Social confinement and mesoscopic localization of epidemics on networks

Guillaume St-Onge,[1,2] Vincent Thibeault,[1,2] Antoine Allard,[1,2] Louis J. Dubé,[1,2] and Laurent Hébert-Dufresne[1,3]

[1]*Département de physique, de génie physique et d'optique,*
Université Laval, Québec (Québec), Canada G1V 0A6
[2]*Centre interdisciplinaire en modélisation mathématique,*
Université Laval, Québec (Québec), Canada G1V 0A6
[3]*Department of Computer Science & Vermont Complex Systems Center, University of Vermont, Burlington, VT 05405*

Recommendations around epidemics tend to focus on individual behaviors, with much less efforts attempting to guide event cancellations and other collective behaviors since most models lack the higher-order structure necessary to describe large gatherings. Through a higher-order description of contagions on networks, we model the impact of a blanket cancellation of events larger than a critical size and find that epidemics can suddenly collapse when interventions operate over groups of individuals rather than at the level of individuals. We relate this phenomenon to the onset of mesoscopic localization, where contagions concentrate around dominant groups.

Standard disease models reduce the complexity of epidemics to simple processes that provide useful insights. In fact, many of the key results of these models provide the foundation for our current understanding and forecasting of novel emerging epidemics [1–3]. The reduction of the complex to the simple is perhaps best embodied by the *mass-action approximation* [4]. This assumption essentially means that we are considering a randomly mixed population, ignoring household structures, social gatherings, and the different behaviors of different individuals. Mass-action models are thus seriously limited, since they focus on the average number of infections caused by each case, the basic reproduction number R_0 [5], and ignore the underlying heterogeneity [6]. There are also conceptual issues with the design of targeted interventions when relying on the mass-action assumption. Where should we target our interventions, and what should be their impact? In this letter, we address these questions through higher-order contact patterns.

Network science provides a natural framework to go beyond the mass-action approximation by considering key features of the structure of contacts among individuals. The simplest generalization is perhaps the *heterogeneous pair approximation*—individuals are nodes categorized by their number of contacts and their state, and the contacts are distinguished by the states of the nodes involved [7]. At this level of sophistication, all pairwise contacts of a given state are still, a priori, equivalent.

One developing area in network science concerns dynamical processes on higher-order representation of networks, i.e., where the network is not simply a conglomerate of pairwise interactions but where interactions occur in a coordinated manner because of a higher-level organization (schools, households, events, etc.) [9]. For dynamics on higher-order networks, one straightforward generalization of the framework just described is the *heterogeneous clique approximation* [10]. Nodes are categorized by their state and *membership*, i.e., the number m of groups to which they belong. The groups are characterized by their size n, and the states of the nodes involved.

Let us consider contagion processes on a simple version of higher-order networks, see Fig. 1. The network is char-
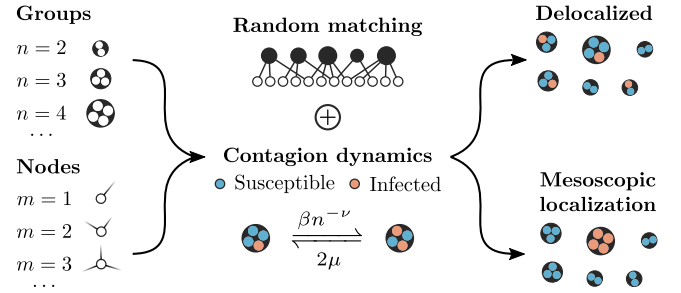
FIG. 1. **Framework for contagions on higher-order networks.** Nodes are assigned to m groups and groups are of various sizes n, distributed according to g_m and p_n. We consider a SIS dynamics where infected nodes transmit the disease in a group of size n at rate $\beta n^{-\nu}$ with $\nu \in [0, 1]$, and recover at rate μ. We characterize the phenomenon of mesoscopic localization, namely the concentration of infected nodes in large groups.

acterized by g_m, the distribution for the memberships m of nodes and p_n, the distribution for the sizes n of groups. We use different heterogeneous distributions for both of them, $g_m \propto m^{-\gamma_m}$ and $p_n \propto n^{-\gamma_n}$, with finite cut-offs $m_{\max}$ and $n_{\max}$ respectively.

