# Spreading dynamics in networks under context-dependent behavior

Giulio Burgio, Sergio Gómez, and Alex Arenas*

*Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain*

In some systems, the behavior of the constituent units can create a "context" that modifies the direct interactions among them. This mechanism of indirect modification inspired us to develop a minimal model of context-dependent spreading. In our model, agents actively impede (favor) or not diffusion during an interaction, depending on the behavior they observe among all the peers in the group within which that interaction occurs. We divide the population into two behavioral types and provide a mean-field theory to parametrize mixing patterns of arbitrary type-assortativity within groups of any size. As an application, we examine an epidemic-spreading model with context-dependent adoption of prophylactic tools such as face masks. By analyzing the distributions of groups' size and type-composition, we uncover a rich phenomenology for the basic reproduction number and the endemic state. We analytically show how changing the group organization of contacts can either facilitate or hinder epidemic spreading, eventually moving the system from the subcritical to the supercritical phase and vice versa, depending mainly on sociological factors, such as whether the prophylactic behavior is hardly or easily induced. More generally, our work provides a theoretical foundation to model higher-order contexts and analyze their dynamical implications, envisioning a broad theory of context-dependent interactions that would allow for a new systematic investigation of a variety of complex systems.

## I. INTRODUCTION

Real systems often exhibit higher-order, group interactions that cannot be reduced to combinations of pairwise interactions [1–3]. Even when an interaction is pairwise (e.g., disease spreading), if the behavior of the involved units is modified by the co-presence of other units (e.g., other people), the interaction cannot be considered in isolation. The group as a whole defines a "context" which alters the (direct) interactions taking place within it. These higher-order interactions are known as "indirect modifications" in ecology [4–6], whereby the functional relationship between two species is altered by the presence of a third, trophically disconnected species.

In opinion formation, for instance, the direct influence that an individual can exert on another can be modified by the opinions [7] or the individual characteristics [8] of the others in the group. In linguistics, when an individual moves to a new region, the learning rate of the local language depends on how often the locals use that language in the presence of that individual [9]. They could bear the cost of switching to another language (e.g., English) for a more inclusive conversation, but eventually slowing down the diffusion of the local one; or could instead stick to it, enforcing its diffusion.

In infectious spreading, recent studies have emphasized the importance of considering the group organization of contacts in understanding and controlling the contagion process [10–13]. In this regard, the adoption of nonpharmaceutical prophylactic behavior, such as face mask wearing and physical distancing, proves to be a viable and effective strategy for epidemic control [14–17]. Existing studies on adaptive network models for the spread of contagious disease have considered agents whose behavior depends on individual characteristics, on the observed behavior among peers, and on external sources of information, within both well-mixed and structured populations (see references [62–89] in the review by Benson *et al.* [18], and the review by Wang *et al.* [19]). However, these studies do not distinguish between potentially infectious contacts that occur in isolation versus those that occur in the co-presence of other individuals. This distinction becomes essential when accounting for behavioral adoption, and our proposed model addresses this issue by introducing higher-order context dependence. This framework provides several benefits over existing models, including greater theoretical flexibility, the ability to gain deeper theoretical insights, and the ability to infer parameter values from real-world data. Moreover, our model allows for the analysis of the effects of changing the frequency and assortativity at each group size, which cannot be captured by simpler adaptive models based on pairwise interactions. Indeed, even though each transmission concerns only two people, the likelihood of transmission is indirectly affected by the way contacts are organized within groups, for the adoption of the prophylactic behavior (e.g., wearing a face mask) depends on the level of adoption an individual observes in the entire group [20–22]. Adoption is thus mutable, contextual: An individual may exhibit one behavior or another depending on the current context.

This implies that an appropriate description of the system needs to take into account either that contacts generally occur within groups, and that, concurrently to the infectious spreading, there is a decision-making process at the group level that

shapes the behavior of the individuals therein. Current models either consider the group structure [10–13] or the disease-behavior coupling [18], but never both. In this work, we fill this gap presenting a general model of context-dependent spreading. This is representative of an entire class of models, each identified by the specific spreading process considered and by how agents' behavior is affected by the context. Note that there exist models where behavior is context-independent, such as epidemic models where individuals can adopt permanent prophylactic tools to avoid transmission (e.g., vaccines), or information spreading models where agents divide into active sources or spreaders and passive consumers.

To describe context-dependent behavior we divide individuals into behavioral types, distinguished by a different propensity to adopt a certain behavior. The correlation between individual characteristics and behavior has been observed in various forms, among which in the adoption of prophylactic measures [23–25]. At the same time, similar individuals not only behave similarly, they are also more likely to interact among them than with others, a phenomenon known in sociology as homophily [26]. The local dynamics—and the global one emerging from it—is thus determined by both, the behavior of the types and the way they mix [27–31]. To account for this, we first provide in Sec. III a mean-field approximation to the higher-order structure of interactions in a population divided into two types. We thus formulate an annealed heterogeneous mixing in hypergraphs, which is a contribution in itself. Then, in Sec. IV, we model the context-dependency of agents' behavior. Together, Secs. III and IV provide the theoretical foundation for the study of context-dependent processes analytically. As an application of our theory, in Sec V we consider an epidemic spreading model incorporating context-dependent adoption of prophylactic behavior, then explored in Sec. VI. There, we reveal a rich phenomenology for the basic reproduction number and the endemic state in relation to the size and type-composition distributions of the groups, unveiling the important dynamical implications of accounting for the indirect modifications of the pairwise contagion created by a higher-order context.

## II. THE MODEL

Let us consider a population of agents within which a certain entity (e.g., a pathogen, a rumour) can spread from one agent to another in a pairwise fashion. Such an interaction generally occurs in the presence of other agents not directly participating to it, i.e., within a group of some size (in particular, of size two if no other agent is present). If the co-present agents can affect the pairwise interaction, the group itself mediates an interaction: An indirect modification. Thus, we henceforth use the term "group" or "higher-order interaction" to indicate a group of agents, defining the context within which the pairwise (direct) interactions conveying the spread take place between any two agents in the group.

Every agent, during each interaction, can either actively behave to modify (hinder or favor) the spread or not. Whether it chooses to adopt such behavior or not is determined by both the behavior it observes among the peers involved in the interaction and some intrinsic characteristics of its own. To account for the latter, suppose to partition the population

into two behavioral types, labeled through letters A and N, standing for "adoption-inclined" and "not adoption-inclined:" accordingly, type A is assigned to those agents which are intrinsically more prone to adopt an active behavior. Then, every time a group of agents forms, each of them chooses a behavior and a spreading event is let to happen within any pair of them. Note the assumption of timescale separation we are implicitly making: The agents decide their behavior instantly, as soon as the group forms, to then let the slow spreading dynamics unfold throughout the duration of the group interaction. This is a reasonable (and convenient) hypothesis in some cases, like the one considered in this work, but could be not in others. Anyway, there is no technical difficulty in relaxing it, still leaving intact the core of the model—indirect modifications generated by context-dependency.

The minimal description given so far does not yet specify how the interactions take place. Among the many possible ones, we henceforth consider the following: At time step $t$, a generic agent $i$ is involved in $k_i^{(n)}(t)$ group interactions of size $n \in \{2, \ldots, n_{\max}\}$. A group of size $n$ can be represented as a hyperedge of cardinality $n$ (or $n$-edge) of a hypergraph [32], whose nodes represent the agents. "Group" and "hyperedge" are synonyms here, as well as "node" and "agent," therefore we will use them indistinctly. The composition of the groups formed at time $t$ can be encoded in the adjacency tensors $\{A_t^{(n)}\}$, where $A_{t\,i_1,\ldots,i_n}^{(n)} = 1$ if the group $\{i_1, \ldots, i_n\}$ is formed at time $t$, and 0 otherwise.

Figure 1 provides an illustrative example of the model, while showing the inadequacy of a pairwise representation for context-dependent processes and the consequent need for higher-order representations. Indeed, since the success of each pairwise transmission depends on the behavior taken on by all the agents in the group, the exact description of the whole process requires knowledge of either the interaction structure (i.e., who gathers with whom, and when) and the functional dependence of the agents' behavior on their type and on the type-composition of a group. The exact specification of the interaction structure requires a huge amount of detailed information. While this is largely unavailable in most of the cases, we are anyway forced to resort to some approximate description if we want to gain clear knowledge from analytical results.

## III. MEAN-FIELD THEORY FOR MIXING IN GROUPS

The coarsest approximation is based on the well-mixed hypothesis, which, assuming that any agent can interact with any other one in the population, allows us to substitute local state variables with global aggregates. Nonetheless, we can enrich and extend this scheme by making the probability for a given interaction to occur be a function of the type of the agents it involves. The structure is thus described by means of probability mass functions providing the probability for each type to participate to groups of various size and type-composition. Put differently, we operate a dimensional reduction based on the assumption that the type of the agents is the sole property needed to infer the interaction structure. By taking into account the way the behavioral types mix, such description allows us to preserve the essential elements that determine the spreading process.
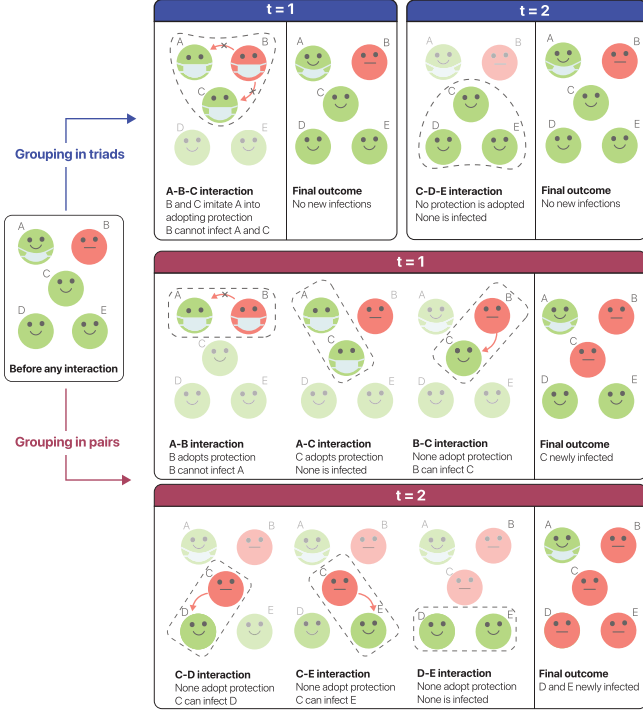
FIG. 1. Example of a context-dependent spreading where the agents can act in order to avoid it (in either directions) by adopting a prophylactic behavior, represented as a worn face mask. At each time $t$, some group interactions (dash-outlined) take place. During each interaction, an infection can occur from any infected (red) agent to any susceptible (green) one. In this example, A-type nodes (represented with a ready-to-use, lowered face mask while no interacting) are assumed to behave actively no matter what the others do, while N-type nodes (those without face mask while no interacting) do it only when they observe another node doing it, i.e., when at least one A-type node is in the group. In this setup, any transmission may be avoided when interacting in groups of three, whereas all the N-type agents may be infected when interacting in pairs. Neglecting the group organization of contacts can thus lead to radically different outcomes, proving the need for a higher-order representation to account for context.

We first present the general mean-field theory valid for any group size, and then we provide explicit formulas for rank-3 hypergraphs.

### A. General case

Consider a population (a set) of $N$ agents (nodes). Let $\rho$ be the proportion of A-type agents, hence $1 - \rho$ the proportion of N-type agents. We assume here that, at each time step, groups form at random, with probabilities depending solely on the types of the agents involved. Accordingly, the value of $A_{t\,i_1,\ldots,i_n}^{(n)}$ depends only on the set $\{X_{i_1}, \ldots, X_{i_n}\}$ of the agents' types, and we replace it with its expected value over the ensemble of hypergraphs parametrized by $N$, $\rho$, and a set of mixing parameters (defined below) regulating the average level of type-(dis)assortativity in a group. To this end, we need an expression for the number $g_{n_A,n_N}^{(n)}$ of $(n_A + n_N)$ edges composed of $n_A$ A-type nodes and $n_N$ N-type ones, being subsets of $n$-edges (i.e., $n_A + n_N \leqslant n \leqslant n_{\max}$), formed at each time step.

