dotted grey lines show the well-mixed mean-field approximation for the same initial conditions. Agreement between the pairwise approximation and simulations is excellent through the rapid spreading phase, with some simulation lag during the plateau phase.

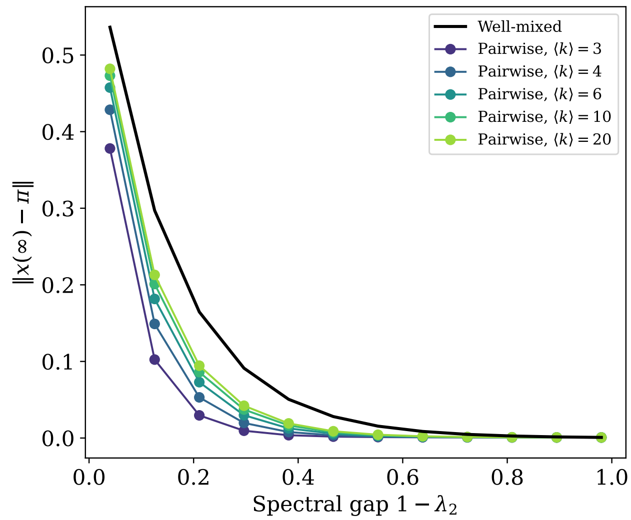

FIG. 6. **Network sparsity drives the pairwise attractor toward the spectral stationary distribution.** Distance from $\boldsymbol{\pi}$, the attractor defined by the spectrum of Q, for the well-mixed approximation (black) and the pairwise approximation on $\bar{k}$-regular graphs at several values of $\bar{k}$ (colored), as a function of the spectral gap $1 - \lambda_2$ of the binary symmetric channel, Q_{BS}. All values were obtained by integrating the corresponding mean-field systems to saturation, with $i_0 = 10^{-3}$ all beginning in the first strain.

cal integration of this system for a dynamics with two strains. The pairwise prediction lies within the confidence band of the simulation across the full trajectory. Two qualitative differences from the well-mixed dynamics are visible in the figure. First, as expected, the onset of spreading is delayed in every pairwise approximation compared to the well-mixed approximation because assumptions about SI edges in the latter create an artificially faster early growth rate [17]. More relevant to the multi-strain dynamics, however, we also find that the saturation plateau is shifted, as both strains approach long-time values closer to the spectral stationary distribution $\boldsymbol{\pi}$ than in the well-mixed case, despite identical Q.

This is further examined in Figure 6, which shows the systematic relationship between graph degree and the location of the attractor predicted by the pairwise approximation. For all $\bar{k}$ tested, the pairwise attractor sits closer to the spectral stationary distribution $\boldsymbol{\pi}$ than the well-mixed attractor, with the effect most pronounced at intermediate spectral gaps and small $\bar{k}$. Heuristically, the pairwise closure encodes the slower per-node spreading rate, extending the effective time available for the channel Q to mix strain identities per unit of forward propagation. The result is that network sparsity acts on the long-time behavior in a manner qualitatively similar to an increase in the channel noise rate—in both cases, the additional mixing pulls the attractor closer to $\boldsymbol{\pi}$.

VI. COMMUNICATION THROUGH A POPULATION

A natural reinterpretation of this model casts it as a model of communication across a population. In this register, the strains of $\mathcal{A}$ are not biological variants but discrete signals or messages, and the channel Q encodes the per-transmission fidelity of message reproduction. As an example, one might imagine the telephone game—a game in which the first speaker produces an utterance that is increasingly distorted as it propagates through a chain of subsequent speakers. The propagation of strains through the population then corresponds to the spread of these progressively distorted utterances, each a mutated descendant of the initial signal. More generally, the model is then interpreted as describing the spread of a noisy signal from an initially informed subset to the entire population, with each peer-to-peer transmission introducing fixed-rate distortion.

a. Setup Let W denote a source that can take one of Ω discrete states with distribution $p(\omega)$. At time $t = 0$, a small fraction i_0 of the population observes the source and encodes its state through a map $\phi : \Omega \to \mathcal{A}$. The remaining population is uninformed and learns about W only through peer-to-peer transmission governed by the Noisy SI dynamics of Sec. II. The composition, $\tilde{\boldsymbol{x}}_0$, of the initial informed subset is then determined by the joint distribution of the source, $p(\omega)$, and the encoding, ϕ.

