CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Contact-based model for epidemic spreading on temporal networks
Authors	Koher, Andreas;Lentz, Hartmut H. K;Gleeson, James P.;Hövel, Philipp
Publication date	2019-08-02
Original Citation	Koher, A., Lentz, H. H. K., Gleeson, J. P. and Hövel, P. (2019) 'Contact-Based Model for Epidemic Spreading on Temporal Networks', Physical Review X, 9(3), 031017. (20pp.) doi: 10.1103/PhysRevX.9.031017
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1103/PhysRevX.9.031017
Rights	© The Author(s) 2019. Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-16 11:39:32
Item downloaded from	https://hdl.handle.net/10468/9330

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Contact-Based Model for Epidemic Spreading on Temporal Networks

Andreas Koher*
Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany

Hartmut H. K. Lentz
Institute of Epidemiology, Friedrich-Loeffler-Institut, Südufer 10, 17493 Greifswald-Insel Riems, Germany

James P. Gleeson
MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland

Philipp Hövel
School of Mathematical Sciences, University College Cork, Western Road, Cork T12 XF64, Ireland and Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany

(Received 17 November 2018; revised manuscript received 31 March 2019; published 2 August 2019)

We present a contact-based model to study the spreading of epidemics by means of extending the dynamic message-passing approach to temporal networks. The shift in perspective from node- to edge-centric quantities enables accurate modeling of Markovian susceptible-infected-recovered outbreaks on time-varying trees, i.e., temporal networks with a loop-free underlying topology. On arbitrary graphs, the proposed contact-based model incorporates potential structural and temporal heterogeneities of the contact network and improves analytic estimations with respect to the individual-based (node-centric) approach at a low computational and conceptual cost. Within this new framework, we derive an analytical expression for the epidemic threshold on temporal networks and demonstrate the feasibility of this method on empirical data.

DOI: 10.1103/PhysRevX.9.031017

Subject Areas: Complex Systems, Interdisciplinary Physics, Nonlinear Dynamics

I. INTRODUCTION

Accurate models of disease progression are valuable tools for public health institutions as they enable detection of outbreak origins [1–4], assessment of epidemic risk and vulnerability [5–8], and containment of the spreading at an early stage [5,9]. Mitigation strategies can thus be evaluated and employed without the need to run a large number of Monte Carlo (MC) realizations.

A fundamental challenge to mathematical epidemiologists is the accurate determination of the critical parameters that separate local and global epidemic outbreaks [10–16]. To this end, the early Kermack-McKendrick model [17] divides a population according to the disease status into compartments of susceptible, infected, and recovered individuals with mass-action equations to determine the transitions between them. Since then, a wide range of improvements has been proposed, including the impact of stochasticity [18–20], non-Markovian dynamics [21–25], and, notably, heterogeneity in the contact structure [11,26–30].

In recent years, the availability of mobility and contact data with a high temporal resolution, so-called temporal networks, offers another opportunity to improve analytical predictions [31–36]. The timing of links between nodes matters, in particular, when the network evolves on a similar timescale as the spreading dynamics, which led to an increasing interest in the interplay between disease and network dynamics [37–42].

One approach to model the states of individual nodes in a network takes the corresponding probabilities directly as variables in a set of coupled dynamic equations [7,13,20,43–47]. We refer to this approach as the individual-based (IB) model, though it is sometimes also called the N-intertwined model [20] or quenched mean field [48,49]. However intuitive, the analytic predictability suffers from the simplifying assumption that epidemic states of adjacent nodes are independent.

*andreas.koher@tu-berlin.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Recently, a change from a node-centric to an edge-centric perspective has been discussed within different frameworks in order to overcome the inherent limitation of the IB model. These approaches include branching processes [50], message passing [23,51], belief propagation [4], and the edge-based compartmental model [39]. So far, however, edge-centric models are mostly limited to static topologies. It thus remains an open challenge to simultaneously account for topological and temporal properties of the underlying contact data and hence improve current predictions of the epidemic threshold [7,52–56].

In this paper, we generalize the dynamic message-passing approach for discrete-time Markovian susceptible-infected-recovered (SIR) spreading [51] to time-evolving networks and derive the epidemic threshold within this new framework. The proposed model takes an edge-centric perspective because the relevant dynamic equations are based on the set of edges. Furthermore, the framework integrates the complete temporal and topological information of the underlying network into the epidemic model. We refer to our approach as contact-based (CB) model and compare numerical predictions with the widely used IB model that takes a node-centric perspective. Within the CB framework, we then derive a new analytic expression of the epidemic threshold for temporal networks and show that the edge-centric approach improves existing results [7,20,43,44,52,53] at a low conceptual and numerical cost. The CB and IB models have been implemented in PYTHON with the source code available on Github [57].

The remainder of this paper is structured as follows: First, we summarize the conceptual framework in Sec. II and formulate in Sec. III the dynamic equations of the IB and CB models. Then, we derive the epidemic threshold for temporal networks within the CB framework in Sec. IV. We compare the edge- and node-centric approaches against MC simulations in Sec. V and close with a discussion in Sec. VI. Appendix A includes an extension to weighted contacts and heterogeneous epidemiological parameters. A network analysis of the German cattle-trade data is given in Appendix B. Further results and applications of the CB model are summarized in Appendix C.

II. CONCEPTUAL FRAMEWORK

We consider a temporal network $\mathcal{G} = [G(0), G(1), \ldots, G(T-1)]$ with N nodes and T snapshots sampled at a constant rate. Although both modeling frameworks can, in principle, account for contact weights that indicate the strength of a connection, we focus on unweighted networks for simplicity and refer to Appendix A for an extension of the model.

Emphasizing the important difference between temporal and static elements, we refer to contacts as time-stamped links $(t, k, l) \in \mathcal{C} \subset \mathcal{T} \times \mathcal{N} \times \mathcal{N}$, thereby denoting with $\mathcal{N}$, $\mathcal{T}$, and $\mathcal{C}$ the set of nodes, time stamps, and contacts, respectively. We further assume that every contact is of constant duration and equal to the sampling time of the temporal network. By edges, we refer to the corresponding static elements in the time-aggregated network. In other words, an edge $(k, l) \in \mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ exists if and only if at least one (temporal) contact is recorded between k and l. Here, we denote with $\mathcal{E}$ the set of edges. Moreover, we assume directed edges throughout the paper and represent an undirected contact as two reciprocal contacts. Following the convention in Ref. [23], we denote with $k \to l$ a directed edge from k to l, and we indicate edge-based quantities in a similar fashion.

As the stochastic process, we assume a discrete-time SIR model, where a node $l \in \mathcal{N}$ represents an individual that is either susceptible, infected, or recovered at a given time t with a corresponding probability $S_l(t)$, $I_l(t)$, and $R_l(t)$, respectively. A susceptible node that is in contact with an infected neighbor contracts the disease with a constant and uniform (per time step) probability β. Furthermore, we treat the transmission events from multiple infected neighbors as independent, and similarly, we interpret potential (integer) edge weights as independent infection attempts (see Appendix A). We do not account for secondary infections within one time step; i.e., only direct neighbors can be affected. Once infected, the individual recovers with a uniform and constant probability μ independently of the infection process and henceforth acquires a permanent immunity.

Concerning the contact data, we focus our numerical analysis first on a face-to-face interaction network between 100 conference participants [58]. This so-called proximity graph has a resolution of 20 s, and the observation time is limited to the first 24 h. If necessary, we extend the data set with a periodic boundary condition in time. The time-resolved contacts enable the study of spreading of airborne diseases as well as the propagation of ideas and rumors.

As an illustrative example, we present in Fig. 1 the time-dependent probability that a selected node in the proximity graph is either susceptible (yellow), infected (red), or recovered (green). The results are derived from 10^4 MC simulations

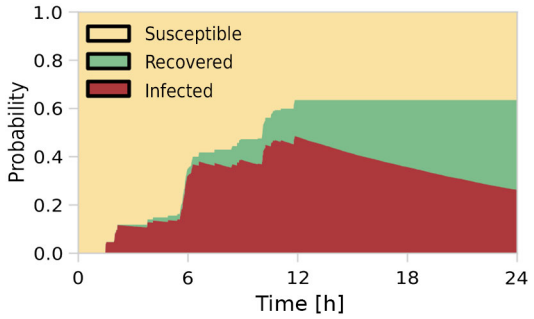


FIG. 1. Illustrative examples of a simulated epidemic outbreak from a single initially infected node. Colors give the probability that another arbitrarily selected node is in the susceptible (yellow), infected (red), or recovered (green) state, respectively. Simulation parameters: $\mu = 2.85 \times 10^{-4}$, $\beta = 100\mu$, 10^4 MC realizations.

with the same initially infected node. The trajectories reflect the bursty activity of the underlying temporal network [58] within the first 12 h and the subsequent inactive nighttime.

As a second source of data with direct relevance to public health, we consider an excerpt of the national German livestock database HI-Tier [59]. This temporal network comprises the movement of cattle between farms in Germany for the year 2010 with daily resolution. Within the observation window of 365 days, more than 3 million transactions have been recorded between over 180 000 farms and traders, respectively. For more details on the graph, see Appendix B. Cattle trade is considered an important transmission route for livestock-related diseases such as foot-and-mouth disease (FMD), which broke out in the United Kingdom in 2001 with an estimated cost of 8 billion pounds sterling [60]. Therefore, the analysis of the corresponding spatiotemporal graphs is highly relevant to public health institutions.

III. DYNAMIC EQUATIONS

In this section, we present the mathematical framework to model the stochastic SIR process as outlined in the Introduction and Sec. II. Our main focus is the CB model, but in order to facilitate a direct comparison between the node and the edge-based approach, we begin with a short overview of the IB model.

A. Individual-based model

In the IB model, the marginal probabilities $S_l(t)$, $I_l(t)$, and $R_l(t)$ for all $l \in \mathcal{N}$ directly enter a set of $3 \times N$ coupled dynamic equations. The probability for l to contract the infection from k upon a temporal contact is given by $\beta I_k(t)$. For convenience, we introduce an indicator function with $a_{k\to l}(t) = 1$ if a (directed) contact from k to l exists at time t and $a_{k\to l}(t) = 0$ otherwise. Then, the probability for node l to receive no infection at time t from any of its neighbors factorizes by assumption to $\prod_k [1 - \beta a_{k\to l}(t) I_k(t)]$ and $k \in \mathcal{N}$. With this result, the marginal probability $S_l(t+1)$ can be expressed by the probability $S_l(t)$ to be susceptible in the previous time step t and not contract the infection within the interval $[t, t+1)$. In the IB model, the joint probability factorizes, and we obtain

$$S_l(t+1) = S_l(t) \prod_{k\in\mathcal{N}} [1 - \beta a_{k\to l}(t) I_k(t)]. \tag{1}$$

Here, the crucial simplification is to treat the epidemic states of l and its neighbors as mutually independent, which is sometimes referred to as neglecting dynamic correlations [61].

The marginal probability $I_l(t+1)$ follows from two independent contributions: (i) The outflux $\mu I_l(t)$ indicates the transition from the infected to the recovered state. (ii) The influx $\Delta S_l(t) = S_l(t) - S_l(t+1)$ reflects the probability that node l is newly infected at time $t+1$. Combining both contributions leads to

$$I_l(t+1) = (1-\mu) I_l(t) + S_l(t) \left\{ 1 - \prod_{k\in\mathcal{N}} [1 - \beta a_{k\to l}(t) I_k(t)] \right\}. \tag{2}$$

The set of $2 \times N$ coupled dynamic equations (1) and (2) thus constitutes the IB model for temporal networks. The remaining marginal probability $R_l(t)$ to find node l in the recovered state follows from the conservation condition $S_l(t) + I_l(t) + R_l(t) = 1$ for all $l \in \mathcal{N}$. Finally, we assign a probability $z_l = S_l(0)$ that node l is initially susceptible, as well as $I_l(0) = 1 - z_l$ and $R_l(0) = 0$ throughout the paper.

Though intuitive and, in many cases, sufficient from a modeling perspective, the limits of the IB model are difficult to estimate due to the *ad hoc* factorization of the joint probability in Eq. (1). Even for the simplest network with two nodes connected by an undirected static edge, the IB approach can deviate significantly from the expected outcome as illustrated in Ref. [62]. In their example, recovery is neglected for simplicity, and only the first node is infected initially with some probability $0 < z_1 \leq 1$. Counterintuitively, the probabilities to find each node in the infected state converge to $I_1(\infty) = I_2(\infty) = 1$ according to the IB model, independent of the initial condition z_1. This convergence occurs because integrating Eqs. (1) and (2) admits a probability flux from the outbreak location to the adjacent node and back to its origin again. This mutual reinfection, coined the echo chamber effect in Ref. [62], appears because we neglect the fact that the probability I_2 to find the second node in the infected state is conditioned on the state of the first node, and thus the factorization in Eq. (1) is not justifiable.

In an arbitrary network, an initially infected node leads to a cascade of secondary infections within which all marginal probabilities are highly correlated. An accurate model excludes these previously infected nodes from those that can potentially contract the infection in the future. In the next section, we discuss how a shift from a node-centric to an edge-centric view can take into account some such dependencies.