From this we can then derive an expression for the probability that an agent of a certain type takes part in a group of any given type-composition. Denoted with $e^{(n)}$ the number of $n$-edges formed at each time step, $g_{n_A,n_N}^{(n)}$ is related to it through

$$p_{n_A,n_N}^{(n)} = \frac{g_{n_A,n_N}^{(n)}}{\binom{n}{n_A + n_N} e^{(n)}}, \tag{1}$$

being $p_{n_A,n_N}^{(n)}$ the probability that sampling $n_A + n_N$ nodes from a randomly chosen $n$-edge [there are $\binom{n}{n_A + n_N} e^{(n)}$ ways of doing it], the sample consists of $n_A$ A-types and $n_N$ N-types. Also, let $p_{l_A,l_N|m_A,m_N}^{(n)}$ be the conditional probability that, given a set of $m_A$ A-type and $m_N$ N-type nodes in an $n$-edge, sampling a set of other $l_A + l_N$ nodes in it ($m_A + l_A + m_N + l_N \leqslant n$), the sample consists of $l_A$ A-types and $l_N$ N-types. Since there are $\binom{m_X + l_X}{m_X}$ ways of choosing $m_X$ nodes out of $m_X + l_X$, $X \in \{A, N\}$, and $\binom{n - m_A - m_N}{l_A + l_N}$ ways of choosing $l_A + l_N$ nodes among the $n - m_A - m_N$ remaining ones, we get

$$p_{l_A,l_N|m_A,m_N}^{(n)} = \frac{\binom{m_A+l_A}{m_A}\binom{m_N+l_N}{m_N}}{\binom{n-m_A-m_N}{l_A+l_N}} \frac{g_{m_A+l_A,m_N+l_N}^{(n)}}{g_{m_A,m_N}^{(n)}}. \tag{2}$$

In particular, $p_{l_A,l_N|1,0}^{(n)}$ ($p_{l_A,l_N|0,1}^{(n)}$) is the sought probability that an A-type (N-type) node takes part in a group of size $n$ which includes other $l_A$ A-types and $l_N$ N-types. Therefore, indicated with $k_{1,0}^{(n)}$ ($k_{0,1}^{(n)}$) the expected $n$ degree of an A-type (N-type) node, i.e., the expected number of groups of size $n$ to which it takes part in per time step, and given $l_A + l_N = n - 1$, then $\kappa_{l_A,l_N|1,0}^{(n)} = p_{l_A,l_N|1,0}^{(n)} k_{1,0}^{(n)}$ ($\kappa_{l_A,l_N|0,1}^{(n)} = p_{l_A,l_N|0,1}^{(n)} k_{0,1}^{(n)}$) is the expected number of groups of size $n$ in which an A-type (N-type) takes part with other $l_A$ A-types and $l_N$ N-types. These numbers, beside $N$ and $\rho$, specify completely the interaction structure.

To compute them, we need an expression for $g_{n_A,n_N}^{(n)}$. For mixed subsets, i.e., with $n_A, n_N \geqslant 1$, we have (see Appendix A 1 for the detailed derivation)

$$g_{n_A,n_N}^{(n)} = \frac{N_{0,1} k_{0,1}^{(n)}}{n_N} \binom{n-1}{n_A + n_N - 1}\binom{n_A + n_N - 1}{n_A}$$
$$\times p_{0,1|0,1}^{(n)} p_{0,1|0,2}^{(n)} \cdots p_{0,1|0,n_N-1}^{(n)}$$
$$\times p_{1,0|0,n_N}^{(n)} p_{1,0|1,n_N}^{(n)} \cdots p_{1,0|n_A-1,n_N}^{(n)}, \tag{3}$$

and for uniform N-type and A-type subsets,

$$g_{0,n_N}^{(n)} = \frac{N_{0,1} k_{0,1}^{(n)}}{n_N} \binom{n-1}{n_N - 1} p_{0,1|0,1}^{(n)} p_{0,1|0,2}^{(n)} \cdots p_{0,1|0,n_N-1}^{(n)}, \tag{4}$$

$$g_{n_A,0}^{(n)} = \frac{N_{1,0} k_{1,0}^{(n)}}{n_A} \binom{n-1}{n_A - 1} p_{1,0|1,0}^{(n)} p_{1,0|2,0}^{(n)} \cdots p_{1,0|n_A-1,0}^{(n)}. \tag{5}$$

Inserting Eqs. (3) to (5) into Eq. (2) we are thus expressing $p_{l_A,l_N|m_A,m_N}^{(n)}$ only in terms of conditional probabilities of finding a single node of given type, i.e., of the form $p_{1,0|m_A,m_N}^{(n)}$ and $p_{0,1|m_A,m_N}^{(n)}$. The latter are the only free variables left. To close the theory, given $m_A + m_N = m \leqslant n - 1$, we parametrize them as follows:

$$p_{1,0|m_A,m_N}^{(n)} = \alpha_{m,m_A}^{(n)} \rho, \quad \text{if } m_A \leqslant m_N, \tag{6}$$

$$p_{0,1|m_A,m_N}^{(n)} = \alpha_{m,m_A}^{(n)}(1 - \rho), \quad \text{if } m_A > m_N. \tag{7}$$

Their counterparts are found from normalization, e.g., $p_{0,1|m_A,m_N}^{(n)} = 1 - \alpha_{m,m_A}^{(n)}\rho$ if $m_A \leqslant m_N$. Parameter $\alpha_{m,m_A}^{(n)}$ governs the mixing probability in a $(m+1)$ edge, subset of an $n$-edge, conditioned on the presence of $m_A \leqslant m$ A-type nodes in it: $\alpha_{m,m_A}^{(n)} = 1$ corresponds to homogeneous mixing, in which case it is only the proportion of the types in the population to determine the expected mixing; for $m_A \neq m_N$, $\alpha_{m,m_A}^{(n)} < 1$ ($\alpha_{m,m_A}^{(n)} > 1$) indicates (dis)assortativity of the majority in the subset towards the remaining node, which is therefore more (less) likely than expected to be of the same type of the majority; for $m_A = m_N$ (occurring only for $m$ even), $\alpha_{m,m_A}^{(n)} < 1$ ($\alpha_{m,m_A}^{(n)} > 1$) indicates asymmetric preference towards N-type (A-type) nodes.

Since $m_A \in \{0, \ldots, m\}$, there are $m + 1$ parameters characterizing the mixing within $(m + 1)$ edges. However, only one of them is free, eventually implying that the interaction structure is uniquely determined by at most $\binom{n_{\max}}{2}$ parameters (see Appendix A 2 for proof).

Part of our analysis will focus on how the dynamics is affected by how type-assortativity is distributed among the various group sizes. To this end, let us define the level of pairwise homophily (or just "homophily") in $n$-edges, $h^{(n)}$, as the probability that two agents are of the same type when interacting within a group of size $n$,

$$h^{(n)} = \frac{N_{1,0}k_{1,0}^{(n)}p_{1,0|1,0}^{(n)} + N_{0,1}k_{0,1}^{(n)}p_{0,1|0,1}^{(n)}}{N_{1,0}k_{1,0}^{(n)} + N_{0,1}k_{0,1}^{(n)}}$$

$$= 1 - \rho(1-\rho)\frac{\alpha_{1,0}^{(n)}k_{0,1}^{(n)} + \alpha_{1,1}^{(n)}k_{1,0}^{(n)}}{k^{(n)}}, \quad (8)$$

where $k^{(n)} = \rho k_{1,0}^{(n)} + (1-\rho)k_{0,1}^{(n)}$ is the average $n$ degree. Note that this is only one of the several ways in which assortativity can be quantified within groups of size $n > 2$. Indeed, one can consider a weaker notion of assortativity by defining an index for each one of the group compositions where one type is majoritarian (e.g., for triads, in addition to the index counting configurations where the three nodes are of the same type, another one counting those where two of them are of the same type) [33]. In this sense, assortativity in pairs is exceptional, for there is only one way of being majoritarian. We choose to focus on pairwise assortativity because the direct interactions we will consider—those mediating the spreading—are pairwise, hence knowing how frequently different pairs form is of primary importance here. Nonetheless, from the previous analysis we know that fixing $h^{(n)}$ alone is not sufficient to fully specify the group organization. Therefore, unless otherwise specified, we fix the higher-order mixing parameters ($\alpha_{m,m_A}^{(n)}$, $m > 1$) and vary instead the pairwise ones ($\alpha_{1,m_A}^{(n)}$, $m_A = 0, 1$). In particular, chosen a value for the average homophily $h = \sum_{n=2}^{n_{\max}} h^{(n)}(n-1)k^{(n)}/k$, given $k = \sum_{n=2}^{n_{\max}}(n-1)k^{(n)}$ the average pairwise degree (i.e., the average degree over the graph projection of the hypergraph), we vary the type-assortativities $\{h^{(n)}\}$ and analyze how this affects the dynamics. Using Eq. (8), we find

$$h = 1 - (\alpha_{1,0} + \alpha_{1,1})\rho(1-\rho), \quad (9)$$

where $\alpha_{1,0} = \sum_{n=2}^{n_{\max}} \alpha_{1,0}^{(n)}(n-1)k_{0,1}^{(n)}/k = \sum_{n=2}^{n_{\max}} \alpha_{1,1}^{(n)}(n-1)k_{1,0}^{(n)}/k = \alpha_{1,1}$ is the average pairwise mixing parameter.

## B. Pairs and triads

As a minimal application of this formalism, we will consider agents interacting only within pairs (2-edges) and triads (3-edges). To start with, Eqs. (6) and (7) read

$$p_{1,0|0,1}^{(2)} = \alpha_{1,0}^{(2)}\rho, \quad (10)$$

$$p_{0,1|1,0}^{(2)} = \alpha_{1,1}^{(2)}(1-\rho), \quad (11)$$

for $n = 2$, and

$$p_{1,0|0,1}^{(3)} = \alpha_{1,0}^{(3)}\rho, \quad (12)$$

$$p_{0,1|1,0}^{(3)} = \alpha_{1,1}^{(3)}(1-\rho), \quad (13)$$

$$p_{1,0|0,2}^{(3)} = \alpha_{2,0}^{(3)}\rho, \quad (14)$$

$$p_{1,0|1,1}^{(3)} = \alpha_{2,1}^{(3)}\rho, \quad (15)$$

$$p_{0,1|2,0}^{(3)} = \alpha_{2,2}^{(3)}(1-\rho), \quad (16)$$

for $n = 3$. Choosing then $\alpha_{1,0}^{(2)}$, $\alpha_{1,0}^{(3)}$, and $\alpha_{2,1}^{(3)}$ as the free parameters, one finds the relations

$$\alpha_{1,1}^{(n)} = \frac{k_{0,1}^{(n)}}{k_{1,0}^{(n)}}\alpha_{1,0}^{(n)} \ (n = 2, 3), \quad (17)$$

$$\alpha_{2,0}^{(3)} = \frac{\alpha_{1,0}^{(3)}(1 - \alpha_{2,1}^{(3)}\rho)}{1 - \alpha_{1,0}^{(3)}\rho}, \quad (18)$$

$$\alpha_{2,2}^{(3)} = \frac{\alpha_{1,1}^{(3)}\alpha_{2,1}^{(3)}\rho}{1 - \alpha_{1,1}^{(3)}(1-\rho)}. \quad (19)$$

We eventually find the following expressions for the mixing probabilities in triads conditioned on a single node:

$$p_{2,0|0,1}^{(3)} = \alpha_{1,0}^{(3)}\rho\,\alpha_{2,1}^{(3)}\rho, \quad (20)$$

$$p_{1,1|0,1}^{(3)} = 2\alpha_{1,0}^{(3)}\rho(1 - \alpha_{2,1}^{(3)}\rho), \quad (21)$$

$$p_{0,2|0,1}^{(3)} = (1 - \alpha_{1,0}^{(3)}\rho)(1 - \alpha_{2,0}^{(3)}\rho), \quad (22)$$

$$p_{0,2|1,0}^{(3)} = \alpha_{1,1}^{(3)}(1-\rho)(1 - \alpha_{2,1}^{(3)}\rho), \quad (23)$$

$$p_{1,1|1,0}^{(3)} = 2\alpha_{1,1}^{(3)}(1-\rho)\alpha_{2,1}^{(3)}\rho, \quad (24)$$

$$p_{2,0|1,0}^{(3)} = [1 - \alpha_{1,1}^{(3)}(1-\rho)][1 - \alpha_{2,2}^{(3)}(1-\rho)], \quad (25)$$

which, together with Eqs. (10) to (13), specify completely the mixing patterns for a single node within groups of size $n = 2, 3$. Note that, to ensure that all the mixing probabilities lie in [0, 1], the intervals of variation of $\alpha_{1,0}^{(2)}$, $\alpha_{1,0}^{(3)}$, and $\alpha_{2,1}^{(3)}$ must fulfill the following constraints:

$$\alpha_{1,0}^{(2)} \in \left[0, \ \min\left\{\frac{1}{\rho}, \frac{1}{1-\rho}, \frac{k_{1,0}^{(2)}}{(1-\rho)k_{0,1}^{(2)}}\right\}\right], \quad (26)$$

$$\alpha_{1,0}^{(3)} \in \left[0, \ \min\left\{\frac{1}{\rho}, \frac{1}{1-\rho}, \frac{k_{1,0}^{(3)}}{(1-\rho)k_{0,1}^{(3)}}, \frac{k^{(3)}}{3\rho(1-\rho)k_{0,1}^{(3)}}\right\}\right], \quad (27)$$

$$\alpha_{2,1}^{(3)} \in \left[ \max\left\{ 0, \frac{2\alpha_{1,0}^{(3)}\rho - 1}{\alpha_{1,0}^{(3)}\rho^2} \right\}, \ \min\left\{ \frac{1}{\rho}, \frac{1 - \alpha_{1,1}^{(3)}(1-\rho)}{\alpha_{1,1}^{(3)}\rho(1-\rho)} \right\} \right].$$
(28)

In particular, the forth upper bound in Eq. (27) guarantees the feasibility of the structure by constraining the lower bound for $\alpha_{2,1}^{(3)}$ to stay below its upper bound.

When keeping fixed the average homophily $h$ (hence $\alpha_{1,0}$), we get the additional relation

$$\alpha_{1,0}^{(3)} = \frac{1}{2k_{0,1}^{(3)}} \left( \alpha_{1,0}k - \alpha_{1,0}^{(2)}k_{0,1}^{(2)} \right),$$
(29)

with $k = k^{(2)} + 2k^{(3)}$. Equation (29) puts further bounds on $\alpha_{1,0}^{(2)}$ in order for $\alpha_{1,0}^{(3)}$ to satisfy Eq. (27).