This setup formalizes a class of scenarios in which information about a localized or hard-to-access source must propagate through a population by serialized, noisy communication. Examples include rumor cascades originating from a small number of eyewitnesses, news propagating from journalists through informal networks of readers, or scientific findings moving from primary literature into popular understanding. In each case one is not interested in the quantity of any particular strain, but how the distribution of strains at initial encoding and throughout the cascade preserves source information.

A. Applications in Information Spreading

a. Mutual Information A natural observable emerging from this framing is the mutual information

$$I[W;x_t]=\sum_{\omega\in\Omega}\sum_{k\in\mathcal{A}}p(\omega)\tilde{x}_k(t)\log\frac{\tilde{x}_k(t)}{\sum_{\omega'\in\Omega}p(\omega')\tilde{x}_k(t)}$$

between the source and the distribution of messages at time t, which quantifies how much about W can be inferred from sampling the population at time, t.

As a functional of the trajectory $\boldsymbol{x}(t)$, the long-term behavior of this quantity is also predominantly determined by the stationary distribution $\boldsymbol{\pi}$ of the channel Q. This implies that the long-time mutual information $I[W;x(\infty)]$ is bounded above by a function of i_0 and the spectrum of Q, and approaches zero as $i_0\to 0$ regardless of how informative the encoding f is at the source. Information about W is degraded by the same spectral mechanism that contracts the simplex of strain compositions toward $\boldsymbol{\pi}$.

b. Implications for Information Spreading This perspective makes several qualitative predictions that distinguish the present framework from simpler models of information spread. First, the asymptotic information content of the cascade depends on the source distribution and the encoding only through their projection onto the sub-leading eigenspace of Q. Encodings that exploit the slowly-mixing directions of Q (i.e., those for which $\phi(\omega)$ varies along eigenvectors with λ_m close to 1) retain more information than encodings aligned with rapidly mixing directions, even at fixed i_0.

Second, comparisons between the well-mixed and pairwise approximations in Sec. V acquires a communicative interpretation. In this light, we can read the results as saying that sparse contact networks dilute information about W more aggressively than dense ones, because the slower per-node spreading rate gives the channel more opportunity to mix message identities before saturation. Thus, the location of the long-time attractor is shifted toward $\boldsymbol{\pi}$, which carries no information about the source.

As a full information-theoretic treatment lies beyond the scope of this work, the aim of this section is to indicate that the analytical structure developed in Secs. III-V carries directly into the communicative setting. Morover, it is to show that the spectral characterization of the attractor governs the long-time fidelity of population-level communication in the same way that it governs the strain composition of an epidemic.

VII. DISCUSSION

The Noisy SI model studied in this work occupies a deliberately minimal position in the multi-strain compartmental literature. By dropping cross-immunity, competition, recovery, and structural constraints on the mutation kernel, what remains is a model whose mean-field dynamics admit an exact closed-form solution and whose long-time behavior is characterized entirely by the spectrum of a single matrix. The principal contributions of this work—the exact solution in the well-mixed regime, the spectral form of the attractor, and the pairwise extension to structured populations—are analytical results that hold for arbitrary column-stochastic Q and that we have validated against stochastic simulations.

The simplicity of the model is both a strength and a limitation. Real epidemic processes typically involve recovery, immunity, and rich strain interactions that the present framework abstracts away. These abstractions may nevertheless be reasonable in regimes where spreading is fast relative to recovery or immune response and where strain interactions are weak, as well as in non-epidemiological applications such as the communicative setting developed in Sec. VI. A separate caveat applies to the mean-field approximations themselves, which hold cleanly only on fully connected populations and on $\bar{k}$-regular graphs. Extensions to heterogeneous-degree, modular, or temporally varying contact structures would require correspondingly extended closure schemes. We expect the spectral characterization of Sec. IV to remain qualitatively intact in these extensions, since the location of the attractor is determined by Q alone, but the rate of convergence and the magnitude of the sparsity-induced shift identified in Sec. V will likely acquire topology-dependent corrections.

Finally, our communicative reinterpretation presented in Sec. VI makes available comparisons with a litany of models and empirical findings concerning the social transmission of heterogeneous content. For instance, iterated learning experiments have shown that transmission chains converge to systematic distortions of input signals reflecting biases of the transmitting agents [18] and studies of online information spreading have documented that different categories of content exhibit markedly different propagation dynamics on the same underlying network [19–21]. Perhaps the most salient analogy comes from the Paris School tradition in cultural evolution, which has framed the convergence of representations across a population in terms of “cultural attractors” [22]. So, while our model is considerably simpler than any that may be found in these literatures, it shares the basic observation that the long-time composition of messages in a population is shaped in part by the noise and mutation inherent to transmission, rather than the original signal alone. The closed-form solution and spectral characterization developed here may thus be useful for studying this kind of convergence in a setting where the underlying dynamics are fully analytically accessible.