B. Contact-based model

We begin with a slightly different approach to the marginal probability $S_l(t)$. First, we note that l is susceptible at time t, if it was susceptible initially [with probability $S_l(0) = z_l$] and has not contracted the infection from any of its neighbors up to time t. We assign the probability $\Phi_l(t)$ to the latter statement. Thus, without introducing any approximation at this stage, we can write

$$S_l(t) = z_l \Phi_l(t). \tag{3}$$

In order to determine $\Phi_l(t)$, we make the assumption that the underlying time-aggregated graph is a tree (ignoring directionality). Then, different branches originating in

node l are independent as long as l remains susceptible, and thus $\Phi_l(t)$ factorizes. However, if node l contracts a disease from a neighbor k with some probability and passes it on to another node k', then the corresponding probabilities I_k and $I_{k'}$ are clearly correlated. A simple solution that allows different branches to nonetheless be treated as independent is to prevent a probability flow through the root node in the first place. From a graph-theoretic perspective, this solution corresponds to the (virtual) removal of all out-directed contacts from the root node. This approach does not modify the dynamics of the node under consideration because it can still contract the disease, and once infected, the recovery process is independent of the topology. However, the idea considerably reduces the amount of bookkeeping that would otherwise be necessary if we accounted for the correlations directly. The singular node l is said to be a cavity node or in the cavity state [23,51], a concept closely related to the test-node assumption [39] and the idea of cut vertices [63]. With this concept, we can factorize $\Psi_l(t)$ and thus obtain

$$S_l(t) = z_l \prod_{k \in \mathcal{N}_l} \theta_{k\to l}(t). \tag{4}$$

Here, we introduce the probability $\theta_{k\to l}(t)$ that no disease has been transmitted from node k to the cavity node l up to time t.

The change in perspective towards an edge-centric analysis introduces new auxiliary dynamic quantities such as $\theta_{k\to l}(t)$. These quantities are defined on the set of edges $\mathcal{E}$ of the time-aggregated network, and thus the number of dynamic variables scales with L, the number of edges.

In order to obtain a system of dynamic equations, we focus on our first edge-centric variable $\theta_{k\to l}$. Initially, no disease was transmitted such that $\theta_{k\to l}(0) = 1$ for all edges $(k, l) \in \mathcal{E}$. Henceforth, the dynamic quantity reduces only (i) upon a temporal contact indicated by $a_{k\to l}(t)$ and (ii) if the adjacent node k is infected without having transmitted the disease earlier to the cavity node l—we denote the corresponding probability by $I_{k\to l}(t)$. Hence, the out-flow of probability is given by $\beta a_{k\to l}(t) I_{k\to l}(t)$, leading to our first dynamic equation

$$\theta_{k\to l}(t+1) = \theta_{k\to l}(t) - \beta a_{k\to l}(t) I_{k\to l}(t). \tag{5}$$

Next, the probability $I_{k\to l}(t)$ evolves according to three contributions. (i) It decreases with the recovery probability μ and (ii) with the probability β to infect its target node upon a temporal contact. These processes are independent and may contribute simultaneously with the joint probability $\beta\mu$. (iii) $I_{k\to l}(t)$ increases with the probability $\Delta S_{k\to l}(t) = S_{k\to l}(t) - S_{k\to l}(t+1)$ that k is newly infected by at least one of its incident neighbors *excluding* the cavity node l. In sum and with the initial condition $I_{k\to l}(0) = 1 - z_k$, these contributions lead to

$$I_{k\to l}(t+1) = (1-\mu)[1 - \beta a_{k\to l}(t)] I_{k\to l}(t) + \Delta S_{k\to l}(t). \tag{6}$$

Finally, we consider the probability $S_{k\to l}(t)$ that node k, adjacent to the cavity node l, is susceptible. Since k is not affected by the state of l, it stays susceptible if it does not contract the disease from any of its remaining, incident neighbors $j \in \mathcal{N}_k \backslash l$. It has been shown in Ref. [64] that the corresponding probability $\Phi_{k\to l}(t) = \prod_{j\in\mathcal{N}_k\backslash l} \theta_{j\to k}(t)$ factorizes, and thus, similar to Eq. (3), we find $S_{k\to l}(t) = z_l \Phi_{k\to l}(t)$ or, equivalently,

$$S_{k\to l}(t+1) = z_k \prod_{j\in\mathcal{N}_k\backslash l} \theta_{j\to k}(t+1). \tag{7}$$

The disease progression in the CB framework is fully characterized by Eqs. (5) and (6), a set of $2L$ coupled equations. Equation (7) is introduced here for convenience only and can be substituted into Eq. (6). Next, we return to the node-centric quantities. To this end, we note that $S_l(t)$ has already been determined in Eq. (4). The remaining marginals I_l and R_l are equivalent to the IB model and given by the conservation condition, as well as the transition to the recovered state in Eqs. (8) and (9), respectively:

$$I_l(t+1) = 1 - S_l(t+1) - R_l(t+1), \tag{8}$$

$$R_l(t+1) = R_l(t) + \mu I_l(t). \tag{9}$$

The CB model is exact for temporal networks, where the undirected, time-aggregated graph has a tree structure and is therefore loop-free. Most realistic networks, however, contain a large number of loops such as triangles in social graphs, where two friends are likely to have many more friends in common. Here, the CB model nevertheless appears to be "unreasonably effective" (cf. Ref. [65]) and improves predictions significantly with respect to the IB approach as we will see in Sec. V. For further extensions to the model that include heterogeneous infection and recovery probabilities, as well as weighted contacts, see Appendix A.

IV. EPIDEMIC THRESHOLD

The parameters that mark the epidemic threshold can be derived by examining small perturbations around the disease-free state. If such perturbations die out, then any outbreak remains local, but if the perturbation grows, then a global epidemic may occur. We consider a linearization of the dynamic equations (5)–(7), which will give rise to a criticality condition, determining the epidemic threshold. We begin with the ansatz $\theta_{k\to l}(t) = 1 - \delta_{k\to l}(t)$ and $z_l = 1 - \epsilon_l$, where $\delta_{k\to l}(t), \epsilon_l \ll 1$ are small perturbations around the disease-free state for all nodes l and edges (k, l). Thus, Eq. (5) becomes

$$\delta_{k\to l}(t+1) = \delta_{k\to l}(t) + \beta a_{k\to l}(t) I_{k\to l}(t). \tag{10}$$

In Eq. (7), we keep the linear terms of the Taylor expansion, which transforms the product into a corresponding sum:

$$S_{k\to l}(t+1) = (1-\epsilon_k)\prod_{j\in\mathcal{N}_k\setminus l}[1-\delta_{j\to k}(t+1)] \tag{11a}$$

$$\approx 1-\epsilon_k-\sum_{j\in\mathcal{N}_k\setminus l}\delta_{j\to k}(t+1) \tag{11b}$$

$$= S_{k\to l}(t)+\beta\sum_{j\in\mathcal{N}_k\setminus l}a_{j\to k}(t)I_{j\to k}(t). \tag{11c}$$

In Eq. (11b), we substituted the dynamic equation (10) and identified $S_{k\to l}(t)$ in the next step. From the resulting Eq. (11c), we can read the linearized form of $\Delta S_{k\to l}$, which allows us to decouple the dynamic equations for $I_{k\to l}$:

$$I_{k\to l}(t+1) \approx (1-\mu)[1-\beta a_{k\to l}(t)]I_{k\to l}(t) + \beta\sum_{j\in\mathcal{N}_k\setminus l}a_{j\to k}(t)I_{j\to k}(t). \tag{12}$$

Next, we rewrite the remaining set of L dynamic equations in a compact, matrix-based formulation and therefore introduce the vectors $\boldsymbol{I}(t)$ and $\boldsymbol{a}(t)$ with elements $I_{k\to l}(t)$ and $a_{k\to l}(t)$, respectively. To this end, we also express the linear operation $\sum_{j\in\mathcal{N}_k\setminus l}a_{j\to k}(t)$ in Eq. (12), which acts on the elements $I_{k\to l}(t)$ of the state vector, through the temporal unweighted nonbacktracking matrix $\boldsymbol{B}(t)$:

$$B_{k\to l,j\to k'}(t) = \begin{cases} a_{j\to k'}(t) & \text{if } k'=k, \text{ and } j\neq l \\ 0 & \text{otherwise.}\end{cases} \tag{13}$$

In other words, $B_{k\to l,j\to k'}(t)=1$ if the contact (t,j,k') at time t is incident on the edge (k,l) (implying $k'=k$), and additionally $j\neq l$. Otherwise, we have $B_{k\to l,j\to k'}(t)=0$. It is only the nonbacktracking property $j\neq l$ that sets $\boldsymbol{B}$ apart from the adjacency matrix of the ordinary line graph. For temporal networks, a subtle distinction has to be made between the first and the second index of the $L\times L$ dimensional matrix $\boldsymbol{B}$: The first corresponds to an out-directed (*static*) edge $(k,l)\in\mathcal{E}$ of the underlying aggregated network and can be interpreted as a potential contact in the future. The second, however, is an incident (*temporal*) contact $(t,j,k')\in\mathcal{C}$ from node j to k' at time t. We also introduce the diagonal matrix $\mathrm{diag}(\mathbf{1}-\beta\boldsymbol{a}(t))$, with diagonal elements given by the vector $\mathbf{1}-\beta\boldsymbol{a}(t)$. Here, we denote by $\mathbf{1}$ the vector of all ones. With these definitions, we rewrite Eq. (12) as

$$\boldsymbol{I}(t+1) = [(1-\mu)\mathrm{diag}(\mathbf{1}-\beta\boldsymbol{a}(t))+\beta\boldsymbol{B}(t)]\boldsymbol{I}(t). \tag{14}$$

The explicit solution to the state vector $\boldsymbol{I}(T)$ at final observation time T is formally given by $\boldsymbol{I}(T)=\boldsymbol{P}(\beta,\mu)\boldsymbol{I}(0)$, where the so-called infection propagator $\boldsymbol{P}$ [54] is introduced for notational convenience:

$$\boldsymbol{P}(\beta,\mu) = \prod_{t=0}^{T-1}[(1-\mu)\mathrm{diag}(\mathbf{1}-\beta\boldsymbol{a}(t))+\beta\boldsymbol{B}(t)]. \tag{15}$$

In order to evaluate the asymptotic behavior, we assume a periodic boundary condition in time, i.e., $\boldsymbol{B}(t)=\boldsymbol{B}(t+T)$. This allows us to assess the vulnerability of the temporal network through the spectral radius of the propagator $\boldsymbol{P}$. In particular, we find that a SIR-type outbreak is asymptotically stable under small perturbations, i.e., remains confined to a small set of nodes, as long as the spectral radius satisfies $\rho[\boldsymbol{P}(\beta,\mu)]<1$. Thus, the phase transition is given by the criticality condition

$$1=\rho\left(\prod_{t=0}^{T-1}[(1-\mu)\mathrm{diag}(\mathbf{1}-\beta\boldsymbol{a}(t))+\beta\boldsymbol{B}(t)]\right). \tag{16}$$

Note that for irreducible and non-negative matrices, the largest eigenvalue is simple and positive according to the Perron-Frobenius theorem [66]. Assuming $0\leq\beta,\mu<1$, a sufficient condition for temporal networks is to restrict contacts to the giant strongly connected component (GSCC) of the underlying time-aggregated graph. In Sec. V B, we fix the recovery probability μ and determine the critical infection probability β_{crit} as the root of $f(\beta)=1-\rho[\boldsymbol{P}(\beta,\mu)]$ for different empirical networks.

We conclude this section with a discussion on the static network limit. In the so-called quenched regime, the disease evolves on a much faster timescale than the dynamic topology and thus operates on an effectively static network with $\boldsymbol{B}(t)\equiv\boldsymbol{B}(0)\equiv\boldsymbol{B}$ and $\boldsymbol{a}(t)\equiv\mathbf{1}$ for all times t. As in the temporal analysis, we restrict the network to the GSCC so that the Perron-Frobenius theorem [66] applies. In this limit, the dynamic equations (5)–(7) reduce to the dynamic message-passing formulation in Ref. [51]. Moreover, Eq. (15) now becomes a product $\prod_{t=0}^{T-1}\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)=[\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)]^T$ of T identical, single time-step propagators

$$\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu) = (1-\mu)(1-\beta)\mathbb{1}+\beta\boldsymbol{B}, \tag{17}$$

where $\mathbb{1}=\mathrm{diag}(\mathbf{1})$ denotes the identity matrix.

The spectral radius in Eq. (16) factorizes to $\rho[\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)^T]=\rho[\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)]^T$, and it follows that the criticality condition Eq. (16) reduces to $\rho[\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)]=1$. Furthermore, we find from basic linear algebra that $\rho[\boldsymbol{P}_{\mathrm{fast}}(\beta,\mu)]=(1-\mu)(1-\beta)+\beta\rho(\boldsymbol{B})$, and hence we obtain the corresponding static threshold condition

$$\left(\frac{\beta}{\beta+\mu-\beta\mu}\right)_{\mathrm{crit,fast}}=\frac{1}{\rho(\boldsymbol{B})}. \tag{18}$$

The criticality condition in Eq. (18) deviates from the continuous-time result in Refs. [16,23]. In the derivation presented here, the term $\beta\mu$ in Eq. (18) accounts for the

simultaneous events when a node infects a neighbor and recovers within the same time step.

