Mapping spreading dynamics: From time respecting shortest paths to bond percolation

Dijana Tolić,1 Kaj-Kolja Kleineberg,2 and Nino Antulov-Fantulin*2
1 Rudjer Bosković Institute, Laboratory for Machine Learning and Knowledge Representations, Bijenička Cesta 54, 10 000 Zagreb, Croatia
2 ETH Zürich, Swiss Federal Institute of Technology, COSS, Clausiusstraße 50, 8092 Zurich, Switzerland
(Dated: September 12, 2017)

We propose a mapping of spreading dynamics to an ensemble of weighted networks, where edge weights represent propagation time delays. In this mapping, shortest paths in the weighted networks preserve the temporal causality of spreading. Furthermore, for efficient sampling, we construct a Markov Chain (Gibbs sampler) over elements of an ensemble of mapped weighted networks. Our framework provides insights into the local and global spreading dynamics from arbitrary source nodes and the scaling of the average propagation time for Markovian and non-Markovian processes. Furthermore, it enables efficient source detection and helps to improve strategies for time-critical vaccination. Our framework overcomes the limitations of previous methods such as tree-like assumptions of message passing, omitting dynamical correlations with mean-field approximations, or setting all initial conditions upfront for the kinetic Monte Carlo method. Finally, in a limit of process time, we establish the connection of our mapping with bond percolation.

I. INTRODUCTION

Spreading processes [1–3], like the spread of infectious diseases, information, or failures, naturally occur in socio-technological systems. A broad class of these systems can be described as networks, whose connectivity patterns have been shown to be heterogeneous [4]. Whereas infinite time properties are well understood [2–3], the actual dynamical properties of spreading processes on complex networks still lack understand-
ing. In order to exactly solve the spreading process one must be able to evaluate the evolution of the process configurations. The evolution of the probabilities of configurations is given as a solution to a set of coupled differential equations called master equation [5–6], which cannot be solved directly in most cases. Therefore, different approximations are used: (i) neglecting dynamical correlations by assuming statistical independence at first 
nhood, pair or node level in mean field approxi-
mation [7–10], (ii) neglecting loops in the network structure with message passing or belief propagation techni-
ces [11–13], or (iii) neglecting the temporal evolution by mapping to the percolation theory [14–16].

In this paper, we propose how to exactly transform spreading dynamics to a shortest path problem on an ensemble of weighted networks using a Monte Carlo approach. In this framework, weights represent interaction time delays along edges. We are able to extract a temporal distance measures that describes the expected time the spreading process needs to propagate from one node to another. The mapping is applicable to the generalized Susceptible Infected Recovered (SIR) model [17] without memory (exponential inter-event distribution/Poisson process) and with memory (arbitrary inter-event distributions). Contrary to the message passing methods [11–13], our method does not neglect loops in the network structure. We show the equivalence between the time respecting paths (shortest paths) on the constructed weighted networks and the propagation times of spreading dynamics. Our results provide new insights into the global and local spreading dynamics on complex networks and have direct applications for source detection and time-critical vaccination. Finally, in a limit of process time, we establish the connection with bond percolation theory [14–16].

II. MAPPING OF SPREADING DYNAMICS

Let us start with the Poissonian SIR model, which is memoryless, such that the next state only depends on the present state of the system. This process is defined by two parameters ($\beta, \gamma$), which describe transitions where susceptible nodes become infected with rate $\beta$ from infected neighbours and transitions where infected nodes recover with rate $\gamma$. This process has exponential inter-event time distributions: $\psi(t) = e^{-\beta t}$ (spread) and $\phi(t) = e^{-\gamma t}$ (recovery). However, many realistic processes have memory [20] and their inter-event time distributions are non-exponential. Therefore, we allow for arbitrary inter-event density distributions ($\psi(t), \phi(t)$) for generalized SIR processes. For a given network $G$ and class of generalized SIR spreading models, we show how to create weighted networks $\{G_k\}$ which encode realizations of the stochastic spreading dynamics. Each weighted network $G_k$ (see Fig. 1h) encodes one possible outcome of the spreading process from every potential source node in a network. In particular, a time-respecting weighted network instance $G_k$ is created by taking the input network $G$ and assigning weights to the edges of the network instance with the Inverse Smirnov transform [21].
\[ \rho_{i,j} = \begin{cases} \Psi^{-1}(x) : & \Psi^{-1}(x) \leq \Phi^{-1}(y) \\ \infty : & \Psi^{-1}(x) > \Phi^{-1}(y) \end{cases} \]