For mathematical convenience, we use a Susceptible-Infected-Susceptible (SIS) dynamics. However, our results have repercussions for a much broader class of dynamical processes, including Susceptible-Infected-Recovered (SIR) dynamics where individuals develop immunity. Infected nodes transmit the disease to susceptible nodes belonging to a same group of size n at rate $\beta n^{-\nu}$ with $\nu \in [0, 1]$, and recover at rate μ. It is equivalent to a standard SIS model on networks formed of cliques [10, 11], but the edges have weights $n^{-\nu}$. We recover an unweighted network for $\nu = 0$. The parameter ν tunes the strength of interactions within groups, which we assume to be decreasing with size. For instance, an individual in a workplace typically interacts with more people than at home, but interactions in a household are stronger.

We track $s_m(t)$, the probability for a node of membership m to be susceptible, and $c_{n,i}(t)$, the probability to observe i infected nodes within a group of size n. Their dynamics are described by the following coupled ordinary differential equa-
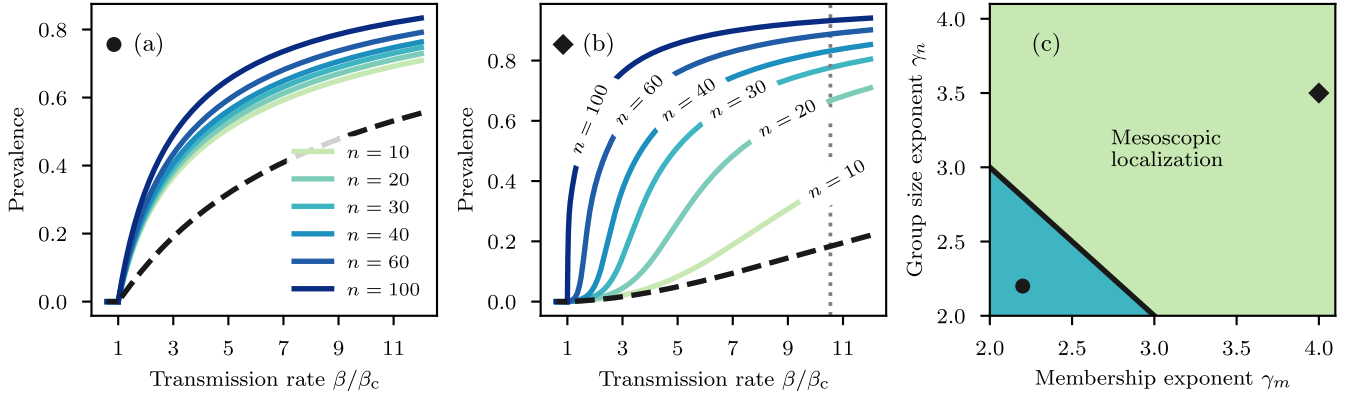
FIG. 2. **Epidemic localization in networks with heterogeneous higher-order structures.** We use $\nu = 0$ for all panels and power-law distributions for the memberships $g_m \propto m^{-\gamma_m}$ and group sizes $p_n \propto n^{-\gamma_n}$. (a)-(b) Solid lines represent the stationary group prevalence and dashed lines represent the stationary global prevalence. (a) For strongly coupled groups ($\gamma_m = \gamma_n = 2.2$), we find a collective phase transition. (b) For weakly coupled groups ($\gamma_m = 4, \gamma_n = 3.5$), we find a phenomenon of mesoscopic localization. While the global prevalence in the population can remain extremely low, larger groups can self-sustain the epidemic. The vertical dotted line is an estimation of the *delocalization* threshold, where the epidemic invades the whole network (see Ref. [8]). (c) Mesoscopic localization is the rule rather than the exception. The solid line, $\gamma_m + \gamma_n = 5$, separates the *delocalized regime* (blue area of strong group coupling) from the *mesoscopic localization regime* (green area of weak group coupling), obtained from Eqs. (8a-b). The circle and diamond markers correspond to the networks in panels (a)-(b).