In contrast to the quenched regime, one can also consider the so-called annealed limit. Then, parameters β and μ are sufficiently small such that no more than one infection or recovery event can take place within the observation time. Therefore, we expand the infection propagator to the first order in β and μ and obtain

$$\boldsymbol{P}_{\text{slow}}(\beta,\mu) = (1-T\mu)\mathbb{1} - T\beta\,\text{diag}(\bar{\boldsymbol{a}}) + T\beta\bar{\boldsymbol{B}}. \quad (19)$$

Here, $\bar{\boldsymbol{a}} = 1/T\sum_t \boldsymbol{a}(t)$ and $\bar{\boldsymbol{B}} = 1/T\sum_t \boldsymbol{B}(t)$ denote the corresponding time-averaged quantities. It is insightful to evaluate simple bounds for the set of parameters $(\beta,\mu)_{\text{crit,slow}}$ that satisfy the threshold condition $\rho(\boldsymbol{P}_{\text{slow}}) = 1$ in the annealed limit. With $1/T \le \bar{\boldsymbol{a}} \le 1$ for all elements in $\bar{\boldsymbol{a}}$, we thus find

$$\left(\frac{\beta}{\beta+\mu}\right)_{\text{crit,slow}} \le \frac{1}{\rho(\bar{\boldsymbol{B}})} \le \left(\frac{\beta}{\beta/T+\mu}\right)_{\text{crit,slow}}. \quad (20)$$

Assuming the upper bound in Eq. (20) overestimates the outbreak risk and can be considered a conservative choice from an epidemiological perspective. This limit is realized for a temporal network where every edge appears exactly once within the observation time, hence $\bar{\boldsymbol{a}} = 1/T$. The lower bound in Eq. (20) is exact in the case of a static network (thus $\bar{\boldsymbol{a}} = 1$) and corresponds to the continuous-time result in Ref. [16]. However, this limit underestimates the outbreak risk, and therefore we conclude with a note of caution when applying results from static network theory directly to time-varying topologies.

V. APPLICATION

A big advantage common to both the node-centric IB and edge-centric CB modeling framework is a significant reduction in computational complexity compared to MC simulations. The CB model requires iteration through all edges at every time step, and thus the time complexity scales with $\mathcal{O}(LT)$. The IB formulation and a single MC realization require $\mathcal{O}(\bar{C}T)$, where $\bar{C}$ denotes the average number of active contacts, which can be significantly smaller than L. Stochastic MC simulations, on the other hand, require a large number of realizations in order to provide reliable statistics. The computational disadvantage of MC simulations becomes even more apparent when we consider a complex quantity such as the epidemic threshold, which requires multiple ensemble averages for different sets of epidemic parameters in order to fit the critical infection probability (see Sec. V B). Equally important, however, is the accuracy of our analytic approach. Therefore, in this section, we compare estimations from the IB and CB mean-field model with MC simulations using empirical data as introduced in Sec. II.

A. Numerical analysis of the mean-field dynamics

We begin with an analysis on the level of individual nodes. In Fig. 2, we show the cumulative infection probability for a small number of example nodes from the conference data set given the same outbreak location. The selection is intended to present qualitatively different trajectories, also demonstrating that deviations between the two models vary considerably. The MC result (blue curve) in Fig. 2(a) corresponds to the introductory example in Fig. 1. Here, a comparison with the analytic estimation shows that the CB approach leads to a substantial improvement to the IB model. Also in Figs. 2(b)–2(d), the trajectories are erratic, as they reflect the sudden changes in the underlying topology, highly individual and yet well approximated by the CB model. For all nodes in the network, we find that the CB model gives a closer upper bound to MC simulations because, unlike the IB framework, it accounts for dynamic correlations between nearest-neighbor states.

Dynamic mean-field models such as the IB and CB framework provide realistic expectation values only if stochastic fluctuations are negligible. In order to illustrate the limitations, we study epidemic outbreaks for three different initially infected nodes in Figs. 3(a)–3(c), respectively. The left column gives the time-resolved distribution of the outbreak size, and the right column presents the final distribution at the end of the three-day observation period. For the ensemble average (blue line), we consider only realizations with more than 20 infected nodes overall. This threshold separates outbreaks that die out early due to stochastic fluctuations and thus permits a direct comparison with estimations from the IB and CB frameworks in green and red, respectively.

We choose the outbreak locations such that the degree of stochasticity increases from top to bottom. In Fig. 3(a1), we

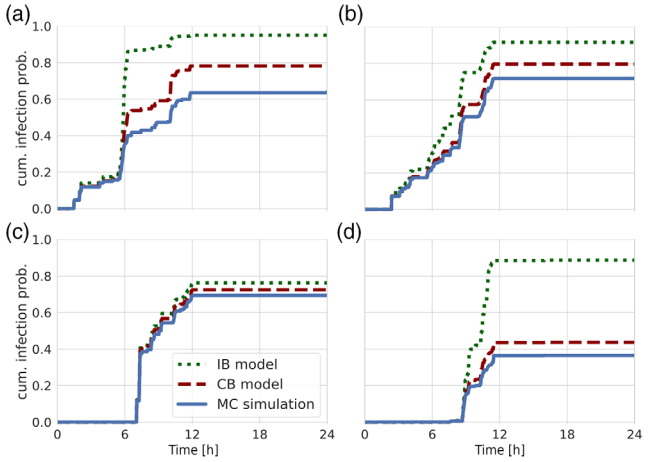


FIG. 2. Epidemic trajectories for four exemplary individual nodes. We compare the cumulative infection probability from MC simulations (blue line) with estimations from the CB model (red dashed line) and the IB approach (green dotted line). Simulated results are averaged over 10^4 MC realizations with the same outbreak location and disease parameters as in Fig. 1.

find a narrow distribution around the ensemble average, which is well approximated by the mean-field models. Minor outbreaks due to early extinctions are well separated in Fig. 3(a2) from large epidemics. In Fig. 3(b), the initially infected node leads to realizations with considerably stronger fluctuations, and in Fig. 3(c), it is barely possible to separate early extinctions at all. Additionally, we observe a second source of stochastic variation, namely, the time at which a disease escalates and hence evolves into a global epidemic. As a consequence, early outbreak sizes may be overestimated significantly before the analytic trajectory approaches the expectation value again [see Fig. 3(c1)].

Remarkably, the performance of both mean-field models varies significantly with the outbreak location, even for the basic reproduction number R_0 well above the epidemic threshold. At the late phase of an outbreak, however, the mean-field models provide good approximations, and consistently with Fig. 2, we find that the CB model outperforms the IB approach. In Appendix C, we demonstrate how a sufficiently large number of initially infected individuals significantly improves the predictability.

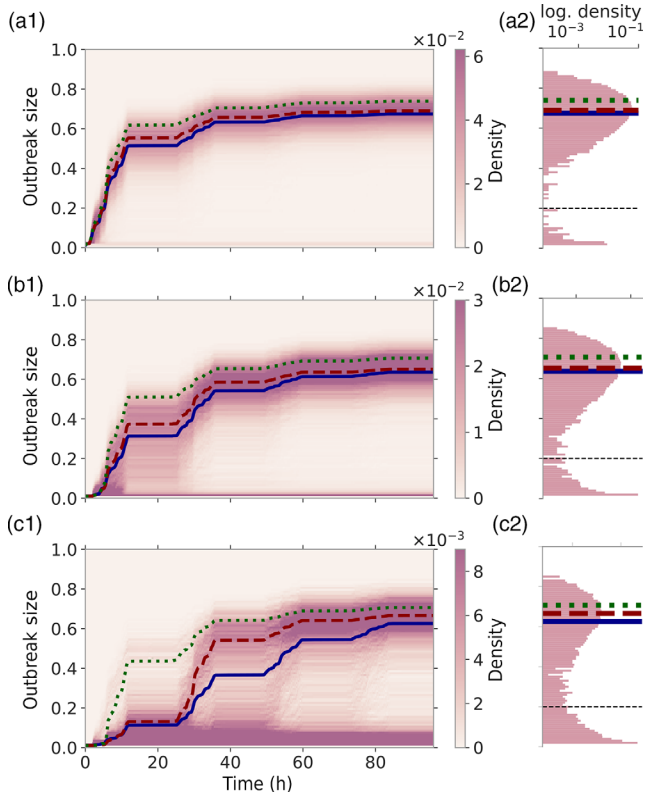


FIG. 3. Left column: Time-resolved and normalized outbreak-size distributions for three different initially infected nodes. Epidemic parameters are as in Fig. 1 and 10^4 MC realizations. The expectation value (blue line) assumes a cutoff at 20% of the population size (black dashed line). The CB and IB models are presented as a red dashed and green dotted line, respectively. Right column: Outbreak-size distribution at the final observation time in logarithmic scale with corresponding expectation values.

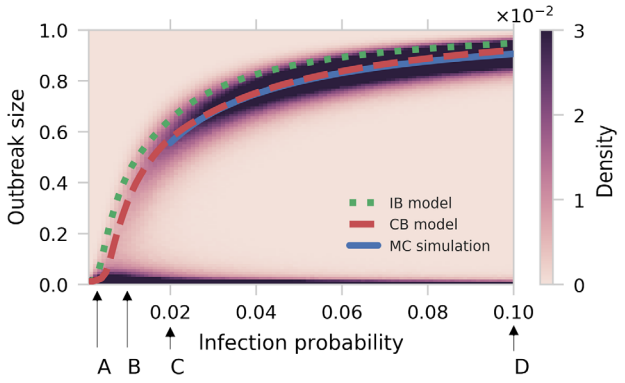


FIG. 4. Distribution of final outbreak sizes as a function of the infection probability β. We perform 10^5 MC realizations for every value of β, each starting with one randomly chosen outbreak location. For $\beta > 0.02$, we show the averaged final outbreak size (blue line) with a cutoff at 20% of the population size. Estimations from the CB and IB models are presented as red dashed and green dotted lines, respectively. Labeled arrows at the bottom mark infection probabilities that correspond to Figs. 5(a)–5(d), respectively.

Another source of stochasticity is the choice of disease parameters β and μ, respectively. We focus on the final outbreak size, averaged over all outbreak locations. The distribution as a function of the infection probability β (see Fig. 4) shows a percolation-like transition from localized spreading to epidemics that affect a considerable fraction of the network. We apply the same threshold as in Fig. 3 for a direct comparison between the averaged outbreak size and the mean-field models for $\beta > 0.02$. Here, we find that the difference between the expected size and the CB estimation is close to negligible, whereas the IB model consistently overestimates the expected value.

A comparison at low values of the infection probability β becomes unreliable as stochasticity impedes a reasonable distinction between minor and global outbreaks. In order to illustrate the effect, we present in Fig. 5 the outbreak-size distribution for different values of β as marked by the arrows in Fig. 4. This representation highlights the transition from the subcritical to the supercritical parameter domain: The unimodal distribution in Fig. 5(a) characterizes localized outbreaks, whereas the bimodal distribution in Fig. 5(d) clearly separates early extinctions and global epidemics. Next, we focus on the critical infection probability that marks the transition.

B. Epidemic threshold

In Fig. 6(a), we present the region of small β from Fig. 4 in order to focus on the transition from localized outbreaks to the sudden emergence of global epidemics. We determine the critical infection probability β_{crit} (vertical blue line) from the maximum of the relative standard deviation [54], also known as the coefficient of variation [see blue line in Fig. 6(b)]:

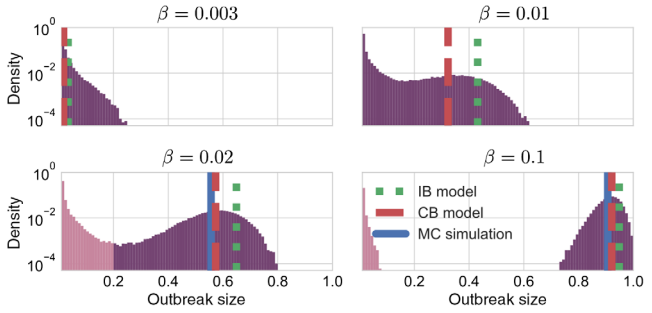


FIG. 5. Distribution of final outbreak sizes for $\beta = 0.003$, 0.01, 0.02, and 0.1, respectively. For $\beta = 0.02$ and 0.1, we mark part of the distribution with outbreak sizes below the given threshold of 20% by a lighter color tone and neglect this contribution to the averaged value (blue vertical line). The expected outbreak size from MC simulations and the estimations from the CB and IB models are plotted as blue, red dashed, and green dotted vertical lines, respectively.

$$c_v = \frac{\sqrt{\langle \sigma^2 \rangle - \langle \sigma \rangle^2}}{\langle \sigma \rangle}. \tag{21}$$

Here, we denote with $\langle \sigma \rangle$ and $\langle \sigma^2 \rangle$ the first and second moments of the outbreak-size distribution. The coefficient of variation captures the intuition that fluctuations dominate the outbreak-size distribution close to the transition. Indeed, c_v diverges at the critical point for infinitely large networks, indicating a second-order phase transition [67].