VIII. ACKNOWLEDGEMENTS

The authors would like to thank Péter L. Simon of Eötvös Loránd University for his invaluable contributions to solving the well-mixed mean-field approximation.

[1] W. O. Kermack, A. G. McKendrick, and G. T. Walker, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **115**, 700 (1997).

[2] R. M. Anderson and R. M. May, Infectious Diseases of Humans: Dynamics and Control (OUP Oxford, Oxford, UK, 1991).

[3] B. Karrer and M. E. J. Newman, Competing epidemics on complex networks, Physical Review E **84**, 036106 (2011).

[4] C. Castillo-Chavez, W. Huang, and J. Li, Competitive Exclusion in Gonorrhea Models and Other Sexually Transmitted Diseases, SIAM Journal on Applied Mathematics **56**, 494 (1996).

[5] S. Funk and V. A. A. Jansen, Interacting epidemics on overlay networks, Physical Review E **81**, 036118 (2010).

[6] M. Girvan, D. S. Callaway, M. E. J. Newman, and S. H. Strogatz, Simple model of epidemics with pathogen mutation, Physical Review E **65**, 031915 (2002).

[7] B. J. M. Williams, G. St-Onge, and L. Hébert-Dufresne, Localization, epidemic transitions, and unpredictability of multistrain epidemics with an underlying genotype network, PLOS Computational Biology **17**, e1008606 (2021).

[8] C. E. Shannon and W. Weaver, The mathematical theory of communication (Univ. of Illinois Press, Urbana, IL, 1998).

[9] T. M. Cover and J. A. Thomas, ELEMENTS OF INFORMATION THEORY, (Wiley-Interscience, Hoboken, NJ, 2nd ed., 2006).

[10] D. J. C. MacKay, Information theory, inference, and learning algorithms (Cambridge University Press, Cambridge, UK ; New York, 2003).

[11] D. J. Daley and D. G. Kendall, Epidemics and Rumours, Nature **204**, 1118 (1964).

[12] D. P. Maki and M. Thompson, Mathematical models and applications: with emphasis on the social, life, and management sciences (Prentice-Hall, Englewood Cliffs, N.J, 1973).

[13] J. T. Davis, N. Perra, Q. Zhang, Y. Moreno, and A. Vespignani, Phase transitions in information spreading on structured populations, Nature Physics **16**, 590 (2020).

[14] A. Rapoport and L. I. Rebhun, On the mathematical theory of rumor spread, The Bulletin of Mathematical Biophysics **14**, 375 (1952).

[15] I. Z. Kiss, J. C. Miller, and P. L. Simon, Mathematics of Epidemics on Networks: From Exact to Approximate Models, Interdisciplinary Applied Mathematics, Vol. 46 (Springer International Publishing, Cham, 2017).

[16] I. Z. Kiss, E. Kenah, and G. A. Rempała, Necessary and sufficient conditions for exact closures of epidemic equations on configuration model networks, Journal of Mathematical Biology **87**, 36 (2023).

[17] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Reviews of Modern Physics **87**, 925 (2015).

[18] S. Kirby, T. Griffiths, and K. Smith, Iterated learning and the evolution of language, Current Opinion in Neurobiology **28**, 108 (2014).

[19] D. M. Romero, B. Meeder, and J. Kleinberg, Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter, in Proceedings of the 20th international conference on World wide web, WWW '11 (Association for Computing Machinery, New York, NY, USA, 2011) pp. 695–704.

[20] S. Goel, A. Anderson, J. Hofman, and D. J. Watts, The Structural Virality of Online Diffusion, Management Science 10.1287/mnsc.2015.2158 (2015).

[21] J. Berger and K. L. Milkman, What Makes Online Content Viral?, Journal of Marketing Research **49**, 192 (2012).

[22] N. Claidière, T. C. Scott-Phillips, and D. Sperber, How Darwinian is cultural evolution?, Philosophical Transactions of the Royal Society B: Biological Sciences **369**, 20130368 (2014).