tions [10],

$$\frac{\mathrm{d}s_m}{\mathrm{d}t} = \mu(1 - s_m) - mrs_m , \tag{1a}$$

$$\frac{\mathrm{d}c_{n,i}}{\mathrm{d}t} = \mu(i + 1)c_{n,i+1} - \mu i c_{n,i} + (n - i + 1)\{\beta n^{-\nu}(i - 1) + \rho\}c_{n,i-1}$$
$$- (n - i)\{\beta n^{-\nu}i + \rho\}c_{n,i} , \tag{1b}$$

with the mean fields $r(t)$ and $\rho(t)$ defined as

$$r(t) = \frac{\sum_{n,i}\beta n^{-\nu}i\,(n - i)c_{n,i}(t)p_n}{\sum_{n,i}(n - i)c_{n,i}(t)p_n} , \tag{2a}$$

$$\rho(t) = r(t)\frac{\sum_m(m - 1)\,ms_m(t)g_m}{\sum_m ms_m(t)g_m} . \tag{2b}$$

If we take a susceptible node and select a random group to which it belongs, $r(t)$ is the mean infection rate associated to that group. Now if we pick a susceptible node in a group, $\rho(t)$ is the mean infection rate received from all external groups (i.e., excluding the one we picked the node from). Without loss of generality, we set $\mu = 1$ hereafter.

An important feature of this framework is that Eq. (1b) is an *approximate master equation* : it describes the full range of possible states for groups of size n, while assuming a mean-field coupling between them. As we show in Ref. [8], the agreement with Monte-Carlo simulations is excellent. The global prevalence in the network—the average fraction of infected nodes—is then

$$I(t) = \sum_m [1 - s_m(t)]\,g_m , \tag{3}$$

and the group prevalence is

$$I_n(t) = \sum_i \frac{i}{n}\,c_{n,i}(t) . \tag{4}$$

In Fig. 2(a) and Fig. 2(b), we show the stationary prevalence (global and within groups) for two different networks, obtained using Eq. (1). As expected from standard models, there exists an epidemic threshold β_c for the transmission rate below which epidemics cannot be sustained (see Ref. [8] for an analytical expression of β_c). Above β_c, the disease-free equilibrium of the dynamics becomes unstable, driving the epidemic to invade the network.

What is less expected are the sequential local transitions observed in the second panel [Fig. 2(b)]. For any value of the transmission rate, the outbreak thrives only in groups above a certain size. The epidemic is *self-sustained* locally, and the global prevalence reaches its highest growth rate with β well above the epidemic threshold, a defining feature of *smeared* phase transitions [12, 13]. This is reminiscent of certain infections, such as the bacteria *C. difficile*, mainly found in hospitals with large susceptible populations in close contact [14].

To get some insights on the emergence of this localization phenomenon, we examine the stationary group prevalence, I_n^*, near the absorbing-state. Using a saddle-point approximation valid for large n, we obtain [8]

$$I_n^* \sim \begin{cases} \dfrac{1}{1 - \beta n^{1-\nu}} & \text{if } \beta < n^{\nu-1} \\[2ex] n^{1/2}\,(\beta n^{1-\nu})^n\,e^{-n+n^\nu/\beta} & \text{if } \beta \geq n^{\nu-1} . \end{cases} \tag{5}$$

For $\beta > n^{\nu-1}$, this implies $I_n^* = O\left(n^{1/2}e^{bn}\right)$ with $b > 0$. Therefore, if $\beta_c \to n_{\max}^{\nu-1}$, the group prevalence increases exponentially with n above the epidemic threshold, and the outbreak is *localized* in large groups, as observed in Fig. 2(b). In other words, the behavior of the epidemic threshold dictates whether or not localization is possible for a given network organization.