Dealing with pairs and triads only, it is convenient to switch to a more explicit notation from now on. In particular, we use $k^{(n)}$ instead of $k_{0,1}^{(n)}$, $n = 2, 3$; $\kappa_{N|N}$ and $\kappa_{A|N}$ instead of $\kappa_{0,1|0,1}^{(2)}$ and $\kappa_{1,0|0,1}^{(2)}$, respectively; and $\kappa_{N,N|N}$, $\kappa_{A,N|N}$, and $\kappa_{A,A|N}$ instead of $\kappa_{0,2|0,1}^{(3)}$, $\kappa_{1,1|0,1}^{(3)}$, and $\kappa_{2,0|0,1}^{(3)}$, respectively (analogously when exchanging N with A). Also, we denote the three free parameters we have as $\alpha_2 \equiv \alpha_{1,0}^{(2)}$, $\alpha_3 \equiv \alpha_{1,0}^{(3)}$, and $\beta_{3,1} \equiv \alpha_{2,1}^{(3)}$, and the average pairwise mixing parameter as $\alpha \equiv \alpha_{1,0} = \alpha_{1,1}$.

## IV. CONTEXT-DEPENDENT BEHAVIOR

The next step is to model the behavioral difference between the two types. The probability of actively modifying the spread can be thought of as consisting of a type-specific, context-independent part (e.g., prosociality [34]) and a context-dependent one, i.e., relying on the observed behavior of the others in the group [20]. The probability of being active is thus a dynamic object. In this regard, a simplifying (but, for some applications, seemingly realistic) assumption is that such probability converges to an equilibrium value during a timescale which is much shorter than the duration of a group interaction (e.g., the decision of wearing a face mask may be considered to be made up in the very first moments of a gathering). With this timescale separation, the probability of being active is assumed to instantly attain its equilibrium value, effectively becoming a time-independent quantity.

Now, given a group of size $n$, let $q_X^{(n_X-1,n-1)}$ be the probability that a X-type agent adopts an active behavior when $n_X - 1$ out of the other $n - 1$ agents in the group are of X types as well. In accordance to the meaning we associated with the labels we require that (i) the probability of behaving actively, within a mixed group, is always higher for an A-type agent than for an N-type; and (ii) such probability increases with the number of A-types present in the group. Taken together they imply the following chain of inequalities:

$$q_A^{(n-1,n-1)} \geqslant q_A^{(n_A-1,n-1)} \geqslant q_N^{(n_N-1,n-1)} \geqslant q_N^{(n-1,n-1)}, \quad (30)$$

for $1 \leqslant n_A \leqslant n - 1$, $n_N = n - n_A$. Equation (30) defines the qualitative behavioral difference between the types.

Having access to empirical measurements of the $q_X$s, one could simply fit them in the dynamic equations. In the absence of empirical information, we may suppose different functional forms for them, satisfying Eq. (30). One possibility is to let the $q_X$s emerge dynamically as a result of a behavioral adaptation,

as shown in Sec. IV A. Another one, presented in Sec. IV B, is to postulate binary functional forms for the $q_X$s, where agents of each type are active or inactive in a deterministic way depending on the composition of the groups they take part in. The simplicity of the binary forms will allow us to obtain explicit analytical results, gaining essential insights to interpret more complicated scenarios.

To notice that any of the behavioral dynamics proposed here can be easily generalized to accommodate additional potential factors, such as external sources of influence (e.g., mass media), seen as a background context modifying the baseline probability of adoption, or some coupling with the spreading dynamics (e.g., awareness raised by information on an epidemic).

### A. Social contagion model

Here we let agents' behavior emerge through a process of social contagion [23,35,36] modeled as a susceptible-infectious-susceptible (SIS) dynamics, where here "S" and "I" represent nonadoption and adoption of the active behavior, respectively, with additional endogenous transitions moving agents from state S to state I and vice versa at rates depending solely on their type. These rates quantify some cost of adopting an active behavior (e.g., the discomfort of wearing a face mask) and the will to bear that cost (e.g., due to prosociality or vulnerability to a disease). Denoting with $c_X$ the I-to-S rate and with $b_X$ the S-to-I rate, $X \in \{A, N\}$, we describe the behavioral dynamics via the following system (see the derivation in Appendix B)

$$\dot{q}_A = \frac{n_N}{n-1}[q_N - q_A] + b_A[1 - q_A] - c_A q_A, \quad (31)$$

$$\dot{q}_N = \frac{n_A}{n-1}[q_A - q_N] + b_N[1 - q_N] - c_N q_N, \quad (32)$$

where we denoted $q_X^{(n_X-1,n-1)}(t)$ simply as $q_X$ for a lighter notation. The first term of each of the two equations contains the context-dependent dynamics, which has the form of a linear consensus formation process [37]; the remaining terms contain the context-independent, type-specific dynamics.

The system admits a unique solution for each possible composition of a group (see Appendix B for details). Interestingly, for a wide range of values of the transition rates ($b$s and $c$s) producing a significant behavioral difference between the types, which is what really motivates their definition, the equilibrium probability of adoption shows a marked nonlinear dependence on the number of A-type (or N-type) agents in the group. This is especially true for $q_N$ ($q_A$ is always high), growing nonlinearly from low values in a uniform N-type group ($n_A = 0$) to medium-high values when all the others in the group are A-types ($n_A = n - 1$).

### B. Binary models

We define here the binary models we call of "easy adaptation" and "hard adaptation." In both of them $q_A = 1$ always, i.e., A-type agents are assumed to behave according to their intrinsic propensity and independently of the context observed

in a group. We then have the following:

(i) In easy adaptation, $q_N$ is 1 when at least an A-type is present (specifically, for rank-3 hypergraphs, in {A,N}, {A,N,N}, and {A,A,N} groups) and 0 otherwise (that is, in uniform N-type groups). N-type agents are easily (maximally) driven by A-types in this scenario.

(ii) In hard adaptation, $q_N$ is 1 when all the other individuals in the group are A-types ({A,N} and {A,A,N} groups) and is 0 otherwise, i.e., N-types are hardly (minimally) driven by A-types in this case.

The only and crucial difference between the two behavioral models is, for rank-3 hypergraphs, the N-type agents' behavior in {A,N,N} groups: The presence of the A-type agent induces active behavior in the two N-types in the easy adaptation model, but not in the hard adaptation one. As shown in the next sections, this difference modifies the spreading process by making it strongly dependent on the behavioral dynamics taking place at group-level, hence on the properties of the higher-order interaction structure.

## V. SPREADING DYNAMICS

Once the interaction structure and the behavior's context-dependency have been modeled, we are ready to investigate how they jointly affect a spreading process taking place upon that structure. For the sake of simplicity, we assume the spreading process to be represented by a SIS dynamics. The latter is a basic model in the study of epidemics, which is precisely the specific application for which the mathematical machinery we developed will be used below. Without loss of generality, we will therefore make use of an epidemiological terminology.

Let thus $\lambda$ be the transmission rate and $\mu$ the recovery rate. We denote with $Y_X \equiv Y_X(t)$ the fraction of agents of type $X \in$ {A, N} in compartment $Y \in$ {S, I} at time $t$. In particular, $I_X$ is the type-specific prevalence for type X, while $I = \rho I_A + (1 - \rho)I_N$ is the prevalence overall. The dynamics can be modeled through the following system of differential equations:

$$\dot{I_A} = -\mu I_A + \lambda S_A(I_A \theta_{A \to A} + I_N \theta_{N \to A}), \tag{33}$$

$$\dot{I_N} = -\mu I_N + \lambda S_N(I_A \theta_{A \to N} + I_N \theta_{N \to N}), \tag{34}$$

with $S_X = 1 - I_X$. Here $\lambda \theta_{X \to Z}$ is the total transmission rate from a X type to a Z type, reading

$$\theta_{A \to A} = \kappa_{A|A} r_A^{(1,1)} s_A^{(1,1)} + \kappa_{A,N|A} r_A^{(1,2)} s_A^{(1,2)} \\ + 2\kappa_{A,A|A} r_A^{(2,2)} s_A^{(2,2)}, \tag{35}$$

$$\theta_{N \to A} = \kappa_{N|A} r_N^{(0,1)} s_N^{(0,1)} + \kappa_{A,N|A} r_N^{(0,2)} s_A^{(1,2)} \\ + 2\kappa_{N,N|A} r_N^{(1,2)} s_A^{(0,2)}, \tag{36}$$

$$\theta_{A \to N} = \kappa_{A|N} r_A^{(0,1)} s_N^{(0,1)} + \kappa_{A,N|N} r_A^{(0,2)} s_N^{(1,2)} \\ + 2\kappa_{A,A|N} r_A^{(1,2)} s_N^{(0,2)}, \tag{37}$$

$$\theta_{N \to N} = \kappa_{N|N} r_N^{(1,1)} s_N^{(1,1)} + \kappa_{A,N|N} r_N^{(1,2)} s_N^{(1,2)} \\ + 2\kappa_{N,N|N} r_N^{(2,2)} s_N^{(2,2)} 2, \tag{38}$$

where $\lambda r_X^{(n_X-1,n-1)} s_Z^{(n_Z-1,n-1)}$ is the effective transmission rate for an interaction occurring within a $n$-edge composed of $n_X$

X-type and $n_Z$ Z-type agents. This is written as the product of an out-going (i.e., depending on the infected, an X-type) transmission probability $r_X^{(n_X-1,n-1)}$ and an in-going (i.e., depending on the susceptible, a Z-type) transmission probability $s_Z^{(n_Z-1,n-1)}$. With the same notation used in Sec. IV, the superscript $(n_X - 1, n - 1)$ indicates how many of the other $n - 1$ nodes in the group are of type X (subscript), given at least an X type is present. In other words, $r_X$ and $s_X$ encode the context within which a direct interaction occurs from the perspective of a X-type agent. Their form is dictated by the specific mechanism through which transmission can be modified and, via their dependence on the agents' probability to adopt active behavior, by how the mechanism is affected by the size and the type composition of a group. For prophylactic mechanisms, $r_X$ and $s_X$ act as reduction factors of the transmission probability.

Going back to Eqs. (33) and (34) and linearizing them around the epidemic-free equilibrium, $(I_A, I_N, S_A, S_N) \approx (0, 0, 1, 1)$, the associated Jacobian matrix reads

$$\mathbf{J} = \begin{pmatrix} \lambda \theta_{A \to A} - \mu & \lambda \theta_{N \to A} \\ \lambda \theta_{A \to N} & \lambda \theta_{N \to N} - \mu \end{pmatrix}, \tag{39}$$

i.e., $\mathbf{J} = \lambda \mathbf{\Theta} - \mu \mathbb{I}_{2 \times 2}$, with $\mathbf{\Theta}$ being the matrix with entries $\{\theta_{X \to Z}\}$, and $\mathbb{I}_{2 \times 2}$ the $2 \times 2$ identity matrix. The basic reproduction number $R$ is then given by the largest eigenvalue of the next-generation matrix (NGM) [38], NGM $= (\lambda/\mu)\mathbf{\Theta}$, as

$$R = \frac{\lambda k_{\text{eff}}}{\mu}. \tag{40}$$

with

$$k_{\text{eff}} = \frac{1}{2}[\theta_{A \to A} + \theta_{N \to N} + \sqrt{(\theta_{A \to A} - \theta_{N \to N})^2 + 4\theta_{N \to A}\theta_{A \to N}}] \tag{41}$$

being the largest eigenvalue of $\mathbf{\Theta}$, representing the effective—for spreading—average pairwise degree. Imposing $R = 1$ provides the epidemic threshold $\lambda_c = \mu/k_{\text{eff}}$, above which the epidemic-free equilibrium is unstable and an endemic state is reached. If there is no modification of the transmission (i.e., $r_X = s_X = 1, \forall X$) and there is no correlation between type and degree ($k_{0,1}^{(n)} = k_{1,0}^{(n)} = k^{(n)}, \forall n$), all agents are equivalent and mixing is irrelevant, hence $k_{\text{eff}} = k$, $R = R_0 \equiv \lambda k/\mu$ ($\lambda_c = \mu/k$), and the homogeneous mean-field threshold of the SIS model on networks is recovered.

Notice that it is possible to get a closed form for the endemic solution of Eqs. (33) and (34). However, the complicated form of the solution does not allow us to draw any noteworthy conclusion directly from it, hence we do not report its expression here. Nonetheless, an approximate solution that preserves the qualitative behavior of the system can be found under high prophylactic efficacy (see Appendix D). Moreover, we anticipate that no qualitative changes are observed in the results presented below when, instead of a SIS dynamics, a susceptible-infectious-recovered (SIR) one is considered. In that case, the results refer to the final attack rate (total fraction of agents that got infected during an outbreak) rather than to the equilibrium prevalence.

### Application to face-mask adoption

To describe an epidemic spreading in the presence of face masks, we need to specify a suitable form for the

probabilities $r_X$ and $s_X$ appearing in Eqs. (35) to (38). To this end, let $\varepsilon_{\text{out}}$, $\varepsilon_{\text{in}} \in [0, 1]$ be the out- and in-going efficacy of a face mask, respectively. The quantities $\varepsilon_{\text{out}} q_X^{(n_X-1,n-1)}$ and $\varepsilon_{\text{in}} q_X^{(n_X-1,n-1)}$ are then the probabilities that an X-type individual wears a mask and that the mask avoids, respectively, out-going transmission when the X type is infectious, and in-going transmission when the X type is exposed to an infection. In more general terms, they quantify how strongly the spreading process can be affected by an agent's behavior: For perfect efficacy, the success of a spreading event is decided in a deterministic way by the eventually adopted behavior; on the contrary, for null efficacy, spreading and behavior fully decouple. Thus we have

$$r_X^{(n_X-1,n-1)} = 1 - \varepsilon_{\text{out}} q_X^{(n_X-1,n-1)}, \tag{42}$$

$$s_X^{(n_X-1,n-1)} = 1 - \varepsilon_{\text{in}} q_X^{(n_X-1,n-1)}. \tag{43}$$

In the following we assume $\varepsilon_{\text{out}} = 0.9$ and $\varepsilon_{\text{in}} = 0.5$. The effect of varying each efficacy will be discussed whenever significant.