where \( x \) and \( y \) are uniform random numbers in \([0, 1]\), \( \Phi^{-1}(x) \) and \( \Psi^{-1}(y) \) are inverse functions of the cumulative inter-event distributions \( \Phi(t) = \int_0^t dt' \phi(t') \) and \( \Psi(t) = \int_0^t dt' \psi(t') \). The quantities \( \Psi^{-1}(y) \) and \( \Phi^{-1}(x) \) respectively represent the samples of the transmission and recovery time obtained with the Inverse Smirnov transform of inter-event distributions. In the special case of the Poissonian SIR process one obtains

\[ \rho_{i,j} = \begin{cases} -\ln(x)/\beta : & -\ln(x)/\beta \leq -\ln(y)/\gamma \\ \infty : & -\ln(x)/\beta > -\ln(y)/\gamma \end{cases} \]  

If nodes recover faster than a transmission occurs, the weight is set to infinity to indicate no transmission through the edge.

We denote the distance as shortest paths on weighted networks, \( d_{\chi_{ij}}(v_i, v_j) = \min_{v_j \in \chi_{ij}} \sum_{(k,l) \in \chi_{ij}} \rho_{k,l} \), where \( \chi_{ij} \) is the set of all possible paths from node \( v_i \) to node \( v_j \) on network \( G_k \) and \( \rho_{k,l} \) denotes the weights defined in Eq. (1). Importantly, this distance is equivalent to the propagation time from node \( i \) to \( j \), i.e. \( t(v_i \to v_j) = d_{\chi_{ij}}(v_i, v_j) \) (time respecting path equivalence, see Supplementary Material, Sec. 1). This holds for \( \beta/\gamma \gg 1 \) i.e. rapid spreading (see Supplementary Material, sec. 3), where one can neglect dynamical correlations and approximatively for \( \beta \approx \gamma \). The exact mapping including dynamical correlations is explained in Appendix D.

Instances of the ensemble of weighted networks can be generated independently, or by traversing elements of the ensemble with Markov Chain. The Markov Chain (rejection-free Gibbs sampler – see Supplemental Material sec. 2) consists of transitions \( W_{i,j} = w(G_i \to G_j) \) over the set of weighted networks \( \{G_i\} \), where each weighted network corresponds to one realization of the spreading process from every potential source node. This allows us to construct transitions between weighted networks by changing the weights on a randomly selected edge by assigning a new weight according to Eq. (1) (see Fig. 1). The existence of a stationary distribution of this Markov Chain is guaranteed by detailed balance property and the uniqueness by ergodicity (see Appendix B). After each weight change, all to all shortest paths can be dynamically recalculated \([24, 25]\) with computational complexity \( O(N^2 \log^2 N) \), where \( N \) denotes the number of nodes in a network (see Supplemental Material, Sec. 4 for more details). In other words, the Gibbs sampling—as well as independent sampling—constructs the probability distribution \( P(\mathcal{G}) \) of the weighted network ensemble, such that it is a statistically exact representation of the stochastic process.

From the ensemble of weighted networks, due to ergodicity, the expectation \( \langle f(\mathcal{G}) \rangle \) can be estimated as the
average over \( n \) samples of weighted networks:

\[
\langle f(\mathcal{G}) \rangle = 1/n \sum_{k=1}^{n} f(\mathcal{G}_k) \approx 1/n \sum_{k=1}^{n} f(\mathcal{G}_k) \tag{3}
\]

By changing functions \( f(\mathcal{G}) \), we estimate different properties such as total outbreak size, temporal evolution, expected propagation time or source likelihood (see Appendix C).