Analytically, we determine β_{crit} from the spectral criterion in Eq. (16) for the CB model and similarly within the IB framework [54,68]. The comparison in Fig. 6 shows that

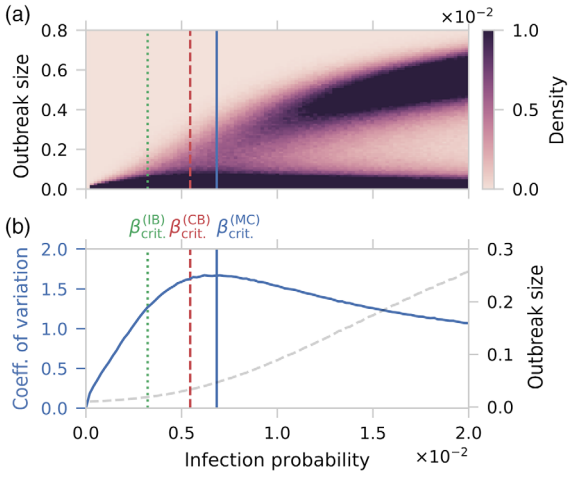


FIG. 6. Estimation of the critical infection probability β_{crit}. (a) Outbreak-size distribution as in Fig. 4 for small values of β. The vertical blue, red dashed, and green dotted lines mark the critical value according to MC simulations ($\beta_{\mathrm{crit}}^{(\mathrm{MC})}$), the CB model ($\beta_{\mathrm{crit}}^{(\mathrm{CB})}$), and the IB approach ($\beta_{\mathrm{crit}}^{(\mathrm{IB})}$), respectively. (b) From the distribution in (a), we derive the coefficient of variation (blue line, left axis) and the mean outbreak size (grey dashed line, right axis).

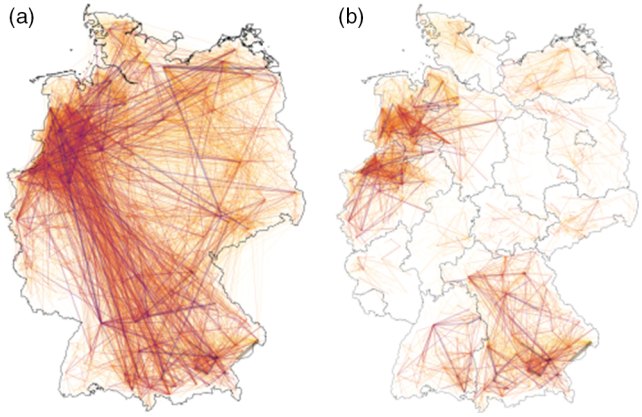


FIG. 7. (a) Cattle trade within Germany. Weighted edges correspond to directed trade relations within the year 2010, whereas the color indicates the accumulated number of traded animals. (b) Cattle trade within the federal states of Germany. We confine the underlying time-aggregated graph to the GSCC and visualize here only edges with a flux of at least 50 animals.

the IB and CB models, marked by a red dashed and green dotted line, respectively, underestimate the critical infection probability from MC simulations (blue line) and thus overestimate the outbreak risk. Consistent with our previous results, we can state that a shift from a node- to an edge-centric framework improves the analytic prediction. In Appendix C, we present similar results for different values of the recovery probability μ. Next, we continue with a realistic application of the epidemic threshold to the German cattle-trade network.

1. Application to German cattle trade

We now consider a completely different data set, where the system size is large and contacts are sparse over time. Our example is a cattle-trade network, where the movements of animals between farms in Germany are recorded on a daily basis. Next, we isolate the trade within each federal state of Germany as visualized in Fig. 7 and restrict trade to the GSCC of the underlying aggregated graph. Disregarding the smallest networks (those with less than 27 nodes), we thus obtain 12 time-varying graphs with sizes varying from 254 to 27 863 nodes and highly heterogeneous topological and temporal features (see Appendix B).

As in the previous section, we assume that premises can be either susceptible, infected, or recovered, and trade events facilitate the transmission of a disease. Unlike before, however, we take into account the number of traded animals during each transaction, i.e., the weight $w_{k\to l}$ of a (temporal) contact from node k to l. To this end, we modify the infection propagator in Eq. (15) and replace β by $1-(1-\beta)^{w_{k\to l}}$ (see Appendix A for more information). In a potential outbreak, we assume that an infected node is detected with a constant probability μ each day, after which it would be isolated and thus removed from the network. As a consequence, highly

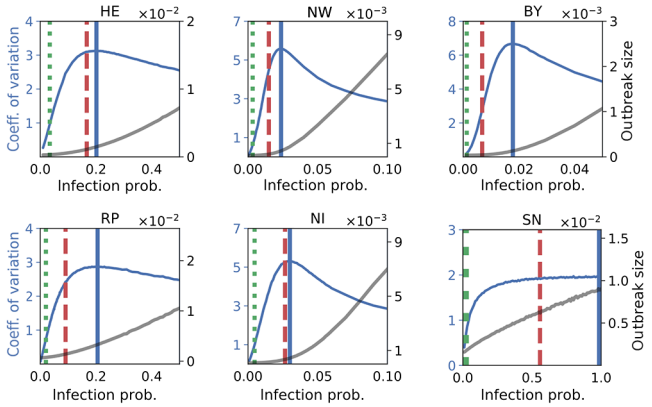


FIG. 8. Detailed threshold analysis for six selected federal states with $\mu = 1/28$ (cf. Table I). We show the simulated mean outbreak size (grey line, right axis) and coefficient of variation (blue line, left axis) averaged over all initially infected nodes. The critical infection probability from MC simulations are shown, and the IB and CB models are presented as vertical blue, green dotted, and red dashed lines, respectively.

infectious diseases such as FMD can be modeled as SIR-type epidemics [60].

In Fig. 8, we compare the critical infection probability similar to Fig. 6 for six selected federal states with different transition characteristics. The critical value derived from MC simulations varies between $\beta_{\mathrm{crit}} = 0.018$ [Bavaria (BY)] and $\beta_{\mathrm{crit}} = 1.0$ [Saxony (SN)]. The latter indicates that outbreaks remain localized for every choice of β due to sparse intrastate trade.

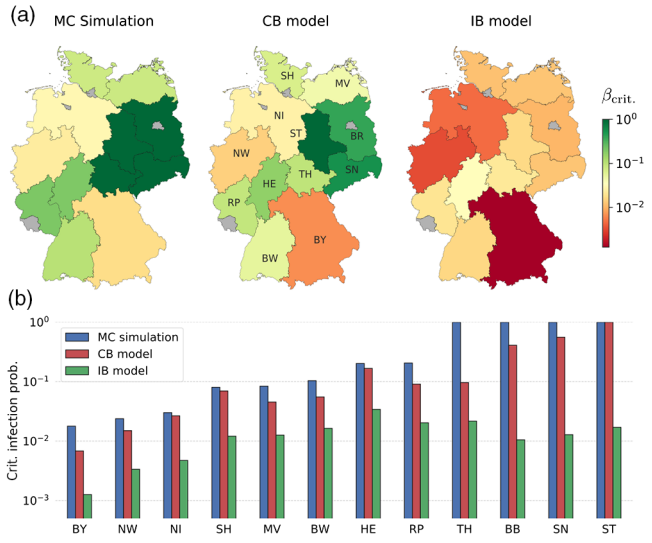


FIG. 9. (a) Spatial variation of epidemic risk due to cattle trade in Germany for $\mu = 1/28$. Federal states are colored according to the critical infection probability β_{crit} as determined from MC simulations, the CB and IB models, respectively (see Figs. 8 and 19 for details). The city states Berlin (B), Hamburg (HH), and Bremen (HB), as well as Saarland (SL), are excluded due to the small network size (see Table I). (b) Critical infection probability β_{crit} in logarithmic scale, sorted from high (left) to low risk (right). Results from MC simulations, and the CB and IB models are presented as groups of blue, red, and green bars, respectively. Disclaimer: A realistic vulnerability analysis requires, for instance, heterogeneous recovery probabilities and complex countermeasures (see Appendix B for details).

As a potential application to public health institutions, we present in Fig. 9(a) the spatial variation of the epidemic risk in terms of β_{crit}. The quantitative comparison in Fig. 9(b) demonstrates that spectral methods provide a lower bound with a varying degree of accuracy depending on the network details. Despite their heterogeneity in size and activity, we find for all networks that the CB model outperforms the IB approach. The detailed results for all states as well as a similar analysis for $\mu^{-1} = 120$ are available in Appendix C.

VI. CONCLUSION

In this paper, we have presented the CB model for epidemic SIR spreading on temporal networks as a conceptually similar framework to the widely used IB approach. Derived from the message-passing framework [23,51], it inherits its accuracy on loop-free topologies and improves analytic estimations with respect to the IB approach for arbitrary time-evolving graphs. Moreover, the focus on edge-based quantities that are updated in discrete time steps allows a seamless integration of temporal interactions. Structurally similar to the node-centric IB model, the proposed CB approach poses a low conceptual barrier and admits application on large graphs.

Importantly, the accuracy of the CB model improves existing approximations of the epidemic threshold, which is a crucial risk measure for public health institutions. To this end, we have studied the largest eigenvalue of the infection propagator matrix, which determines the disease propagation in the low prevalence limit and takes into account the full temporal and topological information up to the observation time. The largest eigenvalue can be easily found through repeated matrix multiplications, i.e., the so-called power method. Without relying on extensive MC simulations and a subsequent parameter fit, the critical value can thus be estimated with efficient, vectorizable tools from linear algebra that are available for most high-level programming languages.

In the application section, we focused first on a social contact graph that can be used to analyze the propagation of airborne diseases as well as the spread of information. Our comparison between MC simulations and analytic estimations from the CB and IB models followed a bottom-up approach: We looked at (i) epidemic trajectories of individual nodes, (ii) averaged trajectories given the same outbreak location, and (iii) the final outbreak size for a range of infection probabilities and with random initial condition. In all cases, the CB model provides a closer upper bound to MC simulations than the widely used IB model. All results based on the conference data set can be reproduced using the PYTHON code provided in Ref. [57].

As a particularly important application, we then compared analytic estimations of the critical infection probability with extensive MC simulations. To this end, we included a case study of livestock trade within 12 federal states in Germany with highly heterogeneous characteristics in terms of size, density, and temporal activity. Consistently, we found that the CB model improves the previously proposed lower bound at a low conceptual and computational cost.

Many excellent results have already been derived within the IB framework for empirical networks and in the context of random graphs (see Ref. [69] for a recent review) that can further improve the CB model. We therefore expect that the conceptual simplicity of the CB framework allows us to integrate features such as non-Markovianity [22], stochastic effects [70], and estimations of uncertainty [71] that are important to realistic disease models on temporal networks. Also, first steps towards higher-order models that go beyond the tree-graph assumption have been proposed in the context of percolation theory [72] and diffusive transport [73,74], and we expect these improvements to be applicable to the CB model as well.

The data are available on [75], and using the source code in Ref. [57], results of this paper can be easily reproduced.

ACKNOWLEDGMENTS

A. K. and P. H. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) in the framework of Collaborative Research Center 910. A. K. acknowledges further support from the German Academic Exchange Service (DAAD) via a short-term scholarship. J. P. G. acknowledges support from Science Foundation Ireland (Grants No. 16/IA/4470 and No. 16/RC/3918). The authors thank Vittoria Colizza, Eugenio Valdano, and Alexandre Darbon for helpful comments.

APPENDIX A: WEIGHTED NETWORKS AND HETEROGENOUS INFECTION AND RECOVERY PROBABILITIES

In order to improve the predictive power of a network model, it is often required to take into account additional information. In the main article, we focus on the temporal dimension. However, another important piece of information is the weight of a contact. The interpretation of weight can range from passenger numbers in the global air traffic network to the impedance in a network of electric components. The distribution of weights in static as well as temporal empirical networks often shows a broad tail [76], and as such, the averaged edge weight can become meaningless because of large fluctuations. It is therefore often required to account for heterogeneous edge weights explicitly in epidemiological models. However, depending on the interpretation, weights may enter the model in different ways. Typically, a time-dependent edge weight $w_{k\to l}(t)$ is considered similar to the conductivity between two nodes k and l in an electric circuit. Translated to an epidemiological context, we would thus scale the infection probability linearly; i.e., $\beta a_{k\to l}(t)$ becomes $\beta w_{k\to l}(t)$ in a weighted network. Another approach, popular in the context of random walks and disease spreading, is to interpret an integer-valued weight $w_{k\to l}(t)$ as a number of parallel and unweighted edges that connect k with l [7,46]. From an epidemiological viewpoint, this idea would translate to $w_{k\to l}$ independent attempts to transmit the disease at time t. Here, the infection probability $\beta a_{k\to l}(t)$ becomes $1-(1-\beta)^{w_{k\to l}(t)}$ in the weighted case. In the main article, we apply the latter interpretation to calculate the epidemic risk in the context of livestock trade (see Fig. 9). Here, weights correspond to the number of animals traded, each of which can infect the target population independently. For small probabilities $\beta \ll 1$, the adjusted infection probability simplifies to $1-(1-\beta)^{w_{k\to l}(t)} \approx \beta w_{k\to l}(t)$.