## VI. RESULTS

To study how the group structure of the contacts affects the spreading dynamics, we vary the interaction structure along two directions. For given average values of degree $k$ and assortativity $\alpha$, we ask what is the effect of distributing (i) the degree ($k_A^{(n)}$ and $k_N^{(n)}$) and (ii) the assortativity ($\alpha_{1,0}^{(n)}$ and $\alpha_{1,1}^{(n)}$) differently at the varying of the group size $n$. To isolate the role of the mixing, we assume that type and degree are hereafter independent random variables. Therefore, *a priori*, $k_A^{(n)} = k_N^{(n)} = k^{(n)}$, $\forall$ $n$, and consequently $\alpha_{1,1}^{(n)} = \alpha_{1,0}^{(n)} \equiv \alpha_n$. We consider individuals gathering in pairs and triads, i.e., $n = 2, 3$.

In the following, we display and analyze the results for the social contagion model of Sec. IV A. To this end, we leverage the basic understanding provided by the binary models of Sec. IV B. We fix the adoption probabilities by taking $c_A = c_N = 0.05$, $b_A = 20 c_A = 1.0$, and $b_N = c_N/20 = 0.0025$, yielding a probability of adoption of around 95% and 5% in a uniform A-type and N-type group, respectively. In an $\{A, N\}$ 2-edge, the adoption probability jumps to around 87% for the N-type agent and decreases to around 91% for the A-types. In an $\{A, N, N\}$ 3-edge, it goes to around 80% for the N-types and to 88% for the A-types, while in a $\{A, A, N\}$ one, to around 89% and 93%, respectively. Apart from the numerical details, what ultimately matters is that the A-type agents act as indirect modifiers of the infectious interactions in a group by lowering the transmission probability through the prophylaxis induced in the N-types.

### A. Varying the degree distribution among group sizes

Suppose that the mixing parameters for each group size, $\alpha_2$, $\alpha_3$, and $\beta_{3,1}$, are given. We then vary $k^{(2)}$ and $k^{(3)}$ respecting the constraint $k = k^{(2)} + 2k^{(3)}$ and study how this affects the dynamics.

### 1. Reproduction number

Let us first consider the basic reproduction number $R$. The behavior of $R$ generally depends on both the mixing pattern ($\alpha_2$, $\alpha_3$, $\beta_{3,1}$) and the type composition of the population ($\rho$), but in a different way depending on the system being closer to the easy or the hard adaptation scenario. This is observed, for example, for the intermediate scenario we used (see Fig. 2)— of which the easy adaptation limit is a good proxy. Assuming an ideal perfect protection in at least one direction (i.e., $\varepsilon_{\text{out}} = 1$ and/or $\varepsilon_{\text{in}} = 1$), we provide explicit conditions for those dependencies (see Appendix C 1 for derivation). In particular, in easy adaptation, there exists a threshold value

$$\tilde{\rho} = \frac{2\alpha_3 - \alpha_2}{\alpha_3 \beta_{3,1}} \tag{44}$$

at which a N-type individual has the same probability of finding at least an A-types—and so be induced to adoption—in a pair or in a triad. That probability becomes lower in a pair when $\rho < \tilde{\rho}$, leading $R$ to increase with $k^{(2)}$ (for fixed $k$); vice versa, it decreases with $k^{(2)}$ when $\rho > \tilde{\rho}$. As a consequence, expressed the critical threshold as the value $\rho = \rho_c$ at which $R = 1$ (e.g., red curves in Fig. 2), the threshold moves up or down with $k^{(2)}$ depending on whether $\rho_c < \tilde{\rho}$ or $\rho_c > \tilde{\rho}$. Notice that $\tilde{\rho} \in [0, 1]$ requires $\alpha_3 \in [\alpha_2/2, \alpha_2/(2 - \beta_{3,1})]$. When $\alpha_3 < \alpha_2/2$, $\tilde{\rho} < 0$ and $R$ thus decreases with $k^{(2)}$, for it is easier for a N-types to find at least an A-type in a pair than in a triad, whatever the fraction $\rho$ of A-types in the population. Said differently, the induced prophylaxis is less frequent in triads than in pairs because the former are comparatively too much assortative, therefore $R$ is minimized by interacting in pairs only. In the opposite case in which $\alpha_3 > \alpha_2/(2 - \beta_{3,1})$, $\tilde{\rho} > 1$ and $R$ thus always increases with $k^{(2)}$, for a N-types is always more likely to find at least an A-types in a triad than in a pair in this case. Among the groups containing type A, while an $\{A, A, N\}$ triad is dynamically equivalent to an $\{A, N\}$ pair, an $\{A, N, N\}$ triad has the advantage of weakening the N-N contact, with the result that $R$ is minimized by interacting in triads only.

Going back to the scenario considered in Fig. 2, we see that, treating it as if it were an easy adaptation scenario with perfect unilateral protection, $\tilde{\rho} \geqslant 1$ for $\alpha_2 \leqslant \alpha_3$ (since $\beta_{3,1} = 1$), while $\tilde{\rho} \approx 0.65$ for $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$, which is anyway not far from the true threshold ($\rho \approx 0.60$). This is consistent with the results reported in Fig. 2, confirming the explanatory power of the simplified dynamics assumed to derive Eq. (44).

In the hard adaptation limit, under the same assumptions (see Appendix C 1), the shape of $R$ is instead always monotonic, increasing or decreasing with $k^{(2)}$ solely depending on whether $\alpha_2$ is smaller or larger than $\alpha_3$, respectively. Indeed, since there is no indirect mechanism at work in the $\{A, N, N\}$ triads in this case, reducing $R$ simply amounts to rise disassortativity and in turn increase the frequency of adoption for the N-types.

All in all, this analysis proves that, depending on the behavioral properties of the agents, solely changing the group-size distribution can move the system from the subcritical (disease-free state) to the supercritical phase (endemic states), and vice versa.
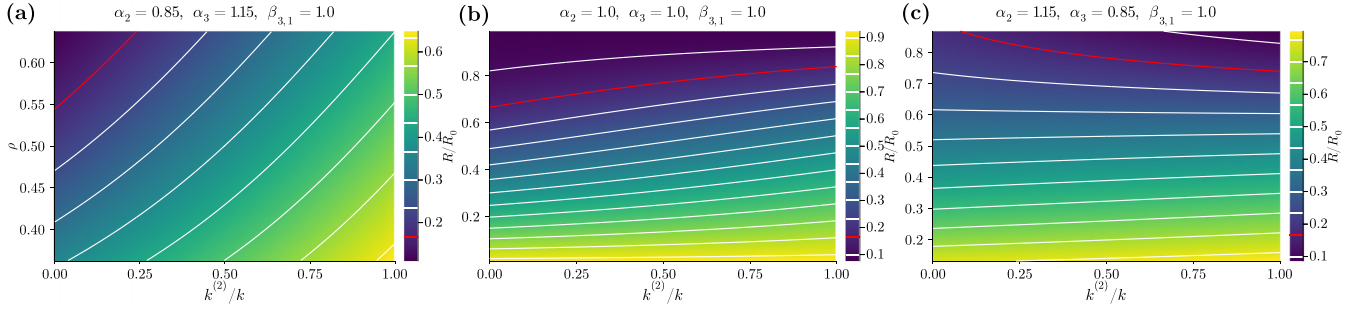
FIG. 2. Normalized reproduction number $R/R_0 = k_{\text{eff}}/k$ computed from Eq. (41) as a function of the proportion of 2-edges, $k^{(2)}/k$, and of the A-type proportion in the population, $\rho$, for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and (a) $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, (b) $\alpha_2 = \alpha_3 = 1.0$, (c) $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. Since $k = k^{(2)} + 2k^{(3)}$, $k^{(2)}/k = 0$ means nodes interacts only within triads, while $k^{(2)}/k = 1$ means they do it just in pairs. White curves report the levels indicated in the respective color bar. The red one indicates the critical curve $R = 1$ ($R/R_0 = 1/6$), above which the disease-free state is stable.

### 2. Prevalence

Looking now at the levels of prevalence at equilibrium, we can first observe that $I_N$ qualitatively behaves like $R$ [compare, for example, Figs. 3(c), 3(f), and 3(i) with Fig. 2], therefore the analysis following Eq. (44)

approximately applies to $I_N$, too. This correspondence between $I_N$ and $R$ is actually expected whenever there is a substantial behavioral difference between the two types and prophylactic efficacy is high. In such case (see Appendix D for proof) the epidemic pressure mainly comes from type N and the dynamics for the latter is well approximated by
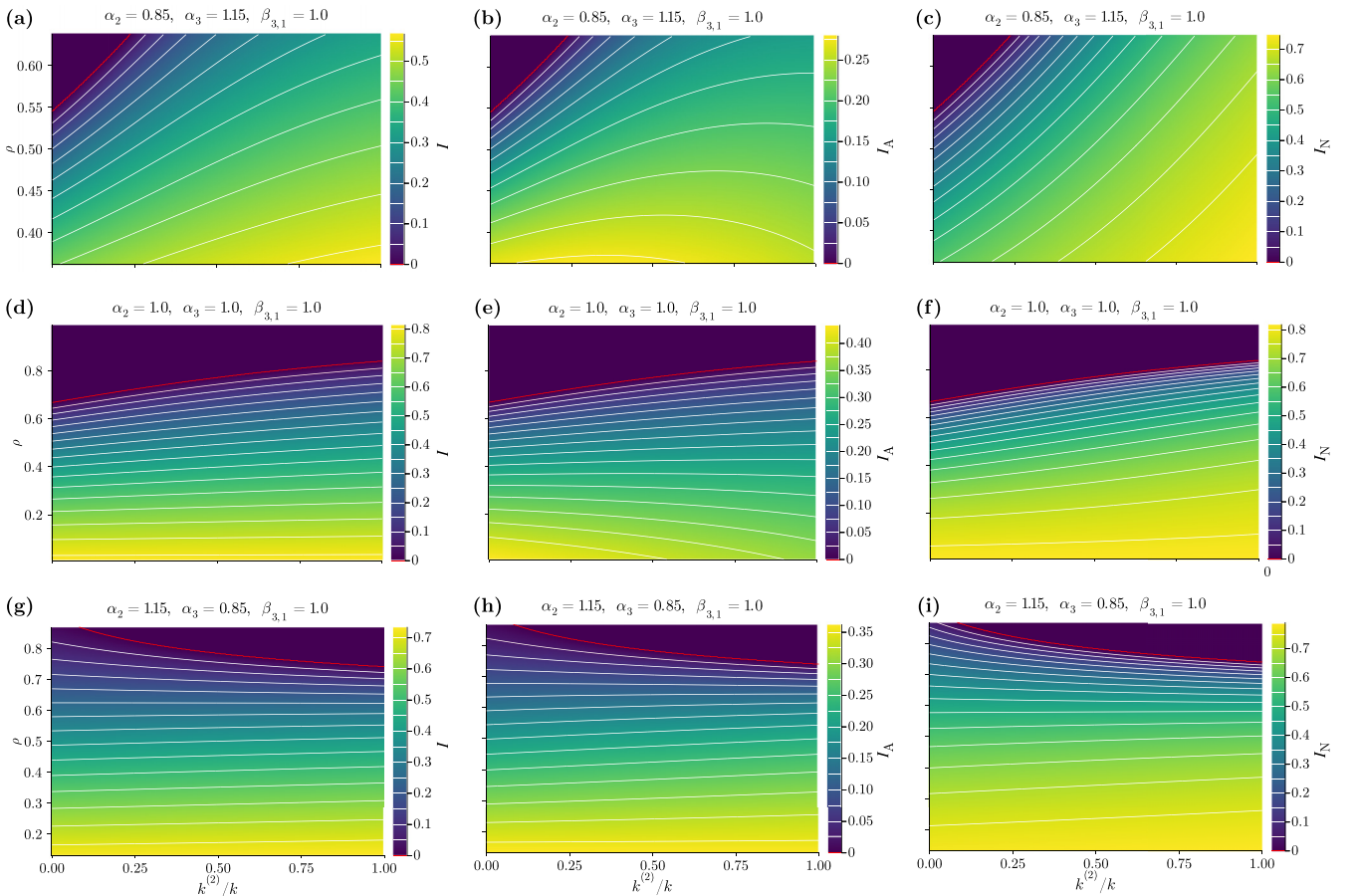


FIG. 3. Equilibrium endemic state [fixed point of the system of Eqs. (33) and (34)] as a function of the proportion of 2-edges, $k^{(2)}/k$, and of the A-type proportion in the population, $\rho$, for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and (a)–(c) $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, (d)–(f) $\alpha_2 = \alpha_3 = 1.0$, (g)–(i) $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. (a), (d), (g) Overall prevalence $I$, (b), (e), (h) prevalence for type A, $I_A$, and (c), (f), (i) prevalence for type N, $I_N$. Since $k = k^{(2)} + 2k^{(3)}$, $k^{(2)}/k = 0$ means nodes interacts only within triads, while $k^{(2)}/k = 1$ means they do it just in pairs. White curves report the levels indicated in the respective color bar. The red one indicates the critical curve $R = 1$, above which the disease-free state is stable.