III. TEMPORAL MAPS – FROM NODE TO NETWORK PERSPECTIVE

Starting from a given node, what is the expected time after which the disease reaches each other node in the network? The answer to this question is encoded in the ensemble of weighted networks. For SI processes (\( \gamma = 0 \)), the expected propagation time is finite and can be estimated as the average over \( n \) samples of weighted networks as

\[
D_{ij} = 1/n \sum_{k=1}^{n} f(\mathcal{G}_k)
\]

where \( f(\mathcal{G}_k) = d_{\mathcal{G}_k}(v_i,v_j) \) denotes the shortest path. In Fig. 1, we show an explicit example from the perspective of a specific source node. The advantage of the temporal maps we have created is they encode this type of information for all possible source nodes. The result is shown in Fig. 2 for three empirical networks: the email network [22], the airport network [24], and the Petser online social network [23]. The plots show the average temporal distance or expected propagation time \( D_{ij} \) for each pair of nodes \( i,j \), and therefore encompass information about the expected spreading dynamics. In particular, we have ordered the nodes by their characteristic spreading timescale, which we define as the expected time after which a given node infects half of the network, \( \tau_i \), as

\[
\tau_i = t\sum \Theta(t-D_{ij}) = N/2,
\]

where \( \Theta(\cdot) \) denotes the Heaviside step function, which is one when \( t \geq D_{ij} \) and zero otherwise. A small value of \( \tau_i \) means that node \( i \) is of high importance or influence.

FIG. 3. Source detection. We compare source detection performance for different methods for the SIR model \( \beta = 0.7, \gamma = 0.3 \) (Fig. a) and \( \beta = 0.7, \gamma = 0.7 \) (Fig. b) at time \( T = 5 \) (discrete time) on the 4-connected lattice (30x30 nodes). For each observed realization, we rank nodes according to their estimated probability score according to the Direct Monte Carlo method [25], in case of a tie we prioritize the node with a higher estimator according to the temporal distance method. We compare different methods as explained in the text: Direct Monte Carlo source probability ranking, source detection ranking by topological distance [24], and the estimation from our mapping (temporal distance). Results are averaged over 30 observed realizations and error bars show one standard deviation, from top to bottom.

In Fig. 3 we observe that, apart from local fluctuations, the average global picture seems to follow a trend which in the figure corresponds to a smooth gradient of the color. We visualize this behavior with the contour lines that are superposed in the figure, and which we created by generating a coarse grained version of the matrices \( D_{ij} \). The contour lines therefore separate regions in which a pair of nodes with the characteristic spreading timescales \( \tau \) and \( \tau' \) has, on average, a temporal distance lower than some value \( D^* \). We take the coincidence of this coarse grained measure with the microscopic temporal distances as evidence for the appropriateness of the definition of the characteristic spreading timescale.
The question how the average propagation time between two individuals scales with the system size is crucial for many real applications and of special theoretic interest. In particular, if this distance scales logarithmically with the system size, an epidemic can spread nearly instantaneously in arbitrary large networks. However, if it scales like \( N^\beta \), in large enough systems the disease will not spread through the network in realistic timescales. We can address this question for the SI-model (\( \gamma = 0 \)) in an elegant fashion by establishing an analogy to the optimal path problem in disordered networks [30–32], which are generated by assigning weights \( \rho \) to edges drawn from some distribution \( f(\rho) \). In [30] the authors have shown that the optimal path length \( \langle d \rangle \) in these weighted networks scales differently with the system size depending on \( f(\rho) \), and distinguish between strong and weak disorder. For the example of Erdos-Renyi networks, in the case of strong disorder the optimal path length scales as \( \langle d \rangle \propto N^{1/3} \), and for weak disorder as \( \langle d \rangle \propto \log N \). The temporal distance we have defined earlier is equivalent to the optimal paths. Therefore, we can use this analogy to study the scaling of the average propagation time with the system size. From [30] it is known that if \( f(\rho) \) is taken as the exponential distribution, only weak disorder can occur, and the average distance scales with the logarithm of the system size for Erdos-Renyi networks. This means that for the memory-less spreading process, where \( \psi \) is the exponential distribution, the average propagation time scales with the logarithm of the system size and diseases spread quickly in systems of any size. However, if the inter-event distribution is for instance given by a lognormal distribution (non-Markovian case), we know from [30] that strong disorder can occur such that the average distance and hence the average propagation time scales like \( \langle d \rangle \propto N^{1/3} \). In this case, the finite time properties of the spreading process are drastically different to the non-Markovian case, as the average propagation time becomes very large as the size of the system increases, such that the disease will not spread globally in realistic timescales.