A second source of heterogeneity that is commonly considered includes heterogeneous infection and recovery probabilities, denoted as $\beta_{k\to l}$ and μ_k, respectively. With these modifications, the dynamic equations (5)–(7) from the main text translate to

$$\theta_{k\to l}(t+1) = \theta_{k\to l}(t) - \Psi_{k\to l}(t) I_{k\to l}(t), \tag{A1a}$$

$$S_{k\to l}(t+1) = z_k \prod_{j\in\mathcal{N}_k\setminus l} \theta_{j\to k}(t+1), \tag{A1b}$$

$$I_{k\to l}(t+1) = (1-\mu_k)[1-\Psi_{k\to l}(t)] I_{k\to l}(t) + \Delta S_{k\to l}(t). \tag{A1c}$$

Here, $\Psi_{k\to l}(t)$ denotes the probability that k infects l at time t given that the former is infected and has not yet transmitted the disease. For weighted networks, we can choose $\Psi_{k\to l}(t) = \beta_{k\to l} w_{k\to l}(t)$ or $\Psi_{k\to l}(t) = 1-(1-\beta_{k\to l})^{w_{k\to l}(t)}$ as discussed above. The linearization of Eqs. (A1a)–(A1c) around the disease-free state leads to

$$\boldsymbol{I}(t+1) = \mathrm{diag}[(\mathbf{1}-\boldsymbol{\mu})\circ(\mathbf{1}-\boldsymbol{\Psi}(t))]\boldsymbol{I}(t) + \boldsymbol{B}_\beta(t)\boldsymbol{I}(t). \tag{A2}$$

Here, the circle denotes the elementwise product. Moreover, the L-dimensional vectors $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}(t)$ integrate the node- and edge-dependent values μ_k and $\Psi_{k\to l}(t)$, respectively. We also generalize the temporal nonbacktracking matrix $\boldsymbol{B}_\beta(t)$ from Eq. (13) to the weighted one:

$$[\boldsymbol{B}_\beta(t)]_{k\to l, j\to k'} = \begin{cases} \Psi_{j\to k'}(t) & \text{if } k' = k, \text{ and } j \neq l \\ 0 & \text{otherwise.} \end{cases} \tag{A3}$$

The largest eigenvalue ρ of the infection propagator determines the asymptotic stability for small perturbations around the disease-free state. Accounting for heterogeneity in β, μ, and contact weights, the criticality condition Eq. (16) from the main text reads

$$1 = \rho\left(\prod_{t=0}^{T-1}\{\mathrm{diag}[(\mathbf{1}-\boldsymbol{\mu})\circ(\mathbf{1}-\boldsymbol{\Psi}(t))]+\boldsymbol{B}_{\beta}(t)\}\right). \quad \text{(A4)}$$

Assuming $\beta \equiv \beta_{k\to l}$ for all edges $k \to l$, we can determine from Eq. (A4) the critical (homogeneous) infection probability β_{crit} given a weighted, temporal network with heterogeneous recovery probabilities μ_k. Similarly, one can assume $\mu \equiv \mu_k$ for all nodes k and thus derive the critical (homogeneous) value μ_{crit} with heterogeneity in the infection probability $\beta_{k\to l}$.

APPENDIX B: GERMAN CATTLE-TRADE NETWORK

The system of traceability of cattle in the EU requires that each animal is identified with ear tags and that each movement, birth, or death event has to be reported within 7 days of the event to the national livestock database. We consider an excerpt of the national German livestock database HI-Tier [59] for the year 2010. The database is administered by the Bavarian State Ministry for Agriculture and Forestry on behalf of the German Federal States. It records 3.2 million animal movements with a total of 13.4 million traded animals between 183 454 premises, such as farms, pastures, slaughterhouses, and traders within the observation window. The location of each animal holding was provided at the resolution of the municipality. We consider each trade event between two premises a temporal contact, and we identify an edge if at least one contact has been recorded. The distribution of edges is highly heterogeneous in terms of geography, degree, and weight. In Fig. 10(a), we observe clusters of trade activity mostly within and between North Rhine-Westphalia (NW), Lower Saxony (NI), Baden-Württemberg (BW), and Bavaria (BY). The number of trading partners, i.e., the node degree, is broadly distributed as demonstrated in Fig. 10(b). Here, we differentiate between in, out, and total degree. Similarly, we find a broad distribution of edge weights in Fig. 10(c), i.e., the number of traded animals along a given edge.

The geographic distribution of nodes in Fig. 11(a) shows dense regions in the northwest and southeast including the above-mentioned federal states NW, NI, BW, and BY. Here, we also find the largest premises in terms of total traded animals: In Fig. 11(c), color and size indicate the node strength, i.e., the aggregated trade volume. The heterogeneous distribution of strength also becomes apparent in Fig. 11(d) where in, out, and total strength are analyzed separately. Finally, we observe in Fig. 11(b) the net flux, i.e., the difference between in- and out-directed trade volume. We display only nodes with at least 500 traded animals in Figs. 11(b) and 11(c).

From a temporal perspective, we find that trade fluctuates between 10^2 and 10^4 active nodes, i.e., farms with at least one trade event on a given day, whereas minima

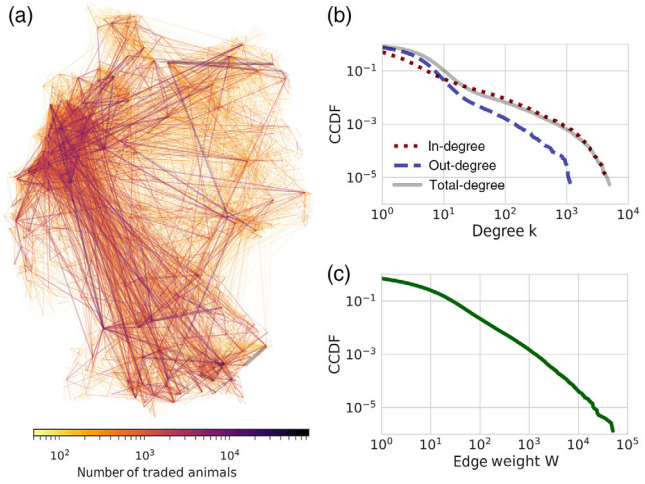


FIG. 10. Degree and weight analysis of cattle trade in the year 2010. Trade data are aggregated from 2010-01-01 to 2010-12-31. (a) Visualization of edges using geolocation with at least 50 animals traded along each link. The color and edge width indicate the aggregated edge weight. (b) Complementary cumulative distribution function (CCDF) of in, out, and total degree. (c) CCDF of the edge-weight distribution.

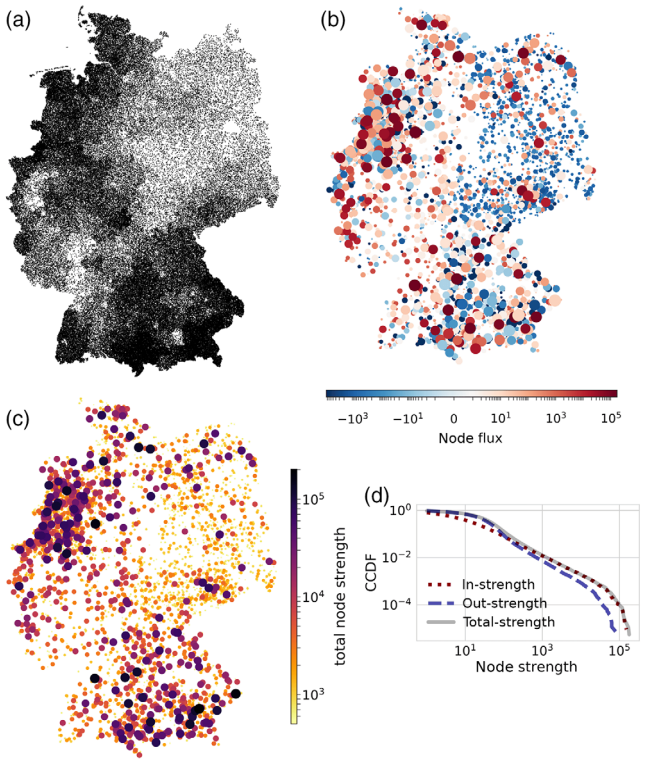


FIG. 11. Node strength and flux analysis of cattle trade in the year 2010. (a) Geographic location of all 183 454 premises that have been recorded. (b) Color indicates net animal flux, i.e., the difference between in- and out-directed animal flux. Node size corresponds to the sum of both [see diagram (c)]. (c) Color and size indicate the total node strength, i.e., the sum of incoming and outgoing animals. (d) CCDF of the in, out, and total strength.

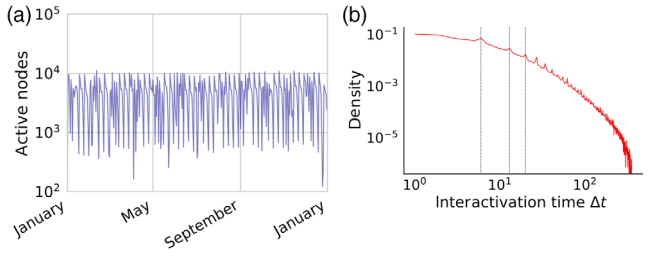


FIG. 12. Temporal analysis cattle trade in the year 2010. (a) Number of premises that trade at least once on a given day. (b) Interevent time distribution, i.e., the interval between subsequent trade events of an arbitrary farm. Vertical lines indicate 7, 14, and 21 days.

appear regularly on the weekends [see Fig. 12(a)]. The weekly pattern is also apparent in the interactivation time distribution, i.e., the time interval between two successive trade events for a given node (see Fig. 12). Here, we find a broad distribution of activity with peaks around 7, 14, and 21 days.

The geographic risk analysis in Sec. V B 1 requires us to separate the network into subgraphs that correspond to the intrastate trade (see Fig. 7). The largest eigenvalue of the corresponding infection propagator allows us to evaluate the outbreak risk within a federal state due to the local movement of infected animals. In Table I, we list the names of all 16 federal states of Germany together with the corresponding ISO abbreviation and basic statistics: the number of nodes, (static) edges, and

TABLE I. Federal states of Germany and the corresponding abbreviation, together with basic network statistics: number of nodes, number of (static) edges, and number of (temporal) contacts. In all cases, we confined the network to the GSCC. We mark the smallest networks, i.e., the city states Berlin, Hamburg, and Bremen, as well as Saarland, with an asterisk.

Federal state	ISO code	Nodes	Edges	Contacts
Schleswig-Holstein	SH	3570	13 541	49 748
Hamburg*	HH	2	2	8
Lower Saxony	NI	12 838	61 044	272 579
Bremen*	HB	4	6	21
North Rhine-Westphalia	NW	9826	42 835	209 677
Hesse	HE	3622	12 498	51 287
Rhineland-Palatinate	RP	1526	5386	29 184
Baden-Württemberg	BW	9168	34 434	150 171
Bavaria	BY	27 863	128 596	550 047
Saarland*	SL	26	52	343
Berlin*	BE	0	0	0
Brandenburg	BB	715	2144	11 535
Mecklenburg-Vorpommern	MV	844	2852	21 864
Saxony	SN	690	1935	12 369
Saxony-Anhalt	ST	254	714	3422
Thuringia	TH	344	957	5361

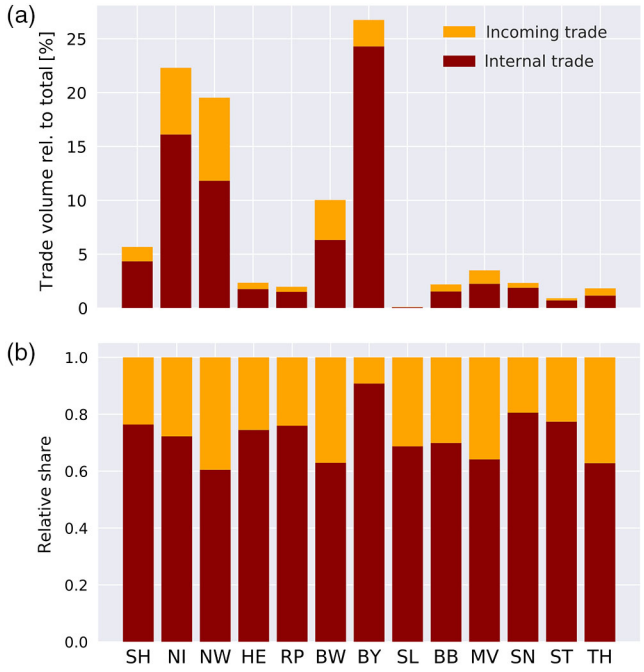


FIG. 13. Error estimation after restricting trade to the federal states. (a) Intrastate trade in terms of trade volume relative to total trade as red bars. Yellow bars on the top indicate the fraction of incoming animals. All bars together sum up to 100%. (b) Sum of internal and incoming trade volume normalized for each federal state. See Table I for a list of states and their abbreviations.

(temporal) contacts in the GSCC. The city states Berlin, Hamburg, and Bremen as well as Saarland, which is a particularly small state in terms of nodes, are marked with an asterisk and are not considered for risk analysis in Fig. 9.

Separating the trade network into subgraphs as visualized in Fig. 7(b) inevitably reduces the outbreak risk as the neglected cross-border edges would otherwise facilitate the disease transmission. In Fig. 13(a), we find that a considerable fraction of trade is directed across federal states and has thus been removed. This case applies, in particular, to the federal states NI, NW, and BW. Similarly, we find that the ratio between intrastate and in-directed trade lies between 0.6 (NW) and 0.9 (BY). Thus, we conclude that a considerable fraction of trade across borders is being neglected in the geographic risk analysis in Fig. 9.

It is also important to stress that we use the same parameter μ across all federal states and thus assume a uniform detection probability. In reality, federal states with a large number of premises tend to enforce stricter hygiene and intervention standards so that the actual epidemic risk for states such as BY and NW is much lower. A realistic evaluation for public health must therefore include heterogeneous recovery (detection) probabilities on the level of states or individual nodes as discussed in Ref. [77] and Appendix A, as well as a complex disease response that

includes trade restrictions, increased awareness, and higher biosecurity.