In Ref. [8], we show that for power-law distributions of

membership and group size, we have the following behavior

$$\beta_c^{-1} \sim \Omega(g_m, p_n; \nu) + n_{\max}^{1-\nu} , \qquad (6)$$

where we define the *coupling* between groups as

$$\Omega(g_m, p_n; \nu) = \left(\frac{\langle m(m-1) \rangle}{\langle m \rangle} \right) \left(\frac{\langle n^{1-\nu}(n-1) \rangle}{\langle n \rangle} \right) . \qquad (7)$$

Asymptotic analysis of Eqs. (6) and (7) in the limit $n_{\max} \to \infty$ reveals the conditions for which $\beta_c \to n_{\max}^{\nu-1}$, i.e., the conditions necessary for a localized epidemic [8]. They require $\nu < 1$, and are always met whenever $\gamma_m \geq 3$. If $2 < \gamma_m < 3$, the conditions are then satisfied only if

$$2 < \gamma_n + \nu < 3 \ \text{ and } \ 3 - \gamma_n + \alpha(3 - \gamma_m) < 1 , \quad (8a)$$

or

$$\gamma_n + \nu \geq 3 \ \text{ and } \ \alpha(3 - \gamma_m) + \nu < 1 , \qquad (8b)$$

where $\alpha \geq 0$ relates the two cut-offs $m_{\max} \sim n_{\max}^{\alpha}$. These conditions define *mesoscopic localization* and distinguish the *localized* regime from the *delocalized* regime. We give some examples with $\nu = 0$ and $\alpha = 1$ in Fig. 2(c).

More intuitively, the conditions (8a-b) can be interpreted as the result of a competition between a collective and a local activation [15]. All pairs of (γ_m, γ_n) below the dark dividing line in Fig. 2(c) are associated with a strong group coupling ($\Omega \gg n_{\max}^{1-\nu}$), whereas pairs above the line correspond to a weak group coupling ($\Omega \ll n_{\max}^{1-\nu}$).

One important observation is that a large fraction of the structural parameter space (γ_m, γ_n) corresponds to the mesoscopic localization regime [green region in Fig. 2(c)], making it the rule rather than the exception. Moreover, the delocalized regime [blue region in Fig. 2(c)] is the parameter sub-space where the underlying networks are dense, i.e, where the average number of contacts of a node, proportional to $\langle n(n-1) \rangle$, diverges in the asymptotic limit. Since real-world networks are generally sparse [16, 17], it is reasonable to assume mesoscopic localization may occur in many real-world networks with a higher-level organization. The results of Fig. 2 extend nicely to cases with $\nu > 0$ (see Ref. [8], Appendix E) where the localized regime still dominates in a large portion the structural parameter space.

It is worth noting that we only observe localization for a certain portion $\beta \in [\beta_c, \beta^*]$ of the bifurcation diagram, where β^* is the *delocalization threshold* [dotted line in Fig. 2(b)]. In Ref. [8], we show how to estimate β^* and discuss other details of the localization regimes, notably the effects of finite cut-offs.

Perhaps most surprising about mesoscopic localization is how strong the effect can be. Even at low overall prevalence, we observe intense but local outbreaks in large groups. This simple observation justifies targeted interventions on these groups, analogous to school closures and the cancellation of large social or professional events. While such closures may seem excessive given the low prevalence found in the general
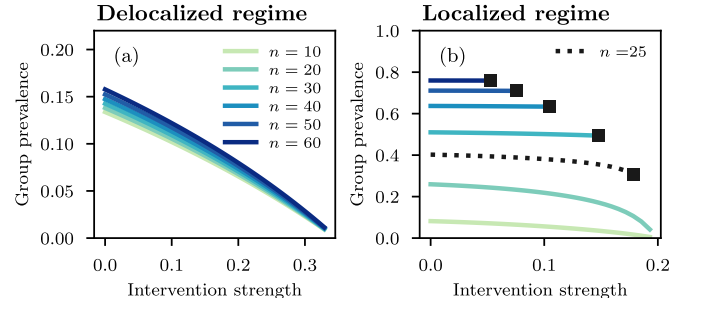
FIG. 3. **Local impact of structural interventions in delocalized and localized epidemics.** (a)-(b) Prevalence within groups of size n against the intervention strength [Eq. (9)]. We use the networks from Fig 2(a-b) with transmission rates adjusted to have similar global prevalence for both regimes without intervention. (a) In the delocalized regime, using $\beta \approx 0.0041$, we find a similar benefit of the intervention among all groups. (b) In the mesoscopic localization regime, using $\beta = 0.07$, we find a different story. Large groups that have not been removed by the intervention are barely affected, until the intervention is strong enough to cause a global collapse of the epidemic. Square markers indicate when groups of a particular size are removed.