To conclude, although it has recently been shown that the infinite time steady states of non-Markovian processes can be mapped to Markovian ones [33], their finite time dynamics can be dramatically different.

IV. APPLICATIONS

Our framework provides a series of direct applications. Here, we discuss three possible applications: computationally efficient source detection, time critical vaccination, and node influence ranking.

A. Source detection

For a given snapshot \( r^* \) of a spreading process at time \( t_0 \) on a network, how does one detect the source node that generated the snapshot? Source detection [12, 13, 26, 28, 35, 36] is an important theoretical and practical problem in network science. It is at least as hard as any problem in NP-complete class [35, 37], because it belongs to \#P-complete class [38] and we approach it by probabilistic kernel estimations [28, 39] of the likelihood that source node generated the observed snapshot. Our framework can be used to perform this task, as the ensemble of weighted networks encompasses information about different realizations of the spreading process from all possible sources.

For the observed realization \( r^* \), our source probability estimates for every node \( v_i \) are constructed from ensemble averages \( s_{ij} \propto \sum_m f(\mathcal{G}_j) \). From weighted network \( \mathcal{G}_j \), we extract realization based on temporal distances and calculate the similarity \( \varphi \in [0, 1] \) to the observed realization \( r^* \). Similarity \( \varphi \) is the normalized number of equal corresponding states between two realizations. Finally,
we use the Gaussian kernel function
\[ f(G_j) = \exp(-(\varphi_j - 1)^2/a^2) \] for probability estimations [28] [39] (see Appendix C).

We compare our source probability estimations with direct Monte Carlo approach [28] and topological distance approach [29]. The direct Monte Carlo approach generates large number of simulations \(n = 10^5 - 10^6\), from each potential source node \(v_i\), and source probability estimation \(s_i \propto \sum \tau_k\) where \(n_i\) is the number of simulations that match the observed realization \(r^*\). The topological distance approach calculates the average topological distance \(\langle d_i \rangle\) from each potential source \(v_i\) to other infected nodes in the observed realization \(r^*\) and source estimation \(s_i \propto \langle d_i \rangle^{-1}\). The result is shown in Fig. 3a, where source ranked probability estimates for different methods are shown. Indeed, the method based on the ensemble of weighted networks \((n = 10^5\) instances) performs better than a topological distance estimation \([29]\) and correctly assigns a high probability to the source candidate identified by the direct Monte Carlo method [28].

B. Time critical vaccination

Now, let us consider the following vaccination scenario. We observe at time \(t_0\) the outbreak of a disease and we have a certain amount, \(m\), of vaccines (see Fig. 4b). We can vaccinate \(m\) individuals instantaneously, but the vaccines will only be effective after some time \(\Delta t\) (a vaccinated individual can get infected before \(t_0 + \Delta t\), and in this case this vaccine portion is wasted, see Fig. 4b). Who shall we vaccinate? Our framework can be used to improve the vaccination strategy. The idea is that, at time \(t_0\), we vaccinate only individuals who—with high probability—will not get infected before \(t_0 + \Delta t\). Therefore, we estimate the probability that a node \(i\) will not be infected before \(t_0 + \Delta t\) as
\[ \bar{p}_i = 1/n \sum_{k=1}^{n} \Theta(d_{G_k}(v_\theta, v_i) - t_0 - \Delta t), \] where \(\theta\) denotes the source node and \(\Theta(\cdot)\) is one when \(d_{G_k}(v_\theta, v_i) \geq t_0 + \Delta t\) and zero otherwise. We vaccinate nodes proportional to \(\bar{p}\). In Fig. 4b we show using Petset network that this procedure performs better than randomly choosing from all susceptible nodes at \(t_0\). Interestingly, vaccinating proportional to degree (hubs strategy) performs even worse than the fully random strategy, since the hubs usually are infected earlier, which increases the chance to waste vaccine portions.