a one-type population SIS where both $I_N$ and $R$ are largely determined by $\theta_{N\to N}$. On the other hand, since $\theta_{N\to A}$ can show a different and opposite dependence on the parameters compared with $\theta_{N\to N}$, $I_A$ might not generally follow $R$, as seen, for instance, comparing Figs. 3(b), 3(e), and 3(h) with Fig. 2. The results for $I_A$ thus require a more detailed inspection. For the considered scenario, they broadly overlap with those obtained in the easy adaptation one (compare with Fig. S2 of the Supplemental Material [39]). A clear exception is the case $\alpha_2 = \alpha_3$ [Fig. 3(e)], for which, in easy adaptation, $I_A$ always increases with $k^{(2)}$. The discrepancy is explained by observing that, on one hand, such increase gradually vanishes when $\rho$ approaches small values. Indeed, with a low proportion of A-type individuals in the population, the indirect mechanism weakening the N-N interactions which makes $I_N$, and consequently $I_A$, smaller, is largely compensated by the high epidemic pressure that anyway emerges via the more frequent N-type uniform groups. On the other hand, in the hard adaptation limit (see Fig. S4 of the Supplemental Material [39]), $I_A$ decreases rapidly for low values of $\rho$. In this limit, since there is no indirect modification of the N-N interactions, $I_N$ is not directly affected by group size, whereas $I_A$ decreases when trading {A, N, N} triads for {A, N} pairs thanks to the gained bilateral protection. {A, N} and {A, N, N} are thus the groups that regulate the dependence of $I_A$ on $k^{(2)}$, but also those within which the A-type individuals interact the most when $\rho$ is small. Combining the fact that in the intermediate scenario $q_N$ in a {A, N, N} triad is not as high as in a {A, N} pair, thus departing from the easy adaptation limit, with the fact that by decreasing $\rho$ the effect of $k^{(2)}$ on $I_A$ gets weaker in easy adaptation but stronger in hard adaptation, results in the decrease of $I_A$ seen in Fig. 3(e) for low enough $\rho$. Such sort of competition between easy and hard adaptation, determining what we observe in an intermediate scenario, is certainly present also for $\alpha_2 \neq \alpha_3$. However, it has no qualitative impact when $\alpha_2$ and $\alpha_3$ are not too close to each other. In such case, the considered intermediate scenario has the same qualitative behavior than the easy adaptation one. Accordingly, for $\alpha_2 < \alpha_3$ (notice $\beta_{3,1} = 1$ here), increasing $k^{(2)}$, not only makes the weakening of the N-N interactions sparser, but also increases the rate at which N-N interactions take place, yielding a rapid increase of $I_N$ (as we already know from $R$) which, in turn, pushes up $I_A$. At the same time, especially for lower $\rho$, type A benefits from the increase in homophily provided by a higher $k^{(2)}$, as it isolates it from the more infectious type N. Eventually, when $\rho$ is not too high and the increase of $I_N$ gets slower [see Fig. 3(c)], the two contributions to $I_A$ become comparable, giving the latter the nonmonotonic shape observed in Fig. 3(b). This holds whenever efficacy is high at least in one way, so that there is a substantial difference between the epidemic pressures of the two types; otherwise, homophily looses importance and $I_A$ just grows driven by $I_N$.

Finally, for $\alpha_2 > \alpha_3$ [see Figs. 3(h) and 3(i)], $I_A$ mainly behaves as $I_N$. Indeed, since pairs are now less assortative than triads, type A cannot benefit from mixing to compensate for the larger exposure to type N implied by a higher $k^{(2)}$.

All in all, it is $I_N$ to be the main driver of the observed phenomenology in the easy adaptation scenario. This leads to the conclusion that, by exploiting the indirect weakening of
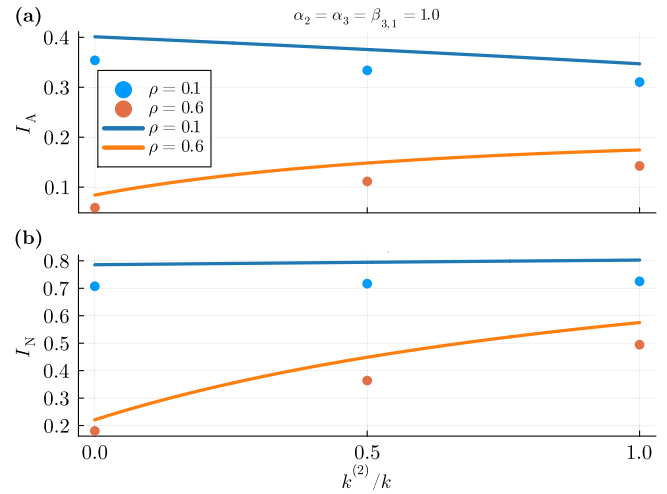


FIG. 4. Monte Carlo results for the equilibrium endemic state as a function of the proportion of 2-edges, $k^{(2)}/k$, for a regular rank-3 hypergraph of $N = 1000$ nodes and pairwise degree $k = k^{(2)} + 2k^{(3)} = 12$. Here, $\lambda = 0.1$, $\mu = 0.2$, and $\alpha_2 = \alpha_3 = \beta_{3,1} = 1.0$ (homogeneous mixing). (a) Prevalence for type A, $I_A$, and (b) for type N, $I_N$. Each point is obtained by averaging over 2500 runs with random initial conditions (infectious state and type assignment). Error bars, representing standard deviations, are smaller than point size. Solid lines represent the results found under the mean-field approximation.

the N-N contacts, the system (made few exceptions) benefits from interacting more in triads than in pairs. In the limit of hard adaptation that role is instead played by $I_A$ (see Sec. S2.1 of the Supplemental Material for a detailed discussion about the results in this scenario). As observed before, while type N is not directly affected by group size in this case, type A benefits from the bilateral protection induced in {A, N} pairs, making pairs preferable to triads for the system overall. What is found for any intermediate scenario, as the one we derived from the social contagion model, eventually depends on how close this is to either one limit or the other.

At last, we considered quenched contact structures (see Appendix E for the numerical implementation of the model) to test whether the presence of topological correlations changed the phenomenology predicted by the mean-field model. Specifically, keeping fixed the mixing parameters to some values, we varied $k^{(2)}/k$ for rank-3 hypergraphs and looked at the equilibrium endemic state. As Fig. 4 shows, the qualitative behavior is accurately reproduced by the mean-field model. Note that the systematic overestimation of the prevalence is actually expected for approximation schemes that ignore dynamical correlations when the system is enough above the epidemic threshold [40].

## B. Varying the type-assortativity distribution among group sizes

We briefly summarize here the results obtained when distributing the assortativity differently between 2- and 3-edges. That is, we fix $\alpha$ and vary $\alpha_2$, with the constraint $\alpha_3 = \alpha + (\alpha - \alpha_2)k^{(2)}/(2k^{(3)})$ [Eq. (29)].

In Appendix C 2 we prove that the reproduction number increases with $\alpha_2$ in the easy adaptation scenario, whereas it decreases in the hard adaptation one.

The same holds also for each type-specific prevalence, hence for the prevalence overall (see Figs. S3 and S5 of the Supplemental Material [39]) and is explained as follows: In the easy adaptation limit, a higher (lower) $\alpha_2$ ($\alpha_3$) means a lower rate at which {A, N, N} triads form, implying that less N-N interactions are weakened. As a consequence, the N-types prevalence, $I_N$, increases. At the same time, since in this scenario the A-N interactions are dynamically equivalent whether they occur isolated or within a triad, the A-types probability of infection is not directly affected by the assortativity distribution. However, since $I_N$ increases, $I_A$ increases as well (unless $\varepsilon_{in} = 1$), although in a milder way being the A-N contacts always bilaterally protected. In hard adaptation, instead, there is no indirect mechanism modifying the N-N interactions, hence $I_N$ is not directly affected by the assortativity distribution. On the other hand, rising (lowering) $\alpha_2$ ($\alpha_3$) means increasing the rate at which the A-N interactions are bilaterally protected, implying a lower $I_A$. In any case, infection from an A-types is unlikely (unless $\varepsilon_{in}$ and $\varepsilon_{out}$ are both small), therefore $I_N$ is only slightly reduced by $\alpha_2$.

As already observed in Sec. VI A, what to expect in any intermediate behavioral scenario depends then on how close it is to either one of the two limit scenarios, the results being generally an interpolation between the two (see Sec. S3 of the Supplemental Material [39] for the results found in the intermediate scenario used in Sec. VI A).

To close the analysis, note that the same qualitative effects of increasing (decreasing) $\alpha_2$ ($\alpha_3$) are also implied by increasing the higher-order mixing parameter $\beta_{3,1}$, for the frequency of the {A, N, N} triads is either way reduced (see Sec. S3 and Fig. S6(b) of the Supplemental Material [39]). While this should not be surprising, as $\beta_{3,1}$ determines the structure as much as any pairwise parameter ($\alpha_2$ and $\alpha_3$), it remarks the fact that exclusively relying on pairwise information may not guarantee an accurate description of the system.

## VII. CONCLUSIONS

We define a minimal model of context-dependent spreading where, during an interaction, an agent either actively behaves to alter the diffusion or not depending on the behavior it observes among the co-present peers. Considering populations where agents are divided into two types (encoding different intrinsic inclinations to take on active behavior), we provide what a mean-field approximation for heterogeneous mixing in hypergraphs, allowing to parametrize mixing patterns of arbitrary type-(dis)assortativity within groups of any size. Choosing then a (spreading) dynamics, the theory makes it possible to obtain analytical results that provide a basic understanding of the system.

Referring to an epidemic spreading model where context-dependency concerns the adoption of prophylactic behavior like face-mask wearing, we have shown that accounting for the behavioral dynamics unfolding at the higher-order level of organization of the interactions can lead to important deviations from what can be expected based on pairwise information alone. We have revealed that the direction and the magnitude of those deviations primarily depend on the properties of the behavioral dynamics; more specifically, on how effective are the (non) adoption-inclined agents in inducing

(inhibiting) adoption of active behavior on the others. Then, depending also on the proportion of adoption-inclined agents in the population and the way type-assortativity distributes among the group sizes, and, secondarily, on the prophylactic efficacy, gathering more often in pairs than in triads (larger groups) can either facilitate or impede the spreading. We have proven this analytically for the basic reproduction number and shown how exclusively changing the group-size distribution can determine whether an outbreak will be subcritical and eventually vanish—or supercritical—leading to endemicity.

More specifically, either for the reproduction number and the prevalence, we can conclude that in general (made few exceptions) when prophylactic behavior is easy to induce, then the system generally benefits by interacting in larger groups, for many otherwise unprotected contacts would be now protected thanks to the elicited adoption. Moreover, the benefit increases when the smaller (larger) groups are the more (dis)assortative. On the contrary, when adoption is hard to induce (e.g., when some majority rule applies), smaller (and disassortative) groups become preferable.

Looking at each type-specific prevalence, we have seen that the type of an agent, not only—as expected—strongly affects its infection risk but also changes the qualitative dependence of that risk on the parameters characterizing the contact structure. Observing this becomes especially important when, for instance, adoption is mostly driven by the vulnerability to the spreading disease. For example, if elderly and young people have very different chances of suffering from severe symptoms—as it is for COVID-19 [41] or influenza [42]—they may also have a dissimilar propensity to protect themselves and the others [43–46]. In such cases, the overall prevalence might not be the most useful indicator, for the prevalence within the vulnerable (and probably adoption-inclined) class does not generally follow the same phenomenology.

Our work stimulates a new direction in the discourse on the interplay between behavior and epidemic spreading. It helps to understand the role of contextual information in decisions to reduce the impact of infectious spreading and provides theoretical support for epidemic control policies based on public gathering restrictions.

In the end, the mean-field approximation provided here opens new possibilities in the modeling of higher-order systems, specifically serving as a basic theory to investigate virtually any process for which mutable and contextual factors can be relevant. Inspired by growing evidence in ecology, we have studied one of such processes and suggested a few more, but perhaps many others have not yet been recognized. We hope that other researchers will be encouraged to look for mechanisms of indirect interaction also in other systems and address the highly challenging problems their complexity entails.