APPENDIX C: FURTHER APPLICATIONS

From the detailed, node-level infection trajectory, we can estimate the infection arrival time from a given outbreak location to all remaining nodes. For that purpose, we extend the contact sequence periodically in time until the infection probabilities are negligible. Then, we derive the infection arrival time to a single node from the corresponding cumulative infection probability (see Fig. 2) as follows: (i) The discrete derivative of the cumulative infection probability gives the probability distribution to contract the infection at a given time step. (ii) The expectation value of the probability distribution gives the mean infection arrival time at a single node, corresponding to a scatter point in Fig. 14(a). Here, we compare the expected values from MC simulations with the estimated infection arrival times given by the CB and IB models, respectively. In a perfect prediction, the scattered values would lie on the diagonal, but, as the contact network is far from a treelike structure, the models estimate infection arrival times smaller than the observed values. The comparison between the IB and CB frameworks in Fig. 14(b) shows a considerably smaller relative deviation of CB estimations from the corresponding MC simulations for the given set of disease parameters and outbreak locations.

Another application focuses on the vulnerability of nodes with respect to a given outbreak location. Again, we assume an infinite time horizon and compare the cumulative probability that a node has been infected in the limit $t \to \infty$. As before, we find a good correlation between simulations and the estimated vulnerability in

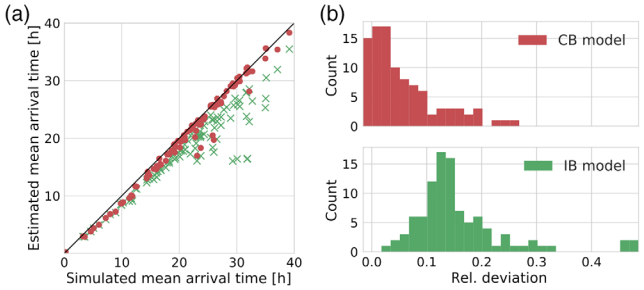


FIG. 14. (a) Comparison between simulated and estimated mean infection arrival times. We extend the data set periodically in time until the outbreak dies out. The discrete derivative of the cumulative infection probabilities (see Fig. 2) yields the infection arrival probabilities of which we take the average value for every node. Results according to the CB and IB models are visualized as red circles and green crosses, respectively. The epidemic starts from the same outbreak origin and disease parameters as in Fig. 1. (b) Histogram over the relative deviation from the simulated infection arrival times. The numerical values are averaged over 10^5 realizations.

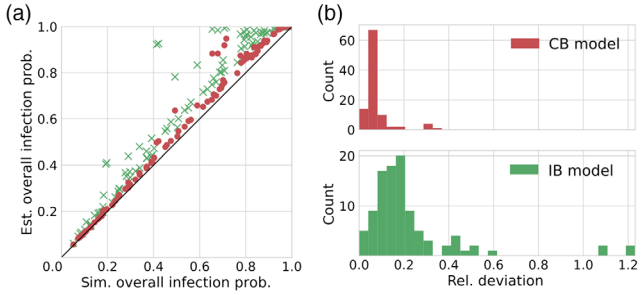


FIG. 15. (a) Comparison between simulated and estimated vulnerability. We compute the cumulative infection probability in the limit $t \to \infty$, also denoted as vulnerability. The comparison with CB and IB estimations is visualized by red circles and green crosses, respectively. Each value corresponds to the vulnerability of a node given the same outbreak location and disease parameters as in Fig. 1. (b) Relative deviation of the estimated values with respect to MC simulations. The numerical values are averaged over 10^5 realizations.

Fig. 15, whereas the CB model consistently outperforms the IB approach and overestimates the expected values surprisingly little given that the underlying aggregated network is fairly dense (the average degree is $\langle k \rangle \approx 19$) and far from being treelike.

1. Trajectories averaged over outbreak locations

For some applications, we may be interested in the trajectory of a global epidemic, averaged over outbreak locations. A sufficiently large number of initially infected nodes would then avoid complications with the early outbreak phase [39,78]. In this case, we adjust the MC simulations such that every node is infected independently with a given probability $1 - z_l = 1 - z$, $\forall\ l \in \mathcal{N}$ at $t = 0$. As for the analytic approach, we only need to set a corresponding homogeneous initial condition, and thus the computational complexity remains the same as in the previous case of one initially infected node.

In Fig. 16 (left column), we observe a narrow, time-dependent distribution of cumulatively infected nodes around the mean value for three different infection probabilities. Without applying any additional threshold, we find a close agreement between the averaged trajectory, and the CB and IB models in all cases. In contrast to Fig. 3 of the main text, we observe in Fig. 16 (right column) only one peak in the distribution due to the large number of initially infected nodes.

One potential application is to calculate the vulnerability of a node as discussed in Ref. [6]. Here, the vulnerability is defined as the probability that a given node is eventually infected by a disease that started somewhere in the network. The value can be used to rank nodes in order to prioritize surveillance or vaccination measures to the nodes that are most likely to contract the disease when resources are limited. In Fig. 17, every curve represents the vulnerability

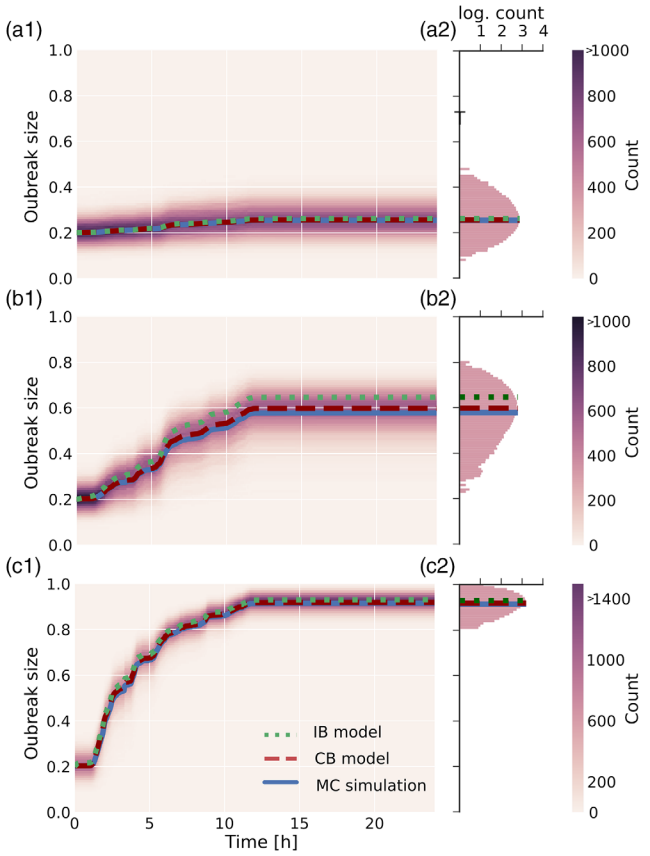


FIG. 16. Cumulative infection probability with a large fraction of initially infected nodes for three different values of β and the same outbreak location as in Fig. 3(a). Left column [panels (a1)–(c1)]: Time-evolving distribution (linear scale) of cumulatively infected individuals for infection probabilities $\beta = 0.01$, 0.02, and 0.1, respectively. We average over outbreak locations with 20% of the network initially infected at random. The mean outbreak size (blue line), averaged over 10^5 realizations and with a standard deviation below 10^{-4}, can thus be compared to the CB and IB models (red dashed and green dotted lines, respectively) with no threshold applied. Right column [panels (a2)–(c2)]: Final distribution (logarithmic scale) together with the averaged values.

of one node as a function of the infection probability β. These results are derived from the CB model, given an initial infection probability of $1 - z = 0.2$. The individual colors correspond to the degree of the node in the underlying time-aggregated graph and serve as a guide to the eye. Interestingly, we find that the ranking, as estimated by the CB model, may change with increasing infection probabilities β as can be seen from the highlighted curve in Fig. 17. This effect has been observed earlier in the context of static networks [6] and indicates that network properties alone are often not sufficient to rank nodes as they do not take into account details of the dynamic system.

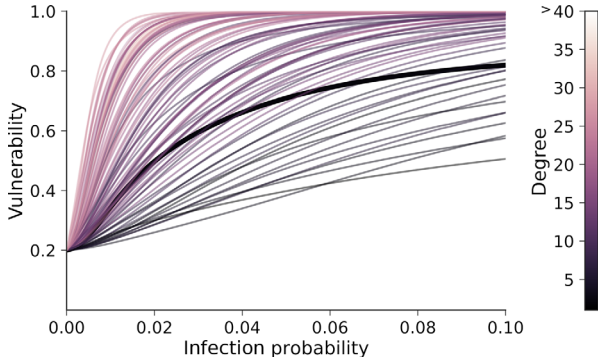


FIG. 17. Vulnerability as a function of the infection probability β estimated from the CB model. Each curve represents the vulnerability of a node, i.e., the probability to contract the infection from a set of randomly chosen outbreak locations. Here, we estimate the vulnerability according to the CB model. Starting from an initial infection probability of $1 - z = 0.2$, we propagate the infection over time until convergence. We stop when the largest increase in vulnerability after 24 h falls below 10^{-3}. The colors indicate the degree of each node in the underlying time-aggregated graph. Moreover, the behavior of one selected node is highlighted.

2. Additional numerical results

The analysis of the conference data set in the main text was limited to a single value of the recovery probability with $\mu = 2.85 \times 10^{-4}$. This choice corresponds to an expected infectious period of about 19.5 h. In addition to the analysis of the main text, we present in Fig. 18 similar results for different values of μ. The left, middle, and right columns correspond to Figs. 4, 6(a), and 6(b), respectively. In all cases, the CB model gives a closer bound to MC simulations as compared to the IB approach.

Next, we complete the analysis in Fig. 8 of the main text on the epidemic threshold for the example of the animal trade network. Again, we assume $\mu = 1/28$ and provide in Fig. 19 a detailed analysis of all federal states, excluding the city states Berlin, Hamburg, and Bremen. The left and middle columns in Fig. 19 provide a similar analysis to Figs. 6(a) and 6(b) of the main text, respectively. In other words, we present the distribution of outbreak sizes for different values of the infection probability β (left column) from which we derive the coefficient of variation c_v (blue line, middle column). The right column presents values of c_v that are close to the peak and a quadratic fit (green line, right column) that determines the numerical estimation of the critical infection probability (blue vertical line). This value can be compared to spectral estimations from the mean-field models. In agreement with previous results, we find that the criticality condition in Eq. (16) of the CB model (see main text) improves previous results of the IB approach.

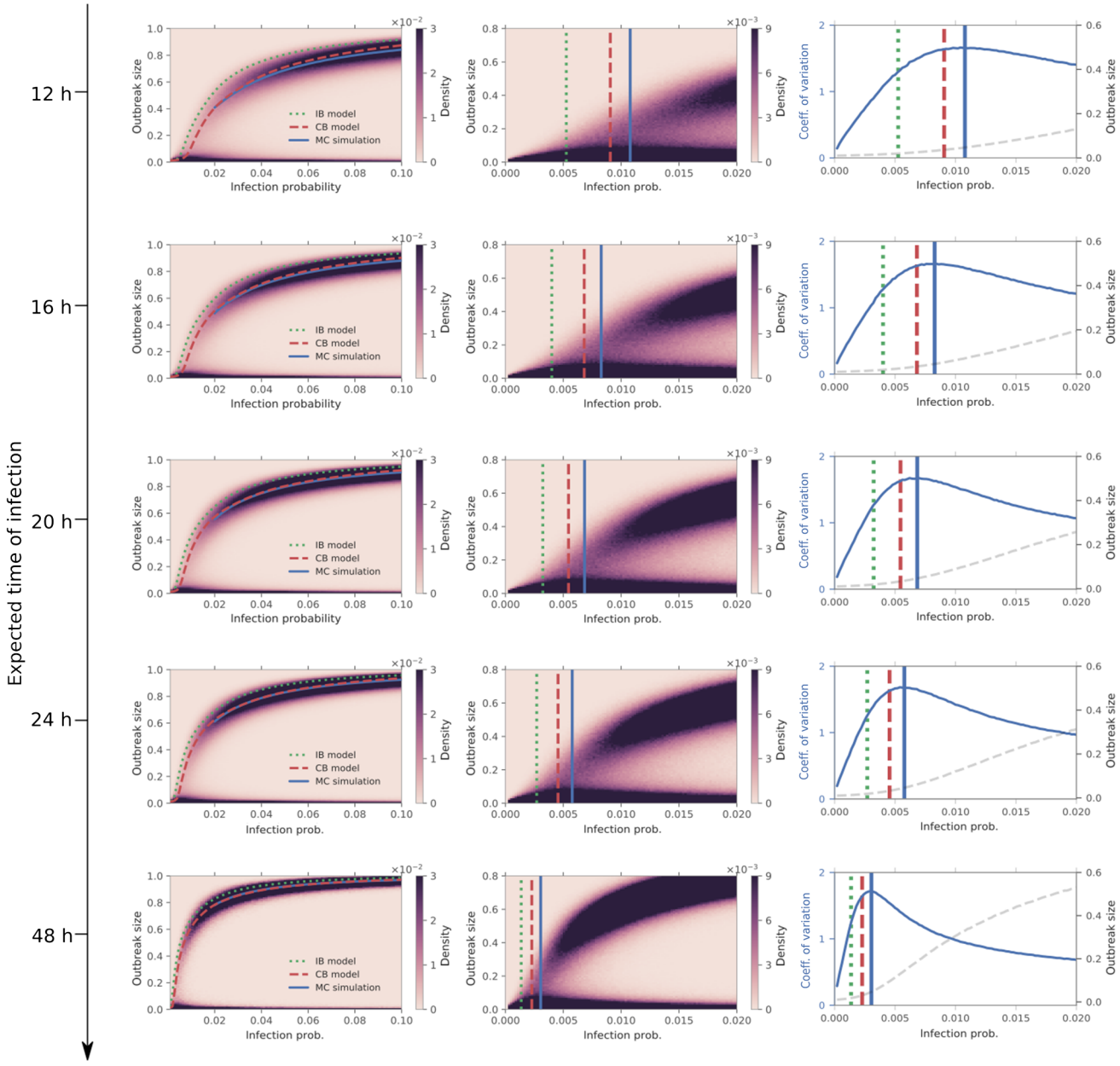


FIG. 18. Comparison between MC simulations and the mean-field models for different values of μ. Every row assumes a fixed recovery probability μ with decreasing values from top to bottom: $\mu = 4.63 \times 10^{-4}$, $\mu = 3.47 \times 10^{-4}$, $\mu = 2.78 \times 10^{-4}$, $\mu = 2.31 \times 10^{-4}$, and $\mu = 1.16 \times 10^{-4}$. These values correspond to an expected infection period of 12, 16, 20, 24, and 48 h. The left column, middle column, and right column correspond to Figs. 4, 6(a), and 6(b), respectively, of the main text.