V. EQUIVALENCE TO BOND PERCOLATION

In the following, we state the connection to bond percolation theory. From the perspective of source node \(v_i\), the shortest path \(d(v_i, v_j)\) indicates the first infection time of node \(v_j\). Therefore all nodes that are not reachable within the temporal distance \(t\) stay in the susceptible state. If a node is reachable within the temporal distance \(t\), then it was infected. By letting \(t\) go to infinity we obtain asymptotic realization equivalence to bond percolation. In particular, the size of the set of nodes \(v_j\) that are reachable in any finite time from the source node \(v_i\), \(|S_p(v_i)|\), is given by
\[ |S_p(v_i)| = \lim_{t \to \infty} |\{d(v_i, v_j) < t\}|. \] The full proof is given in the Appendix A.

Hence, \(|S_p(v_i)|\) denotes the size of the connected bond percolation component, where the transmissibility parameter \([18]\) from source \(v_i\) is given by
\[ p = \int_0^\infty d\tau \phi(\tau) \int_0^\tau d\tau' \psi(\tau'). \] The transmissibility quantifies the probability that an infected node transmits infection along a link to a susceptible neighbour before it recovers. The second integral accounts for the conditional probability of transmission up to the fixed recovery time \(\tau\), and the first integral account for all possible recovery times \(\tau\). Furthermore, to fix dynamical correlations \((\beta \approx \gamma)\) that were missing in [18], we generalize and formalize the transmissibility for the first neighborhood or, more precisely, the probability \(p_{n,k}\) that \(k\) out of \(n\) directed links will be active
\[ p_{n,k} = \int_0^\infty \phi(\tau) d\tau \binom{n}{k} \left(1 - \Psi(\tau)\right)^{(n-k)} \Psi(\tau)^k, \] where \(\Psi(\tau) = \int_0^\tau \psi(t) dt\). For Poissonian process this becomes
\[ p_{n,k} = \binom{n}{k} \frac{\Gamma(k+1)\Gamma\left(\frac{\beta}{\beta}(n-k)+1\right)}{\Gamma(k+1+\frac{\beta}{\beta}(n-k))}. \]
VI. DISCUSSION

We have mapped spreading dynamics on complex networks to an ensemble of weighted networks. This mapping allows us to define a temporal distance as shortest paths between nodes on an ensemble of weighted networks. We concentrate on the stochastic formulation of the generalized continuous and discrete time Susceptible Infected Recovered (SIR) spreading dynamics without memory (exponential inter-event distribution) and with memory (arbitrary inter-event distributions).

Specifically, we overcome the limitations of previous methods such as message passing, mean-field approximations and kinetic Monte Carlo methods. In contrast to message passing, our framework is applicable for arbitrary network structures without neglecting loops in the network structure. Furthermore, our method takes into account dynamical correlations between node states overcoming the assumptions of mean-field like approximations \[8, 10\]. In contrast to the well known historical Gillespie or kinetic Monte Carlo methods \[20, 28, 40–44\], we do not have to specify initial conditions upfront and we can sample new realizations from the previous ones by making local random perturbations on the weighted networks. The average temporal distance provides insights into the local and global dynamical properties of the spreading process. We have shown that for some non-Markovian processes, this average propagation time scales as a polynomial with the system size, even if the underlying network is small world. In stark contrast to the Markovian case, this means that for large enough systems the disease does not spread globally in a realistic timescale. We have shown that our framework allows for direct applications like source detection, time-critical vaccination, and a ranking of nodes according to their importance for the spreading dynamics. Combining our mapping framework with existing vaccination strategies for direct applications like source detection, time-critical vaccination, and a ranking of nodes according to their importance for the spreading dynamics. Combining our framework allows us to define a temporal distance as shortest paths between nodes on an ensemble of weighted networks. This mapping allows us to define a temporal distance as shortest paths between nodes on an ensemble of weighted networks. We concentrate on the stochastic formulation of the generalized continuous and discrete time Susceptible Infected Recovered (SIR) spreading dynamics without memory (exponential inter-event distribution) and with memory (arbitrary inter-event distributions).