## APPENDIX A: MEAN-FIELD MIXING IN GROUPS

### 1. Counting the subsets of given type composition

In this section we derive the expression for $g^{(n)}_{m_A, m_N}$, given in Eqs. (3) to (5). Consider a subset $\{A_1, \ldots, A_{m_A}, N_1, \ldots, N_{m_N}\}$ of $m_A \geqslant 1$ A-type nodes and $m_N \geqslant 1$ N-type nodes. We want to calculate the probability of finding the other $m_A + m_N - 1$ nodes in the subset conditioned on a single node that we can choose to be, for instance, $N_1$, that is, $P(\{A_1, \ldots, A_{m_A}, N_2, \ldots, N_{m_N}\}|N_1)$. There are $(m_A + m_N - 1)!$ different ways of writing this probability, one for each possible order in which the $m_A + m_N - 1$ nodes can be found. We can thus choose one particular ordering, say $(N_2, \ldots, N_{m_N}, A_1, \ldots, A_{m_A})$, and apply recursively the definition of conditional probability [i.e., $P(\{x, y\}) = P(y)P(x|y)$ and $P(\{x, y\}|z) = P(y|z)P(x|\{y, z\})$] to get

$$
\begin{aligned}
P(\{A_1&, \ldots, A_{m_A}, N_2, \ldots, N_{m_N}\}|N_1) \\
&= (m_A + m_N - 1)! \\
&\quad \times P(N_2|N_1) \\
&\quad\quad \vdots \\
&\quad \times P(N_{m_N}|\{N_1, \ldots, N_{m_N-1}\}) \\
&\quad \times P(A_1|\{N_1, \ldots, N_{m_N}\}) \\
&\quad\quad \vdots \\
&\quad \times P(A_{m_A}|\{N_1, \ldots, N_{m_N}, A_1, \ldots, A_{m_A-1}\}). \quad (A1)
\end{aligned}
$$

If we are not interested in the label of the nodes but exclusively in their type, using the notation $m$X to denote a set of $m$ X-type nodes, Eq. (A1) becomes

$$
\begin{aligned}
P(m_A &A, (m_N - 1)N|N) \\
&= \binom{m_A + m_N - 1}{m_A} \\
&\quad \times P(N|N)P(N|2N)\cdots P(N|(m_N - 1)N) \\
&\quad \times P(A|m_N N)\cdots P(A|m_N N, (m_A - 1)A), \quad (A2)
\end{aligned}
$$

where we divided by $(m_N - 1)!$ and $m_A!$, which are the number of permutations of the $m_N - 1$ N-type and $m_A$ A-type nodes in the sequence, respectively, equally contributing to $P(m_A A, (m_N - 1)N|N)$.

If the $m_A + m_N - 1$ nodes are drawn from a set of $n - 1 \geqslant m_A + m_N - 1$ nodes, there are $\binom{n-1}{m_A+m_N-1}$ distinct ways of doing it for any sequence chosen for the $m_A + m_N - 1$ nodes. Then, if there is a total of $N_N$ N-type nodes in the population and, on average, each one is included in $k_N^{(n)}$ $n$-edges per time step, the expected number $g^{(n)}_{m_A, m_N}$ of $(m_A + m_N)$ edges composed of $m_A$ A-type nodes and $m_N$ N-type nodes, subsets

of $n$-edges, can be written as

$$
\begin{aligned}
g^{(n)}_{m_A, m_N} &= \frac{N_N k_N^{(n)}}{m_N} \binom{n-1}{m_A + m_N - 1} \binom{m_A + m_N - 1}{m_A} \\
&\quad \times P(N|N)P(N|2N)\cdots P(N|(m_N - 1)N) \\
&\quad \times P(A|m_N N)\cdots P(A|m_N N, (m_A - 1)A), \quad (A3)
\end{aligned}
$$

where the term $1/m_N$ accounts for the fact that there are $m_N$ N-type nodes to condition upon in the set. If the subset is N-type uniform, we can just take $m_A = 0$ in Eq. (A3) to obtain

$$
\begin{aligned}
g^{(n)}_{0, m_N} &= \frac{N_N k_N^{(n)}}{m_N} \binom{n-1}{m_N - 1} \\
&\quad \times P(N|N)P(N|2N)\cdots P(N|(m_N - 1)N). \quad (A4)
\end{aligned}
$$

Analogous expressions are found by initially conditioning on an A-type. Apart from the more explicit notation used here, Eqs. (A3) and (A4) are identical to Eqs. (3) and (4).

### 2. Number of free parameters

Since $m_A \in \{0, \ldots, m\}$, Eqs. (6) and (7) imply that there are $m + 1$ parameters characterizing the mixing within $(m + 1)$ edges. To see that only one of them is free, observe from Eq. (2) that the number of mixed groups $g^{(n)}_{m_A+l_A, m_N+l_N}$ ($m_A$, $m_N \geqslant 1$) can be expressed in terms of both $p^{(n)}_{l_A, l_N|m_A, m_N}$ and $p^{(n)}_{l_A-1, l_N+1|m_A+1, m_N-1}$ if $l_A = 1$ ($l_N = 0$), establishing a relation between $\alpha^{(n)}_{m, m_A}$ and $\alpha^{(n)}_{m, m_A+1}$, or, alternatively, in terms of both $p^{(n)}_{l_A, l_N|m_A, m_N}$ and $p^{(n)}_{l_A+1, l_N-1|m_A-1, m_N+1}$ if $l_N = 1$ ($l_A = 0$), providing a relation exists between $\alpha^{(n)}_{m, m_A}$ and $\alpha^{(n)}_{m, m_A-1}$. Since there are $m$ distinct mixed group configurations (one for each $m_A \in \{1, \ldots, m\}$), there are $m$ constraints relating the $m + 1$ mixing parameters, which can thus be expressed in terms of only one of them. Those relations can be found by simply substituting Eqs. (6) and (7) into Eqs. (3) to (5) and, as explained before, comparing related pairs of conditional probabilities (or, equivalently, expressing $g^{(n)}_{m_A, m_N}$ using different orders of search). Notice also that the sets $\{\alpha^{(n)}_{m, m_A}\}$ and $\{\alpha^{(n')}_{m, m_A}\}$, $m < \min\{n, n'\}$, satisfy the same set of relations, as the form of Eqs. (6) and (7) does not depend on $m$: $\alpha^{(n)}_{m, m_A}$ and $\alpha^{(n')}_{m, m_A}$ can only differ in their value, given they computed on different sets ($n$ and $n'$ edges, respectively). Therefore, we can find all the relations by just considering the case $m = n - 1$ for each $n \in \{2, \ldots, n_{\max}\}$. Having one free parameter for each $m \in \{1, \ldots, n - 1\}$, we get $n - 1$ parameters to fix for group size $n$. Given $n \in \{2, \ldots, n_{\max}\}$, the interaction structure is then determined by at most $\sum_{n=2}^{n_{\max}}(n - 1) = \binom{n_{\max}}{2}$ parameters.

## APPENDIX B: DERIVATION OF THE SOCIAL CONTAGION MODEL

We detail here the derivation of the social contagion model presented in Sec. IV A. Consider a group of size $n$, composed of $n_A$ and $n_N$ agents of types A and N, respectively, and denote with $q_X(t)$ the probability of adopting an active behavior for a X types at time $t$. An A-type agent in the adopter state is induced (e.g., through peer pressure) to switch to the nonadopter state by the currently nonadopters in the group. These are, on average, $[1 - q_A(t)](n_A - 1) + [1 - q_N(t)]n_N$. If, instead, the

A-type is in the nonadopter state, it is induced to switch to the adopter one by the $q_A(t)(n_A - 1) + q_N(t)n_N$ adopters that, on average, it currently finds in the group. The same holds for a N-type agent given $n_A - 1$ and $n_N$ in the expressions above are replaced by $n_A$ and $n_N - 1$, respectively. Additionally, agents can spontaneously move between the two behavioral compartments at rates depending solely on their type. These rates quantify some cost of adopting an active behavior and the will to bear that cost. Denoting with $c_X$ the I-to-S rate and with $b_X$ the S-to-I rate, for type X, the adaptive behavioral dynamics is described by the following system of two differential equations

$$
\begin{aligned}
\dot{q}_A = (1 - q_A)\left[ q_A \frac{n_A - 1}{n - 1} + q_N \frac{n_N}{n - 1} \right] \\
- q_A\left[ (1 - q_A)\frac{n_A - 1}{n - 1} + (1 - q_N)\frac{n_N}{n - 1} \right] \\
+ b_A(1 - q_A) - c_A q_A,
\end{aligned}
\tag{B1}
$$

$$
\begin{aligned}
\dot{q}_N = (1 - q_N)\left[ q_N \frac{n_N - 1}{n - 1} + q_A \frac{n_A}{n - 1} \right] \\
- q_N\left[ (1 - q_N)\frac{n_N - 1}{n - 1} + (1 - q_A)\frac{n_A}{n - 1} \right] \\
+ b_N(1 - q_N) - c_N q_N,
\end{aligned}
\tag{B2}
$$

where $q_X \equiv q_X^{(n_X-1, n-1)}(t)$ for a lighter notation. The $n - 1$ in the denominator comes from assuming that each of the $n - 1$ sources weighs the same in affecting the behavior of a focal agent. It is immediate to see that the terms involving only one type $[\propto (1 - q_X)q_X]$, as well as the mixed quadratic terms ($\propto q_A q_N$), cancel out. Therefore, we are left with the linear system defined by Eqs. (31) and (32), and the solutions described after it.

In the solution for mixed groups (i.e., $1 \leqslant n_A \leqslant n - 1$), the constants $C_1$ and $C_2$, which depend on all the figuring parameters, read

$$
C_1 = \frac{(n - 1)b_A b_N + \sum_{X=A,N} n_X b_X}{(n - 1)(b_A + c_A)(b_N + c_N) + \sum_{X=A,N} n_X(b_X + c_X)},
\tag{B3}
$$

$$
C_2 = \frac{n - 1}{(n - 1)(b_A + c_A)(b_N + c_N) + \sum_{X=A,N} n_X(b_X + c_X)}.
\tag{B4}
$$

The difference $q_A - q_N = C_2(b_A c_N - b_N c_A)$ requires $b_A c_N > b_N c_A$ in order for A-type agents to be actually more inclined to be adopters than N-types are. We can thus choose $b_A > c_A$ (or $b_A \gg c_A$) and $b_N < c_N$ (or $b_N \ll c_N$). Also, it requires $C_2$ to be large enough and, from Eq. (B4), we see this means considering not too large values for the rates (specifically for $b_A$ and $c_N$, when $b_A \gg c_A$ and $b_N \ll c_N$). The equilibrium adoption probability for the two types as a function of the composition of a group for different values of the A-type adoption rate, $b_A$, is shown in Sec. S1 of the Supplemental Material [39] for groups of $n = 2, 3, 4$ individuals, making evident the nonlinear dependence on the composition. The nonlinearity increases with $b_A$ and tends to disappear only for $b_A$ approaching $c_A$.

## APPENDIX C: DEPENDENCE OF THE BASIC REPRODUCTION NUMBER ON THE STRUCTURAL PARAMETERS

### 1. Degree

The basic reproduction number $R = \lambda k_{\text{eff}}/\mu$ is here studied as a function of the 2 degree $k^{(2)}$, while keeping fixed the total pairwise degree $k = k^{(2)} + 2k^{(3)}$. We assume type and degree to be uncorrelated, i.e., $k_A^{(n)} = k_N^{(n)} = k^{(n)}$. For any fixed $k$, we want to establish the sign of $\partial R/\partial k^{(2)}|_k = (\lambda/\mu)\partial k_{\text{eff}}/\partial k^{(2)}|_k$. To this end, given the complicated expression of $k_{\text{eff}}$ [Eq. (41)], we simplify it by referring to the binary behavioral scenarios presented in Sec. IV B. In both of them, it holds $r_A = 1 - \varepsilon_{\text{out}}$ and $s_A = 1 - \varepsilon_{\text{in}}$. Moreover, we assume perfect protection in at least one way; that is, $\varepsilon_{\text{in}} = 1$ and/or $\varepsilon_{\text{out}} = 1$.

*a. Easy adaptation.* Here, $r_N = 1 - \varepsilon_{\text{out}}$ and $s_N = 1 - \varepsilon_{\text{in}}$ for $1 \leqslant n_N \leqslant n - 1$, with $\varepsilon_{\text{in}} = 1$ and/or $\varepsilon_{\text{out}} = 1$, and $r_N = s_N = 1$ for $n_N = n$. Only $\theta_{N \to N}$ is nonzero, hence $k_{\text{eff}} = \theta_{N \to N}$ and is given by

$$
k_{\text{eff}} = k^{(2)}(1 - \alpha_2\rho) + (k - k^{(2)})[1 - \alpha_3\rho(2 - \beta_{3,1}\rho)],
\tag{C1}
$$

where the two terms multiplying the degrees account for the probability that an N-type individual takes part to a type-N uniform group of size 2 and 3, respectively, for these are the groups where it is not an adopter. Then we obtain

$$
\left.\frac{\partial k_{\text{eff}}}{\partial k^{(2)}}\right|_k = \rho[\alpha_3(2 - \beta_{3,1}\rho) - \alpha_2],
\tag{C2}
$$

and equaling it to zero, we get the solutions $\rho = 0$ and $\rho = \tilde{\rho}$, with

$$
\tilde{\rho} = \frac{2\alpha_3 - \alpha_2}{\alpha_3\beta_{3,1}},
\tag{C3}
$$

given $\alpha_3\beta_{3,1} \neq 0$. Also, note that $\tilde{\rho} \in [0, 1]$ requires $\alpha_3 \in [\alpha_2/2, \alpha_2/(2 - \beta_{3,1})]$. Technically, $\tilde{\rho}$ results from the fact that the frequency of the {A,A,N} and {A,N,N} triads are both quadratic (respectively, increasing and decreasing) functions of $\rho^2$, whereas the frequency of the {A,N} pairs increases linearly with $\rho$. Finally, differentiating Eq. (C2) with respect to $\rho$ and computing it at $\rho = 0$ and $\rho = \tilde{\rho}$, one finds

$$
\left.\frac{\partial}{\partial \rho}\frac{\partial k_{\text{eff}}}{\partial k^{(2)}}\right|_k = \begin{cases} 2\alpha_3 - \alpha_2 & \text{at } \rho = 0 \\ \alpha_2 - 2\alpha_3 & \text{at } \rho = \tilde{\rho}. \end{cases}
\tag{C4}
$$

Thus, given $\tilde{\rho} > 0$ (i.e., $\alpha_3 > \alpha_2/2$), it follows that $\partial k_{\text{eff}}/\partial k^{(2)}|_k$ is positive for $\rho \in (0, \tilde{\rho})$ and negative for $\rho \in (\tilde{\rho}, 1)$; otherwise, $\partial k_{\text{eff}}/\partial k^{(2)}|_k < 0$ for any $\rho \in (0, 1)$. Rephrasing it, $k_{\text{eff}}$ (hence $R$) is an increasing or decreasing function of $k^{(2)}$ depending on whether $\rho$ is lower or higher than $\tilde{\rho}$.