population, these compact organized groups are where most infections will occur.

We now focus on targeted interventions on large groups, modeled by simply forcing a hard cut-off $n_{\max}$ on the distribution p_n. By using a cut-off instead of immunizing the groups, we preserve the membership m of nodes; they will simply belong to groups of smaller sizes.

To compare the effectiveness of interventions across different networks and parametrizations, we define an *intervention strength* (IS) as the fraction of the total edge weights that have been removed. If p_n and $\tilde{p}_n$ are respectively the group size distribution before and after the intervention, then

$$\text{IS} = 1 - \frac{\sum_n n^{1-\nu}(n-1)\tilde{p}_n}{\sum_n n^{1-\nu}(n-1)p_n} . \qquad (9)$$

In Fig. 3, we show the local impact on the group prevalence for such structural interventions. For networks in the delocalized regime [Fig. 3(a)], the intervention appears to reduce the risk of infections at the individual (node) level. As we decrease $n_{\max}$ (and therefore increase the intervention strength), the local prevalence within all groups decreases gradually and homogeneously until an epidemic threshold is reached. This is similar to traditional models where interventions reduce R_0 in a distributed, mass-action way.

In the localized regime [Fig. 3(b)], the intervention has a very different impact. Individuals that would have interacted in groups of size greater than $n_{\max}$ are spared by the intervention, but the large groups of sizes below $n_{\max}$ appear unaffected by the intervention. The main point here is that a local outbreak in certain organized gatherings (e.g. mass transit in urban centers, cruise ships) can persist despite interventions elsewhere. However, once the intervention is strong enough [roughly when $n_{\max}$ is below 25 in Fig. 3(b)], further interventions cause a rapid collapse of the epidemic.
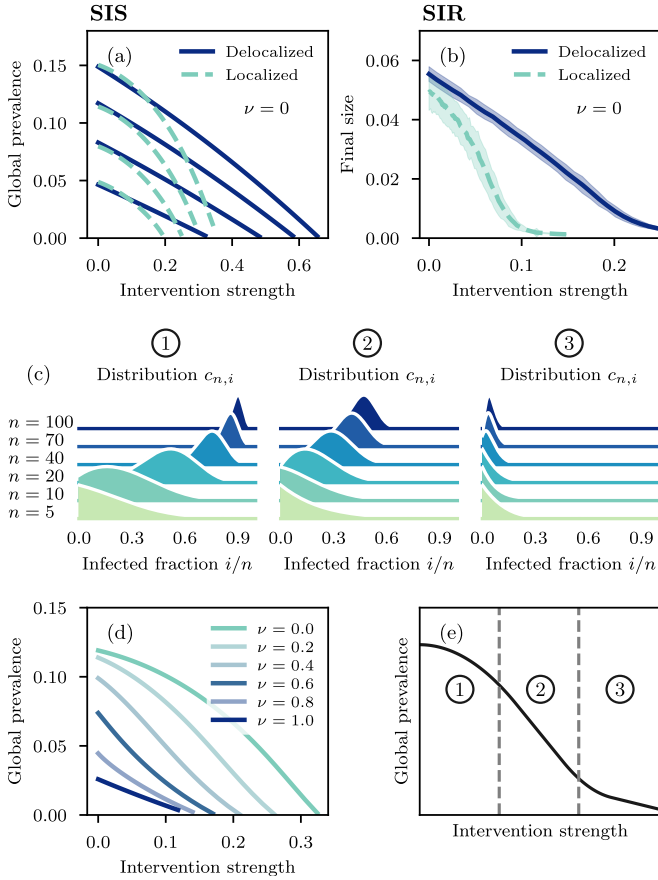
FIG. 4. **Global impact of structural interventions in different localization regimes of SIS and SIR dynamics.** (a) We compare the global prevalence of epidemics against the intervention strength [Eq. (9)] using the same networks as in Fig. 3. We use $\beta \in \{0.004, 0.005, 0.006, 0.007\}$ and $\beta \in \{0.07, 0.09, 0.11, 0.13\}$ for the delocalized and localized regimes. (b) A similar behavior is observed for the final size (fraction of recovered nodes at $t \to \infty$) in simulations of the SIR dynamics [18]. We generate 1000 networks of size $N = 10^6$ and run 10^4 SIR simulations on each network, starting with a single random infected node, with $\beta = 0.0036$ in the delocalized regime and $\beta = 0.083$ in the localized regime. We only kept macroscopic outbreaks, with a final size over 10^{-3}. Solid and dashed lines are sample means and the shaded regions represent twice the standard deviation. (c) Three examples of $c_{n,i}$ distributions for (1) strongly localized, (2) weakly localized, and (3) delocalized epidemics, using ν equal 0, 0.6 and 1 respectively. We used networks with $\gamma_m = \gamma_n = 3.5$ and $\beta = 0.3/\langle k \rangle$, where $\langle k \rangle$ is the average weighted degree before interventions. (d) We look at the impact of interventions for intermediate localization regimes by varying ν with the same networks as in (c). (e) Schematic representation of how the impact of interventions varies as it passes through different localization regimes.