Finally, we provide an additional analysis of the epidemic threshold for the cattle-trade data with $\mu = 1/120$. Results in Fig. 20 are akin to our previous analysis in Fig. 19, except for SL. Here, the spectral condition in Eq. (16) of the CB model predicts that every outbreak remains localized, i.e., $\beta_{\mathrm{crit}}^{\mathrm{CB}} = 1$, whereas MC simulations suggest a transition to global epidemics, hence $\beta_{\mathrm{crit}}^{\mathrm{MC}} < 1$. We attribute the inconsistency to the small size of the network (26 nodes). The spectral approach implicitly assumes an infinitely large network, which is clearly violated in this case.

We summarize the results for $\mu = 1/120$ in Fig. 21, akin to Fig. 9 of the main text. The risk map in Fig. 21(a) visualizes the spatial variability of the outbreak risk, and each group of blue, red, and green bars in Fig. 21(b) provides a quantitative comparison between MC results, the CB model, and the IB approach, respectively.

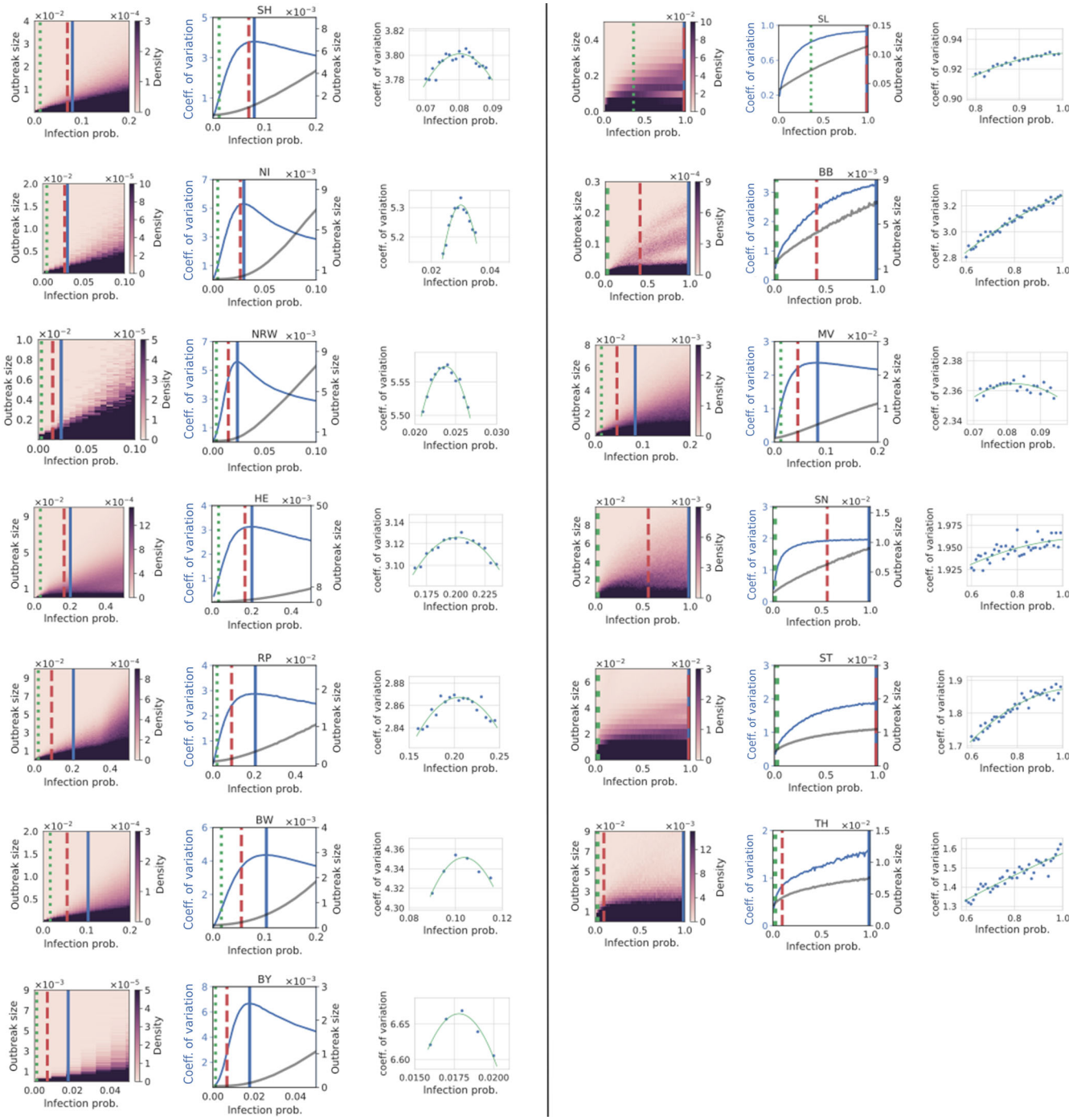


FIG. 19. Detailed analysis of the epidemic threshold for the cattle-trade network with $\mu = 1/28$. The central black line separates the figure into two panels. Every row (in a panel) provides results for a single federal state in Germany, excluding the city states Berlin, Bremen, and Hamburg. Left column: Outbreak-size distribution as a function of the infection probability β. Vertical blue, red, and green lines mark the critical infection probability according to MC simulations, and the CB and IB models, respectively. Middle column: Coefficient of variation c_v (blue line, left axis), mean outbreak size (green dashed line, right axis). Right column: Selected values of c_v for infection probabilities close to the critical value. The quadratic fit (green line) estimates the maximum and hence β_{crit}.

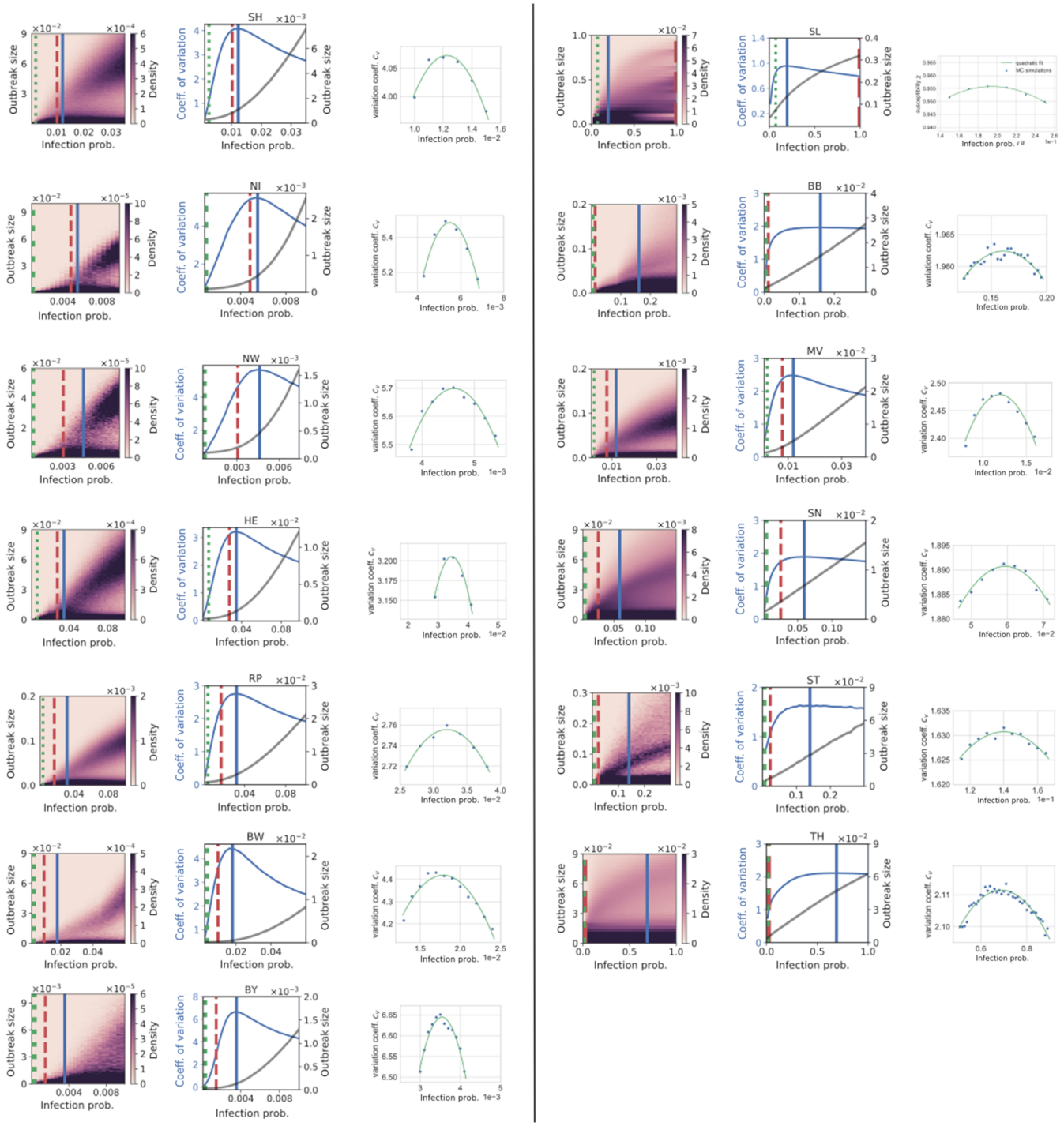


FIG. 20. Detailed analysis of the epidemic threshold for the cattle-trade network with $\mu = 1/120$. The analysis is akin to Fig. 19.

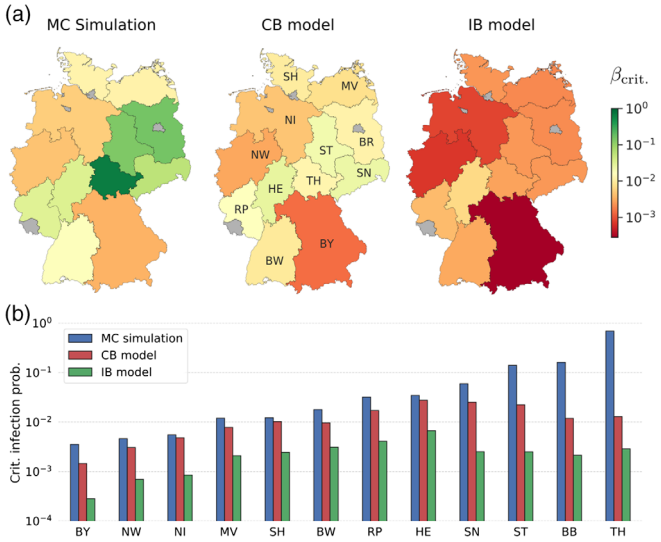


FIG. 21. (a) Spatial variation of epidemic risk for $\mu = 1/120$. The critical infection probability β_{crit} determines the color of the federal states (see Fig. 20 for details). The visualization is akin to Fig. 9 of the main text.

[1] D. Brockmann and D. Helbing, *The Hidden Geometry of Complex, Network-Driven Contagion Phenomena*, Science **342**, 1337 (2013).

[2] F. Iannelli, A. Koher, D. Brockmann, P. Hövel, and I. M. Sokolov, *Effective Distances for Epidemics Spreading on Complex Networks*, Phys. Rev. E **95**, 012313 (2017).

[3] A. L. Horn and H. Friedrich, *Locating the Source of Large-Scale Outbreaks of Foodborne Disease*, J. R. Soc. Interface **16**, 20180624 (2019).

[4] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Lage-Castellanos, and R. Zecchina, *Bayesian Inference of Epidemics on Networks via Belief Propagation*, Phys. Rev. Lett. **112**, 118701 (2014).

[5] J. T. Matamalas, A. Arenas, and S. Gómez, *Effective Approach to Epidemic Containment Using Link Equations in Complex Networks*, Sci. Adv. **4**, eaau4212 (2018).

[6] T. Rogers, *Assessing Node Risk and Vulnerability in Epidemics on Networks*, Europhys. Lett. **109**, 28005 (2015).