*b. Hard adaptation.* In this scenario, $r_N = 1 - \varepsilon_{\text{out}}$ and $s_N = 1 - \varepsilon_{\text{in}}$ for $n_N = 1$, with $\varepsilon_{\text{in}} = 1$ and/or $\varepsilon_{\text{out}} = 1$, and $r_N = s_N = 1$ otherwise. We get

$$
k_{\text{eff}} = k^{(2)}(1 - \alpha_2\rho) + (k - k^{(2)})(1 - \alpha_3\rho),
\tag{C5}
$$

the two terms—notice that both are linear in $\rho$—accounting for the probability that an N-type individual finds at least another N-types in a group of size 2 and 3, respectively, not

being an adopter in such cases. We obtain

$$\left.\frac{\partial k_{\text{eff}}}{\partial k^{(2)}}\right|_k = \rho(\alpha_3 - \alpha_2), \tag{C6}$$

therefore $k_{\text{eff}}$ ($R$) increases or decreases with $k^{(2)}$ solely depending on whether $\alpha_3$ is larger or smaller than $\alpha_2$, respectively.

### 2. Assortativity

Here we study how the basic reproduction number is affected by the way in which assortativity is distributed between 2- and 3-edges. No correlation is assumed between type and degree. Fixed the average pairwise mixing parameter, $\alpha = (\alpha_2 k^{(2)} + 2\alpha_3 k^{(3)})/k$, with $k = k^{(2)} + 2k^{(3)}$, we want to establish the sign of $\partial R/\partial \alpha_2|_\alpha = (\lambda/\mu)\partial k_{\text{eff}}/\partial \alpha_2|_\alpha$

*a. Easy adaptation.* Here, $r_N = 1 - \varepsilon_{\text{out}}$ and $s_N = 1 - \varepsilon_{\text{in}}$ for $1 \leqslant n_N \leqslant n - 1$, and $r_N = s_N = 1$ for $n_N = n$. After some algebra, one gets to

$$\left.\frac{\partial k_{\text{eff}}}{\partial \alpha_2}\right|_\alpha = \frac{k^{(2)}}{2}\rho(1 - \beta_{3,1}\rho)(\varepsilon_{\text{out}} + \varepsilon_{\text{in}} - \varepsilon_{\text{out}}\varepsilon_{\text{in}})$$

$$\times \frac{C + \{C^2 + [2k\alpha(1 - \varepsilon_{\text{out}})(1 - \varepsilon_{\text{in}})]^2\rho(1 - \rho)\}^{\frac{1}{2}}}{\{C^2 + [2k\alpha(1 - \varepsilon_{\text{out}})(1 - \varepsilon_{\text{in}})]^2\rho(1 - \rho)\}^{\frac{1}{2}}} \tag{C7}$$

$$\geqslant \frac{k^{(2)}}{2}\rho(1 - \beta_{3,1}\rho)(\varepsilon_{\text{out}} + \varepsilon_{\text{in}} - \varepsilon_{\text{out}}\varepsilon_{\text{in}})$$

$$\times \frac{C + |C|}{\{C^2 + [2k\alpha(1 - \varepsilon_{\text{out}})(1 - \varepsilon_{\text{in}})]^2\rho(1 - \rho)\}^{\frac{1}{2}}}$$

$$\geqslant 0. \tag{C8}$$

Whatever the value of $C$, which is a function of the various parameters, since the first term in Eq. (C7) is non-negative [recall the condition $\beta_{3,1} \leqslant 1/\rho$, Eq. (28)] and the second term within the square root as well, immediately follow both inequalities. From Eq. (C7) we see that $\partial k_{\text{eff}}/\partial \alpha_2|_\alpha$ is strictly positive when at least one between $\varepsilon_{\text{out}}$ and $\varepsilon_{\text{in}}$ is nonzero, i.e., trivially, whenever there is protection. Additionally, we note that the dependence of $k_{\text{eff}}$ on $\alpha_2$ is reduced by making $1 - \beta_{3,1}\rho$ smaller. From Eqs. (15) and (21) we see that that means lowering the frequency of {A, N, N} triads, confirming them as the responsible for that dependence.

*b. Hard adaptation.* In this scenario, $r_N = 1 - \varepsilon_{\text{out}}$ and $s_N = 1 - \varepsilon_{\text{in}}$ for $n_N = 1$, and $r_N = s_N = 1$ otherwise. With a bit of algebra, one finds

$$\left.\frac{\partial k_{\text{eff}}}{\partial \alpha_2}\right|_\alpha \propto -k^{(2)}\rho(1 - \rho)(1 - \beta_{3,1}\rho)(1 - \varepsilon_{\text{out}})(1 - \varepsilon_{\text{in}})$$

$$\times \left[\varepsilon_{\text{out}} + \varepsilon_{\text{in}} - 2\varepsilon_{\text{out}}\varepsilon_{\text{in}}\right.$$

$$\times \left.\left(\beta_{3,1}\rho + (1 - \beta_{3,1}\rho)\frac{\alpha_2 k^{(2)}}{\alpha k}\right)\right]\alpha k \tag{C9}$$

$$\leqslant -k^{(2)}\rho(1 - \rho)(1 - \beta_{3,1}\rho)(1 - \varepsilon_{\text{out}})(1 - \varepsilon_{\text{in}})$$

$$\times (\varepsilon_{\text{out}} + \varepsilon_{\text{in}} - 2\varepsilon_{\text{out}}\varepsilon_{\text{in}})\alpha k \leqslant 0, \tag{C10}$$

where the constant of proportionality in Eq. (C9) is positive. The first inequality comes from the constraints $\alpha_2 \leqslant \alpha k/k^{(2)}$, ensuring $\alpha_3 \geqslant 0$ [combining Eqs. (27) and (29)], and $\beta_{3,1} \leqslant 1/\rho$ [Eq. (28)], so that the term in square brackets is a decreasing function of $\alpha_2$ and thus takes its minimum value at $\alpha_2 = \alpha k/k^{(2)}$. This yields the first inequality in Eq. (C10) and, since the remaining terms are all non-negative, we finally have $\partial k_{\text{eff}}/\partial \alpha_2|_\alpha \leqslant 0$. Comparing with the previous case, we see that the sign of $\partial k_{\text{eff}}/\partial \alpha_2|_\alpha$ is thus dictated by whether the adaptation by the N-type individuals is easy or hard. From Eq. (C9) we note that, in this scenario, the derivative goes to zero for either full out-going ($\varepsilon_{\text{out}} = 1$) or in-going ($\varepsilon_{\text{in}} = 1$) protection (besides $\varepsilon_{\text{out}} = \varepsilon_{\text{in}} = 0$). In such cases, secondary infections are exclusively generated by N-types but, in absence of indirect modifications, their state is unaffected by how assortativity is distributed among group sizes. In the end, as before, lowering $1 - \beta_{3,1}\rho$ reduces the dependence of $k_{\text{eff}}$ on $\alpha_2$, confirming the role of the {A, N, N} triads.

### APPENDIX D: APPROXIMATE DYNAMICS FOR HIGH PROPHYLACTIC EFFICACY

Let us assume that the probability of adoption is low in type-N uniform groups and high for type A in any group. Considering first the limit of high out-going prophylactic efficacy, i.e., $1 - \varepsilon_{\text{out}} \ll 1$, we have $\theta_{A \to A}, \theta_{A \to N} \approx O(1 - \varepsilon_{\text{out}})$, hence $\theta_{A \to A} \ll \theta_{N \to A}$ and $\theta_{A \to N} \ll \theta_{N \to N}$. In such a regime, Eqs. (33) and (34) approximate to

$$\dot{I}_A \approx -\mu I_A + \lambda(1 - I_A)I_N \theta_{N \to A} y, \tag{D1}$$

$$\dot{I}_N \approx -\mu I_N + \lambda(1 - I_N)I_N \theta_{N \to N}. \tag{D2}$$

Equation (D2) is a standard SIS dynamics for $I_N$, hence its nonzero fixed point (stable for $\lambda\theta_{N \to N} \geqslant \mu$) reads

$$I_N^\star = 1 - \frac{\mu}{\lambda\theta_{N \to N}} \approx 1 - \frac{1}{R}, \tag{D3}$$

where the approximation comes from $k_{\text{eff}} \approx \theta_{N \to N}$ [see Eq. (41)]. As a consequence, $I_N^\star$ and $R$ share the same dependence on the various parameters. The nonzero fixed point for $I_A$ is then

$$I_A^\star = \frac{1}{1 + \frac{\mu}{\lambda\theta_{N \to A}I_N^\star}} \approx \frac{\lambda}{\mu}\theta_{N \to A}I_N^\star. \tag{D4}$$

In particular, if also the in-going efficacy is high, then Eq. (D4) implies $I_A^\star \ll I_N^\star$. From Eqs. (D3) and (D4), is easy to see how $I_N^\star$ and $I_A^\star$ can show different shapes. For instance, everything else left unchanged, an overall increased homophily yields a higher $\theta_{N \to N}$ but a lower $\theta_{N \to A}$, therefore $I_N^\star$ grows (as $R$ does) but $I_A^\star$ can either increase or decrease depending on whether the rate at which $\theta_{N \to A}$ decreases is lower or higher than that at which $I_N^\star$ (growing with $\theta_{N \to N}$) increases. In other words, the reproduction number may not be informative about the endemic equilibrium for type A—not even qualitatively, although it generally may for type N.

Considering now the limit of high in-going prophylactic efficacy, i.e., $1 - \varepsilon_{\text{in}} \ll 1$, we get $\theta_{A \to A}, \theta_{N \to A} \approx O(1 - \varepsilon_{\text{in}})$, implying that $I_A^\star \approx O(1 - \varepsilon_{\text{in}})$ is the only solution. In turn, the latter yields $I_N^\star$ as given—again—by Eq. (D3), the solution of Eq. (D2).
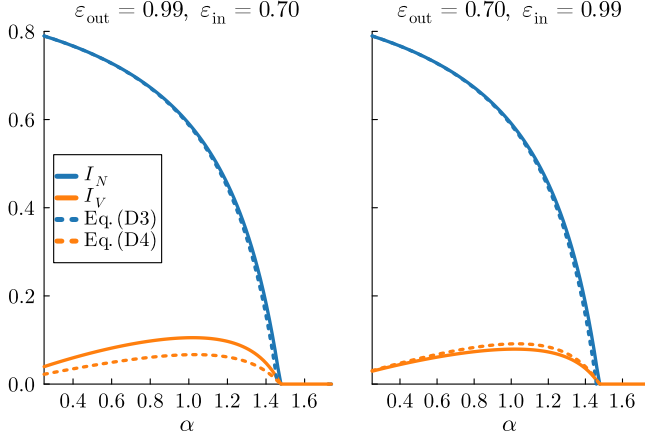
FIG. 5. Exact versus approximated prevalence as given by Eqs. (D3) and (D4) (dotted). Here, $\alpha_2 = \alpha_3 = \alpha$, $\beta_{3,1} = 1.0$, $\rho = 0.45$, $k = k^{(2)} + 2k^{(3)} = 6$, and $\lambda/\mu = 1.0$.

Figure 5 gives an example of how the approximation performs for either high out-going or in-going efficacy. It is less accurate in the former case for type A because, not being the in-going efficacy high enough, the A-to-A infection rate is not so small as assumed. In any case, the approximation preserves the qualitative behavior.

## APPENDIX E: NUMERICAL SIMULATION OF THE MICROSCOPIC SPREADING DYNAMICS

The microscopic SIS spreading dynamics is simulated as a discrete-time Markov process. Letting $\Delta t \ll 1$ be the time-step duration, the local state of the nodes is updated synchronously at the discrete times $t = 0, \Delta t, 2\Delta t, \ldots$, up to a maximum time step. At each step, every node interacts within each of the $k_i^{(n)}$ $n$-edges including it. The contact structure is encoded in the adjacency tensors $\{A^{(n)}\}$ such that $A_{i_1,\ldots,i_n}^{(n)} = 1$ if nodes $i_1, \ldots, i_n$ form a $n$-edge and 0 otherwise. In the following we consider $n = 2, 3$. Denoting with $\sigma_i(t) \in \{0, 1\}$ the binary variable representing the state of node $i \in \{1, \ldots, N\}$ at time $t$ [$\sigma_i(t) = 0$ if susceptible, $\sigma_i(t) = 1$ if infected], we need to compute the transition probabilities of infection, $P(\sigma_i(t + \Delta t) = 1|\sigma_i(t) = 0)$, and recovery, $P(\sigma_i(t + \Delta t) = 0|\sigma_i(t) = 1)$. Finally, let $X_i \in \{A, N\}$ be the type of node $i$.