Note that previous studies \[20, 25, 47\] for estimating the spreading arrival times were concentrated on a global disease spread with different approximations and theoretical understanding of effective distances \[23\] for metapopulation models. However, the exact mathematical equivalences for mapping generalized spreading dynamics at a microscopic level were still lacking. Furthermore, in a limit of process time, we establish the connection with bond percolation theory. Previously the outcome in finite time of the SIR process have been successfully mapped to the percolation \[14, 15, 18, 19\] but not for the finite time. Which bridges the gap between inference problems and percolation theory. Contrary to previous state-of-the-art source inference studies \[12, 13, 28\], estimations with our method require same computational effort to sample realizations at an arbitrary point in time \(t\) and conduct inference not only for discrete time but also for continuous time dynamics. Finally, our mapping to weighted networks naturally provides a dynamical interpretation of the optimal paths problem in disordered networks \[30, 31\].

APPENDIX A: BOND PERCOLATION LIMIT EQUIVALENCE

Here, we prove that, by letting \(t\) go to infinity, we obtain a realization equivalent to the normal bond percolation realization, to which we refer as bond percolation limit equivalence:

\[
|S_p(v_i)| = \lim_{t \to \infty} |\{d(v_i, v_k) < t\}|
\]

\(|S_p(v_i)|\) denotes the size of the set of nodes \(v_k\) that are reachable by any finite shortest path length from the source node \(v_i\). The term \(|S_p(v_i)|\) denotes the size of the connected bond percolation component with transmissibility \(p\) parameter from a source \(v_i\) \[18\].

Next, we show that our mapping establishes links with finite length \(\rho_{ij} < \infty\) with the probability which is equal to the transmissibility \(p\) \[18\]:

\[
p = \int_{0}^{\infty} \gamma e^{-\gamma t} dt \int_{0}^{t} \beta e^{-\beta t} dt = \frac{\beta}{\beta + \gamma}.
\]

In our mapping, the link with finite length \(\rho_{ij} < \infty\) is formed when \(-\ln(x)/\beta \leq -\ln(y)/\gamma\), for \(x\) and \(y\) as uniform random numbers \(\in [0, 1]\).

\[
P(-\ln(x)/\beta \leq -\ln(y)/\gamma) = P(\ln(x) \geq \frac{\beta}{\gamma} \ln(y)) = P(x \geq y^{\frac{\beta}{\gamma}}) = \int_{y^{\frac{\beta}{\gamma}}}^{\infty} \int_{y^{\frac{\beta}{\gamma}}}^{\infty} f_{x,y}(x,y)dxdy = \int_{y^{\frac{\beta}{\gamma}}}^{\infty} f_{y}(y)dy
\]

where \(f_{x,y}(x,y)\) is joint density function of variables \(x\) and \(y\). As they are independent, \(f_{x,y}(x,y) = f_{x}(x)f_{y}(y)\). As both \(x\) and \(y\) are uniform random numbers \(\in [0, 1]\), their densities are \(f_{x}(x) = 1_{[0,1]}\) and \(f_{y}(y) = 1_{[0,1]}\).

\[
P(x \geq y^{\frac{\beta}{\gamma}}) = \int_{-\infty}^{\infty} 1_{[0,1]}(y) \left( \int_{y^{\frac{\beta}{\gamma}}}^{\infty} 1_{[0,1]}(x)dxdy \right) dy = \int_{0}^{1} \left( \int_{y^{\frac{\beta}{\gamma}}}^{\infty} dx \right) dy = \frac{\beta}{\beta + \gamma}
\]
which is equal to transmissibility parameter $p$ \cite{18}. Therefore, we conclude that both bond percolation and time augmented bond percolation are the equivalent stochastic processes for the formation of edges when the finite edge weights are taken into account.