[7] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, *Analytical Computation of the Epidemic Threshold on Temporal Networks*, Phys. Rev. X **5**, 021005 (2015).

[8] E. Valdano, *Source Code of "Analytical Computation of the Epidemic Threshold on Temporal Networks*, github.com/eugenio-valdano/threshold.

[9] F. Altarelli, A. Braunstein, L. Dall'Asta, J. R. Wakeling, and R. Zecchina, *Containing Epidemic Outbreaks by Message-Passing Techniques*, Phys. Rev. X **4**, 021024 (2014).

[10] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, *On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models for Infectious Diseases in Heterogeneous Populations*, J. Math. Biol. **28**, 365 (1990).

[11] R. Pastor-Satorras and A. Vespignani, *Epidemic Spreading in Scale-Free Networks*, Phys. Rev. Lett. **86**, 3200 (2001).

[12] M. E. J. Newman, *Spread of Epidemic Disease on Networks*, Phys. Rev. E **66**, 016128 (2002).

[13] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, *Epidemic Thresholds in Real Networks*, ACM Trans. Inf. Syst. Secur. **10**, 1 (2008).

[14] J. C. Miller, *Spread of Infectious Disease through Clustered Populations*, J. R. Soc. Interface **6**, 1121 (2009).

[15] E. M. Volz and L. A. Meyers, *Epidemic Thresholds in Dynamic Contact Networks*, J. R. Soc. Interface **6**, 233 (2009).

[16] B. Karrer, M. E. J. Newman, and L. Zdeborová, *Percolation on Sparse Networks*, Phys. Rev. Lett. **113**, 208702 (2014).

[17] W. O. Kermack and A. G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*, Proc. R. Soc. A **115**, 700 (1927).

[18] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications*, Mathematics in Medicine Series (Griffin, London, 1975).

[19] P. L. Simon, M. Taylor, and I. Z. Kiss, *Exact Epidemic Models on Graphs Using Graph-Automorphism Driven Lumping*, J. Math. Biol. **62**, 479 (2011).

[20] P. Van Mieghem, J. Omic, and R. Kooij, *Virus Spread in Networks*, IEEE/ACM Trans. Netw. **17**, 1 (2009).

[21] I. Z. Kiss, G. Röst, and Z. Vizi, *Generalization of Pairwise Models to Non-Markovian Epidemics on Networks*, Phys. Rev. Lett. **115**, 078701 (2015).

[22] N. Sherborne, J. C. Miller, K. B. Blyuss, and I. Z. Kiss, *Mean-Field Models for Non-Markovian Epidemics on Networks*, J. Math. Biol. **76**, 755 (2018).

[23] B. Karrer and M. E. J. Newman, *Message Passing Approach for General Epidemic Models*, Phys. Rev. E **82**, 016101 (2010).

[24] S. Gonçalves, G. Abramson, and M. F. C. Gomes, *Oscillations in SIRS Model with Distributed Delays*, Eur. Phys. J. B **81**, 363 (2011).

[25] P. Van Mieghem and R. van de Bovenkamp, *Non-Markovian Infection Spread Dramatically Alters the Susceptible-Infected-Susceptible Epidemic Threshold in Networks*, Phys. Rev. Lett. **110**, 108701 (2013).

[26] R. M. May and A. Lloyd, *Infection Dynamics on Scale-Free Networks*, Phys. Rev. E **64**, 066112 (2001).

[27] M. J. Keeling and K. T. D. Eames, *Networks and Epidemic Models*, J. R. Soc. Interface **2**, 295 (2005).

[28] M. E. J. Newman, A. L. Barabási, and D. J. Watts, *The Structure and Dynamics of Networks* (Princeton University Press, Princeton, NJ, 2006).

[29] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks* (Springer, New York, 2017).

[30] M. A. Porter and J. P. Gleeson, *Dynamical Systems on Networks*, Frontiers in Applied Dynamical Systems: Reviews and Tutorials (Springer, New York, 2016), Vol. 4.

[31] A. Stopczynski, V. Sekara, P. Sapiezynski, A. Cuttone, M. M. Madsen, J. E. Larsen, and S. Lehmann, *Measuring Large-Scale Social Networks with High Resolution*, PLoS One **9**, e95978 (2014).

[32] V. Sekara, A. Stopczynski, and S. Lehmann, *Fundamental Structures of Dynamic Social Networks*, Proc. Natl. Acad. Sci. USA **113**, 9977 (2016).

[33] P. Holme and J. Saramäki, *Temporal Networks*, Phys. Rep. **519**, 97 (2012).

[34] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, *Time-Varying Graphs and Dynamic Networks*, in *Ad-hoc, Mobile, and Wireless Networks. ADHOC-NOW 2011*, edited by H. Frey, X. Li, S. Ruehrup, Lecture Notes in Computer Science, Vol. 6811 (Springer, Berlin, Heidelberg, 2011), pp. 346–359.

[35] L. E. C. Rocha, F. Liljeros, and P. Holme, *Information Dynamics Shape the Sexual Networks of Internet-Mediated Prostitution*, Proc. Natl. Acad. Sci. U.S.A. **107**, 5706 (2010).

[36] N. Eagle, A. Pentland, and D. Lazer, *Inferring Friendship Network Structure by Using Mobile Phone Data*, Proc. Natl. Acad. Sci. U.S.A. **106**, 15274 (2009).

[37] J. C. Delvenne, R. Lambiotte, and L. E. C. Rocha, *Diffusion on Networked Systems Is a Question of Time or Structure*, Nat. Commun. **6**, 7366 (2015).

[38] T. Gross, C. J. D. D'Lima, and B. Blasius, *Epidemic Dynamics on an Adaptive Network*, Phys. Rev. Lett. **96**, 208701 (2006).

[39] J. C. Miller, A. C. Slim, and E. M. Volz, *Edge-Based Compartmental Modelling for Infectious Disease Spread*, J. R. Soc. Interface **9**, 890 (2012).

[40] A. Koher, H. H. K. Lentz, P. Hövel, and I. M. Sokolov, *Infections on Temporal Networks—A Matrix-Based Approach*, PLoS One **11**, e0151209 (2016).

[41] H. H. K. Lentz, A. Koher, P. Hövel, J. Gethmann, T. Selhorst, and F. J. Conraths, *Livestock Disease Spread through Animal Movements: A Static and Temporal Network Analysis of Pig Trade in Germany*, PLoS One **11**, e0155196 (2016).

[42] A. Darbon, E. Valdano, C. Poletto, A. Giovannini, L. Savini, L. Candeloro, and V. Colizza, *Network-Based Assessment of the Vulnerability of Italian Regions to Bovine Brucellosis*, Prev. Vet. Med. **158**, 25 (2018).

[43] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, *Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint*, in *Proceedings of the 22nd International Symposium on Reliable Distributed Systems, Florence, Italy, 2003* (IEEE, New York, 2003), pp. 25–34.

[44] L. E. C. Rocha and N. Masuda, *Individual-Based Approach to Epidemic Processes on Arbitrary Dynamic Contact Networks*, Sci. Rep. **6**, 31456 (2016).

[45] A. Ganesh, L. Massoulié, and D. Towsley, *The Effect of Network Topology on the Spread of Epidemics*, in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, Vol. 2 (IEEE, New York, 2005), pp. 1455–1466.

[46] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, *Discrete-Time Markov Chain Approach to Contact-Based Disease Spreading in Complex Networks*, Europhys. Lett. **89**, 38009 (2010).

[47] M. Youssef and C. Scoglio, *An Individual-Based Approach to SIR Epidemics in Contact Networks*, J. Theor. Biol. **283**, 136 (2011).

[48] W. Wang, Q.-H. Liu, L.-F. Zhong, M. Tang, H. Gao, H. E. Stanley, *Predicting the Epidemic Threshold of the Susceptible-Infected-Recovered Model*, Sci. Rep. **6**, 24676 (2016).

[49] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Epidemic Processes in Complex Networks*, Rev. Mod. Phys. **87**, 925 (2015).

[50] J. P. Gleeson, J. A. Ward, K. P. O'Sullivan, and W. T. Lee, *Competition-Induced Criticality in a Model of Meme Popularity*, Phys. Rev. Lett. **112**, 048701 (2014).

[51] A. Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborová, *Inferring the Origin of an Epidemic with a Dynamic Message-Passing Algorithm*, Phys. Rev. E **90**, 012801 (2014).

[52] V. S. Bokharaie, O. Mason, and F. Wirth, *Spread of Epidemics in Time-Dependent Network*, in *Proceedings of the 19th International Symposium on Mathematical Theory and Network Systems* (2010), pp. 1717–1719, http://fwn06.housing.rug.nl/mtns/?page_id=10.

[53] B. A. Prakash, H. Tong, N. Valler, M. Faloutsos, and C. Faloutsos, *Virus Propagation on Time-Varying Networks: Theory and Immunization Algorithms*, in *Machine Learning and Knowledge Discovery in Databases*, edited by J. L. Balcázar, F. Bonchi, A. Gionis, and M. Sebag (Springer, Berlin, Heidelberg, 2010), pp. 99–114.

[54] E. Valdano, C. Poletto, and V. Colizza, *Infection Propagator Approach to Compute Epidemic Thresholds on Temporal Networks: Impact of Immunity and of Limited Temporal Resolution*, Eur. Phys. J. B **88**, 341 (2015).

[55] E. Valdano, A. Koher, J. Bassett, A. Darbon, P. Hövel, and V. Colizza, *GRAZE: A Generic and Standardized* PYTHON *Package for the Analysis of Livestock Trade Data* (work in progress).

[56] L. Speidel, K. Klemm, V. M. Eguiluz, and N. Masuda, *Temporal Interactions Facilitate Endemicity in the Susceptible-Infected-Susceptible Epidemic Model*, New J. Phys. **18**, 073013 (2016).

[57] A. Koher, *Source Code of "Contact-Based Model for Epidemic Spreading on Temporal Networks*, https://github.com/andreaskoher/Contact_Based_Epidemiology.

[58] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J. F. Pinton, and W. Van den Broeck, *What's in a Crowd? Analysis of Face-to-Face Behavioral Networks*, J. Theor. Biol. **271**, 166 (2011).

[59] The HIT database (Herkunftssicherungs- und Informationssystem für Tiere) is administered by the Bavarian State Ministry for Agriculture and Forestry on behalf of the German federal states, www.hi-tier.de (2012).

[60] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, NJ, 2008).

[61] C. R. Cai, Z. X. Wu, M. Z. Q. Chen, P. Holme, and J. Y. Guan, *Solving the Dynamic Correlation Problem of the Susceptible-Infected-Susceptible Model on Networks*, Phys. Rev. Lett. **116**, 258301 (2016).

[62] M. Shrestha, S. V. Scarpino, and C. Moore, *Message-Passing Approach for Recurrent-State Epidemic Models on Networks*, Phys. Rev. E **92**, 022821 (2015).

[63] I. Z. Kiss, C. G. Morris, F. Sélley, P. L. Simon, and R. R. Wilkinson, *Exact Deterministic Representation of Markovian SIR Epidemics on Networks with and without Loops*, J. Math. Biol. **70**, 437 (2015).

[64] A. Y. Lokhov, Dynamic Cavity Method and Problems on Graphs, Thesis, Université Paris Sud—Paris XI (2014).

[65] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson, *The Unreasonable Effectiveness of Tree-Based*

Theory for Networks with Clustering, Phys. Rev. E **83**, 036112 (2011).
[66] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Philadelphia, 2000).
[67] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, Philadelphia, 1994).
[68] E. Valdano, C. Poletto, A. Giovannini, D. Palma, L. Savini, and V. Colizza, *Predicting Epidemic Risk from Past Temporal Contact Data*, PLoS Comput. Biol. **11**, e1004152 (2015).
[69] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models* (Springer, New York, 2017).
[70] L. D. Valdez, P. A. Macri, and L. A. Braunstein, *Temporal Percolation of the Susceptible Network in an Epidemic Spreading*, PLoS One **7**, 1 (2012).
[71] R. R. Wilkinson, F. G. Ball, and K. J. Sharkey, *The Relationships between Message Passing, Pairwise, Kermack–McKendrick and Stochastic SIR Epidemic Models*, J. Math. Biol. **75**, 1563 (2017).
[72] F. Radicchi and C. Castellano, *Beyond the Locally Treelike Approximation for Percolation on Real Networks*, Phys. Rev. E **93**, 030302(R) (2016).
[73] I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C. J. Tessone, and F. Schweitzer, *Causality-Driven Slow-Down and Speed-Up of Diffusion in Non-Markovian Temporal Networks.*, Nat. Commun. **5**, 5024 (2014).
[74] R. Lambiotte, M. Rosvall, and I. Scholtes, *From Networks to Optimal Higher-Order Models of Complex Systems*, Nat. Phys. **15**, 313 (2019).
[75] Data are available on http://sociopatterns.org.
[76] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *The Architecture of Complex Weighted Networks*, Proc. Natl. Acad. Sci. U.S.A. **101**, 3747 (2004).
[77] A. Darbon, D. Colombi, E. Valdano, L. Savini, A. Giovannini, and V. Colizza, *Disease Persistence on Temporal Contact Networks Accounting for Heterogeneous Infectious Periods*, R. Soc. open sci. **6**, 181404 (2019).
[78] J. C. Miller, *Epidemics on Networks with Large Initial Conditions or Changing Structure*, PLoS One **9**, 1 (2014).