If the interaction occurs in a 2-edge, we have four possible values for the transmission probability, depending on the types of the two involved nodes, say $i$ and $j$. If $i$ is the susceptible node and $j$ the infected one, then the transmission rate is $\lambda r_{X_j}^{(1,1)} s_{X_i}^{(1,1)}$ if $X_i = X_j$ and $\lambda r_{X_j}^{(0,1)} s_{X_i}^{(0,1)}$ if $X_i \neq X_j$. Accordingly, the transmission probability from $j$ to $i$ during the time interval $[t, t + \Delta t]$ is, respectively, $\lambda r_{X_j}^{(1,1)} s_{X_i}^{(1,1)} \Delta t$ and $\lambda r_{X_j}^{(0,1)} s_{X_i}^{(0,1)} \Delta t$. The out-going and in-going probabilities, $r$ and $s$, are given in Eqs. (42) and (43) and are computed at the equilibrium of the adoption dynamics [i.e., using the fixed

point of Eqs. (31) and (32)]. The probability $w_i^{(2)}(t)$ that $i$ does not get infected via any of the 2-edges incident on it thus reads

$$w_i^{(2)}(t) = \prod_{j=1}^{N} A_{i,j}^{(2)}\left[1 - \sigma_j(t)\lambda\tau_{i,j}^{(2)}\Delta t\right], \quad \text{(E1)}$$

where

$$
\begin{aligned}
\tau_{i,j}^{(2)} &\equiv \tau^{(2)}(X_i, X_j) \\
&= \delta_{X_i,X_j} r_{X_j}^{(1,1)} s_{X_i}^{(1,1)} + (1 - \delta_{X_i,X_j}) r_{X_j}^{(0,1)} s_{X_i}^{(0,1)}, \quad \text{(E2)}
\end{aligned}
$$

with $\delta_{x,y}$ being the Kronecker delta.

If the interaction is part of a 3-edge, the transmission probability can take eight different values, depending on the type of all the three nodes, say $i$, $j$, and $l$. Focusing on the interaction between $i$ and $j$, if the former is susceptible and the other one is infected, we have $\lambda r_{X_j}^{(2,2)} s_{X_i}^{(2,2)}$ if $X_i = X_j = X_l$; $\lambda r_{X_j}^{(1,2)} s_{X_i}^{(1,2)}$ if $X_i = X_j \neq X_l$; $\lambda r_{X_j}^{(1,2)} s_{X_i}^{(0,2)}$ if $X_i \neq X_j$ and $X_j = X_l$; and $\lambda r_{X_j}^{(0,2)} s_{X_i}^{(1,2)}$ if $X_i \neq X_j$ and $X_i = X_l$. Then, the probability $w_i^{(3)}(t)$ that $i$ does not get infected within any of the 3-edges it takes part to reads

$$
\begin{aligned}
w_i^{(3)}(t) = \prod_{j=1}^{N} \prod_{l>j}^{N} A_{i,j,l}^{(3)} &\left[1 - \sigma_j(t)\lambda\tau_{i,j;l}^{(3)}\Delta t\right] \\
&\times \left[1 - \sigma_l(t)\lambda\tau_{i,l;j}^{(3)}\Delta t\right], \quad \text{(E3)}
\end{aligned}
$$

where

$$
\begin{aligned}
\tau_{i,j;l}^{(3)} &\equiv \tau^{(3)}(X_i, X_j; X_l) \\
&= \delta_{X_i,X_j}\left[\delta_{X_i,X_l} r_{X_j}^{(2,2)} s_{X_i}^{(2,2)} + (1 - \delta_{X_i,X_l}) r_{X_j}^{(1,2)} s_{X_i}^{(1,2)}\right] \\
&\quad + \left(1 - \delta_{X_i,X_j}\right) \\
&\quad \times \left[\delta_{X_i,X_l} r_{X_j}^{(0,2)} s_{X_i}^{(1,2)} + (1 - \delta_{X_i,X_l}) r_{X_j}^{(1,2)} s_{X_i}^{(0,2)}\right]. \quad \text{(E4)}
\end{aligned}
$$

All in all, the probability that a susceptible node $i$ gets infected via at least one of its contacts is

$$P(\sigma_i(t + \Delta t) = 1|\sigma_i(t) = 0) = 1 - w_i^{(2)}(t)w_i^{(3)}(t), \quad \text{(E5)}$$

while its probability of recovery if infected is simply

$$P(\sigma_i(t + \Delta t) = 0|\sigma_i(t) = 1) = \mu\Delta t. \quad \text{(E6)}$$

The stochastic process is implemented by drawing a value for $u_i \sim \text{Uniform}(0, 1)$ independently for each node $i$, and at each time step. If node $i$ is susceptible, it gets infected if $P(\sigma_i(t + \Delta t) = 1|\sigma_i(t) = 0) > u_i$; if infected, it recovers if $P(\sigma_i(t + \Delta t) = 0|\sigma_i(t) = 1) > u_i$.

The continuous dynamics corresponds to the limit $\Delta t \to 0$, as only the terms linear in $\Delta t$ survive. In the simulation, we took $\Delta t = 0.05$. Using the quasistationary state (QS) method (with 50 stored active states and a 25% probability of update) [47], we let the system run for a transient of 1000 time steps and then averaged over the last 500 time steps.

[1] J. P. Gleeson, Binary-State Dynamics on Complex Networks: Pair Approximation and Beyond, Phys. Rev. X **3**, 021004 (2013).

[2] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Networks beyond pairwise interactions: Structure and dynamics, Phys. Rep. **874**, 1 (2020).

[3] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, What are higher-order networks? arXiv:2104.11329.

[4] J. T. Wootton, Indirect effects and habitat use in an intertidal community: Interaction chains and interaction modifications, Am. Nat. **141**, 71 (1993).

[5] T. Okuyama and B. M. Bolker, Model-based, response-surface approaches to quantifying indirect interactions, in *Ecology and Evolution of Trait-Mediated Indirect Interactions: Linking Evolution, Community, and Ecosystem* (Cambridge University Press, Cambridge, 2012), p. 186.

[6] J. M. Levine, J. Bascompte, P. B. Adler, and S. Allesina, Beyond pairwise mechanisms of species coexistence in complex communities, Nature (London) **546**, 56 (2017).

[7] R. Sahasrabuddhe, L. Neuhäuser, and R. Lambiotte, Modelling non-linear consensus dynamics on hypergraphs, J. Phys.: Complex. **2**, 025006 (2021).

[8] E. Glaeser and C. R. Sunstein, Does more speech correct falsehoods? J. Leg. Stud. **43**, 65 (2014).

[9] R. V. Solé, B. Corominas-Murtra, and J. Fortuny, Diversity, competition, extinction: The ecophysics of language change, J. R. Soc. Interface **7**, 1647 (2010).

[10] Á. Bodó, G. Y. Katona, and P. L. Simon, Sis epidemic propagation on hypergraphs, Bull. Math. Biol. **78**, 713 (2016).

[11] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, Master equation analysis of mesoscopic localization in contagion dynamics on higher-order networks, Phys. Rev. E **103**, 032301 (2021).

[12] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, Social Confinement and Mesoscopic Localization of Epidemics on Networks, Phys. Rev. Lett. **126**, 098301 (2021).

[13] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, Universal Nonlinear Infection Kernel from Heterogeneous Exposure on Higher-Order Networks, Phys. Rev. Lett. **127**, 158301 (2021).

[14] A. Lai, C. Poon, and A. Cheung, Effectiveness of facemasks to reduce exposure hazards for airborne infections among general populations, J. R. Soc. Interface **9**, 938 (2012).

[15] S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, and A. B. Gumel, To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic, Infect. Dis. Model. **5**, 293 (2020).

[16] C. J. Worby and H.-H. Chang, Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic, Nat. Commun. **11**, 4049 (2020).

[17] A. Catching, S. Capponi, M. T. Yeh, S. Bianco, and R. Andino, Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of COVID-19, Sci. Rep. **11**, 15998 (2021).

[18] J. Bedson, L. A. Skrip, D. Pedi, S. Abramowitz, S. Carter, M. F. Jalloh, S. Funk, N. Gobat, T. Giles-Vernick, G. Chowell *et al.*, A review and agenda for integrated disease models including social and behavioural factors, Nat. Hum. Behav. **5**, 834 (2021).

[19] Z. Wang, M. A. Andrews, Z.-X. Wu, L. Wang, and C. T. Bauch, Coupled disease–behavior dynamics on complex networks: A review, Phys. Life Rev. **15**, 1 (2015).

[20] S. E. Asch, Opinions and social pressure, Sci. Am. **193**, 31 (1955).

[21] J. J. V. Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman *et al.*, Using social and behavioural science to support COVID-19 pandemic response, Nat. Hum. Behav. **4**, 460 (2020).

[22] S. E. Bokemper, M. Cucciniello, T. Rotesi, P. Pin, A. A. Malik, K. Willebrand, E. E. Paintsil, S. B. Omer, G. A. Huber, and A. Melegaro, Beliefs about mask efficacy and the effect of social norms on mask wearing intentions for COVID-19 risk reduction, PLoS One **16**, e0258282 (2021).

[23] D. Centola, The spread of behavior in an online social network experiment, Science **329**, 1194 (2010).

[24] D. Centola, An experimental study of homophily in the adoption of health behavior, Science **334**, 1269 (2011).

[25] J. Barceló and G. C.-H. Sheen, Voluntary adoption of social welfare-enhancing behavior: Mask-wearing in spain during the COVID-19 outbreak, PLoS One **15**, e0242764 (2020).

[26] M. McPherson, L. Smith-Lovin, and J. M. Cook, Birds of a feather: Homophily in social networks, Annu. Rev. Sociol. **27**, 415 (2001).

[27] G. Burgio, B. Steinegger, G. Rapisardi, and A. Arenas, Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics, Phys. Rev. Res. **3**, 033128 (2021).

[28] G. Burgio, B. Steinegger, and A. Arenas, Homophily impacts the success of vaccine roll-outs, Commun. Phys. **5**, 70 (2022).

[29] H. Watanabe and T. Hasegawa, Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks, Physica A (Amsterdam, Neth.) **603**, 127760 (2022).

[30] A. K. Rizi, A. Faqeeh, A. Badie-Modiri, and M. Kivelä, Epidemic spreading and digital contact tracing: Effects of heterogeneous mixing and quarantine failures, Phys. Rev. E **105**, 044313 (2022).

[31] T. Hiraoka, A. K. Rizi, M. Kivelä, and J. Saramäki, Herd immunity and epidemic size in networks with vaccination homophily, Phys. Rev. E **105**, L052301 (2022).

[32] A. Bretto, *Hypergraph Theory, An Introduction* (Springer, 2013).

[33] N. Veldt, A. R. Benson, and J. Kleinberg, Higher-order homophily is combinatorially impossible, Sci. Adv. **9**, eabq3200 (2023).

[34] N. Eisenberg, R. A. Fabes, and T. L. Spinrad, *Prosocial Development* (John Wiley & Sons, Inc., New Jersey, 2006).

[35] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis, Infectious disease modeling of social contagion in networks, PLoS Comput. Biol. **6**, e1000968 (2010).

[36] L. Hébert-Dufresne, T. M. Waring, G. St-Onge, M. T. Niles, L. Kati Corlew, M. P. Dube, S. J. Miller, N. J. Gotelli, and B. J. McGill, Source-sink behavioural dynamics limit institutional evolution in a group-structured society, R. Soc. Open Sci. **9**, 211743 (2022).

[37] M. H. DeGroot, Reaching a consensus, J. Am. Stat. Assoc. **69**, 118 (1974).

[38] O. Diekmann, J. Heesterbeek, and M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, J. R. Soc. Interface **7**, 873 (2010).

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.107.064304 for more details about the social contagion model and for further analysis of the results obtained in the binary models.

[40] G. Burgio, A. Arenas, S. Gómez, and J. T. Matamalas, Network clique cover approximation to analyze complex contagions through group interactions, Commun. Phys. **4**, 111 (2021).

[41] R. J. Sorensen, R. M. Barber, D. M. Pigott, A. Carter, C. N. Spencer, S. M. Ostroff, R. C. Reiner Jr, C. Abbafati, C. Adolph, A. Allorant *et al.*, Variation in the COVID-19 infection-fatality ratio by age, time, and geography during the pre-vaccine era: A systematic analysis, Lancet **399**, 1469 (2022).

[42] V. Cozza, H. Campbell, H. H. Chang, A. D. Iuliano, J. Paget, N. N. Patel, R. C. Reiner, C. Troeger, C. Viboud, J. S. Bresee *et al.*, Global seasonal influenza mortality estimates: A comparison of 3 different approaches, Am. J. Epidemiol. **190**, 718 (2021).

[43] E. Badillo-Goicoechea, T.-H. Chang, E. Kim, S. LaRocca, K. Morris, X. Deng, S. Chiu, A. Bradford, A. Garcia, C. Kern *et al.*, Global trends and predictors of face mask usage during the COVID-19 pandemic, BMC Public Health **21**, 2099 (2021).

[44] M. H. Haischer, R. Beilfuss, M. R. Hart, L. Opielinski, D. Wrucke, G. Zirgaitis, T. D. Uhrich, and S. K. Hunter, Who is wearing a mask? gender-, age-, and location-related differences during the COVID-19 pandemic, PLoS One **15**, e0240785 (2020).

[45] R. Sánchez-Arenas, S. V. Doubova, M. A. González-Pérez, and R. Pérez-Cuevas, Factors associated with COVID-19 preventive health behaviors among the general public in Mexico City and the State of Mexico, PLoS One **16**, e0254435 (2021).

[46] W. Kaewpan, K. Rojpaisarnkit, S. Pengpid, and K. Peltzer, Factors affecting face mask-wearing behaviors to prevent COVID-19 among Thai people: A binary logistic regression model, Front. Psychol. **13**, 996189 (2022).

[47] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results, Phys. Rev. E **86**, 041125 (2012).