Let us now look at the global impact of these interventions in Fig. 4. It is important to note that while we have assumed a SIS dynamics, our results hold for a SIR dynamics as well [see the similarity of Fig. 4(a-b)], which is generally a more realistic model for epidemics. The reason is that smeared phase transitions occur for this type of process as well [13].

In Fig. 4(a-b), we observe that the global impact of interventions in the delocalized regime is again similar to mass-action models—the prevalence decreases approximately linearly with the intervention strength. In contrast, in the localized regime, the intervention has a rapid non-linear effect in reducing the global prevalence as its strength is ramped up.

In both Fig. 3(b) and Fig. 4(a-b), it is surprising that the removal of the largest groups does not produce the largest decrease in prevalence, since these are the ones in which we expect most nodes to be infected. Even less clear is what drives the sudden collapse and if we should expect this behavior for all localized epidemics.

To clarify the situation, we need to explore the regimes between the delocalized and the *strongly* localized regimes displayed in Fig. 3(b) and 4(a-b). We break down these regimes in three parts in Fig. 4(c) to simplify the discussion: (1) strongly localized, (2) weakly localized, and (3) delocalized. In the strongly localized regime (1), large groups have a high prevalence, and act as independent entities that are barely affected by interventions elsewhere [as in Fig. 3(b)]. In the weakly localized regime (2), the disease still thrives in large groups, but they are not isolated from one another—interventions in one group now affect the others as well. This is the regime where interventions are most effective. Finally, in the delocalized regime (3), groups act as a whole, but the infection does not thrive in any of them, leading to a lower effectiveness of targeted interventions.

The easiest way to interpolate between the two extremes is to tune the group interaction strength through ν. As shown in Fig. 4(d), the initial impact of interventions changes as ν is increased. With $\nu = 0.6$ for instance, the epidemic is weakly localized, and now the removal of the largest groups produce the largest decrease in prevalence. Further interventions become eventually less effective, as the intervention itself causes the epidemic to shift to a delocalized regime.

More generally, the non-linear decrease of the global prevalence as a function of the intervention strength is explained by transitions between different localization regimes, as illustrated in Fig. 4(e). The sudden collapse for interventions on strongly localized epidemics is thus the result of a shift to a weakly localized regime, where interventions become much more effective.

Altogether, the lesson from Fig. 3 and 4 is that, just as we take heterogeneity of individual risks into account when preferentially vaccinating individuals, we should take heterogeneity of group risks into account when designing interventions. While pathogens operate at the scale of individuals, epidemics themselves interact with our entire social network, which has a modular, hierarchical, higher-order structure. Since we expect real epidemics to experience localization effects, we should aim to leverage their sudden collapse when designing structural interventions, and account for this emergent feature in our models.

Over the last few years, dynamics on higher-order representation of networks have shown time and time again that intuition built from simpler models does not always hold in more

complex scenarios, with examples ranging from competitive dynamics [19] to social contagion [20]. Localization of dynamics over higher-order structures is yet another addition to this list. Our work thus paves the way for a more comprehensive analysis of structural interventions on networks.

[1] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics," Bull. Math. Biol. **53**, 33 (1991).

[2] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics. ii - the problem of endemicity," Bull. Math. Biol. **53**, 57 (1991).

[3] W. O. Kermack and A. G. McKendrick, "Contributions to the mathematical theory of epidemics. iii - further studies of the problem of endemicity," Bull. Math. Biol. **53**, 89 (1991).

[4] O. Diekmann, J. A. J. Metz, and J. A. P. Heesterbeek, "The legacy of Kermack and McKendrick," in *Epidemic Model. Their Struct. Relat. to Data*, edited by D. Mollison (Cambridge University Press, 1995) p. 95.

[5] R. M. Anderson, B. Anderson, and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1992).

[6] L. Hébert-Dufresne, B. M. Althouse, S. V. Scarpino, and A. Allard, "Beyond R_0: the importance of contact tracing when predicting epidemics," medRxiv (2020), 10.1101/2020.02.10.20021725.

[7] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," Rev. Mod. Phys. **87**, 925 (2015).

[8] 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," arXiv:2004.10203 (2020).

[9] 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. (2020), 10.1016/j.physrep.2020.05.004.

[10] L. Hébert-Dufresne, P.-A. Noël, Vincent Marceau, A. Allard, and L. J. Dubé, "Propagation dynamics on networks featuring complex topologies," Phys. Rev. E **82**, 036115 (2010).

[11] M. E. J. Newman, "Properties of highly clustered networks," Phys. Rev. E **68**, 026121 (2003).

[12] T. Vojta, "Rare region effects at classical, quantum and nonequilibrium phase transitions," J. Phys. A. Math. Gen. **39**, R143 (2006).

[13] L. Hébert-Dufresne and A. Allard, "Smeared phase transitions in percolation on real complex networks," Phys. Rev. Res. **1**, 013009 (2019).

[14] L. V. McFarland and W. E. Stamm, "Review of clostridium difficile—associated diseases," Am. J. Infect. Control **14**, 99 (1986).

[15] S. C. Ferreira, R. S. Sander, and R. Pastor-Satorras, "Collective versus hub activation of epidemic phases on networks," Phys. Rev. E **93**, 32314 (2016).

[16] C. I. Del Genio, T. Gross, and K. E. Bassler, "All scale-free networks are sparse," Phys. Rev. Lett. **107**, 178701 (2011).

[17] M. E. J. Newman, *Networks*, 2nd ed. (Oxford University Press, 2018).

[18] G. St-Onge, J.-G. Young, L. Hébert-Dufresne, and L. J. Dubé, "Efficient sampling of spreading processes on complex networks using a composition and rejection algorithm," Comput. Phys. Commun. , 30 (2019).

[19] J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina, "Higher-order interactions stabilize dynamics in competitive network models," Nature **548**, 210 (2017).

[20] I. Iacopini, G. Petri, A. Barrat, and V. Latora, "Simplicial models of social contagion," Nat. Commun. **10**, 1 (2019).