**APPENDIX B: MARKOV CHAIN CORRECTNESS**

We are interested in Markov Chain with state space: $\Omega = \{G_1, G_2, ..., G_i, ...\}$, with transitions $P(G_i \rightarrow G_j)$ between weighted graphs.

Note that graphs $G$ have continuous non-negative real weights and that the density function for weight $\rho$ of a specific link is $f(\tau, \rho) = \phi(\tau)(\psi(\rho)1_{[0,\tau]}(\rho) + \delta(\rho - \infty) \int_{\tau}^\infty d\psi(\rho))$, where $1_{[0,\tau]}(\rho)$ is the identity function which is equal to one on a range $[0, \tau)$ and $\delta(t - \infty)$ is the Dirac delta distribution. Before proceeding, we will use the discrete approximation of the non-negative weights i.e. density functions $\phi(\tau)$, $\psi(\rho)$, $f(\tau, \rho)$ become probability mass functions $g_\phi(\tau)$, $g_\psi(\rho)$, $g_{\phi, \psi}(\tau, \rho)$. This approximation essentially tells us that we use the finite number of possible non-negative weights to approximate density distributions, which is the case for all numeric simulations of continuous functions. This enables us to use the formalism of Markov Chain for discrete state space.

By design, the Markov Chain has stationary distribution $P(G)$, which is equal to the probability that this graph is generated by our mapping:

$$P(G) = \prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l}).$$

(15)

The probability of a particular instance weighted graph due to the independence is a product of probabilities over weights on edges generated by our mapping. The probability of having a weight $\tau_{i,j}$ is $g_\phi(\tau_{i,j})$ and probability of having a weight $\rho_{i,j}$ for a given $\tau_{i,j}$ is $g_{\phi, \psi}(\rho_{i,j} | \tau_{i,j})$.

**Convergence to stationary state**

The existence of stationary distribution of the Markov Chain in discrete time and discrete states is guaranteed by detailed balance property and the uniqueness by ergodicity.

Next, we show why the detailed balance holds if we choose the transition $P(G_i \rightarrow G_j)$ such that the weight of the randomly selected edge $(i^*, j^*)$ is changed by generating new weight with our mapping.

We write the detailed balance condition:

$$P(G_i)P(G_i \rightarrow G_j) = P(G_j)P(G_j \rightarrow G_i).$$

(16)

The probability of generating the particular graph $G_i$ is:

$$P(G_i) = \prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l}),$$

(17)

where $\rho_{k,l}$ denotes the transmission weight on edge $(k, l)$ in $G_i$ and $\tau_{k,l}$ denotes the recovery time for edge $(k, l)$ in $G_i$. Similarly, for graph $G_j$ we can write the probability $P(G_j)$. Now, the transition between $G_i$ and $G_j$ happens by selecting the random edge $(m^*, n^*)$ with uniform probability $c$ and changing the weight on it.

$$P(G_i \rightarrow G_j) = cg_\phi(\tau_{m^*, n^*}) g_{\phi, \psi}(\rho_{m^*, n^*} | \tau_{m^*, n^*}) \times$$

$$\prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l}).$$

(18)

We can extract the edge $(m^*, n^*)$ in $G_i$ from equation (17):

$$P(G_i) = g_\phi(\tau_{m^*, n^*}) g_{\phi}(\rho_{m^*, n^*} | \tau_{m^*, n^*}) \times$$

$$\prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l}).$$

(19)

and similarly for graph $G_j$:

$$P(G_j) = g_\phi(\tau_{m^*, n^*}) g_{\phi}(\rho_{m^*, n^*} | \tau_{m^*, n^*}) \times$$

$$\prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l}).$$

(20)

Now, the ratio is

$$\frac{P(G_i)}{P(G_j)} = \frac{g_\phi(\tau_{m^*, n^*}) g_{\phi}(\rho_{m^*, n^*} | \tau_{m^*, n^*}) \times \frac{\prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l})}{\prod_{A_{\phi, \psi}(\tau, \rho)} g_{\phi}(\rho_{k,l} | \tau_{k,l}) g_\phi(\tau_{k,l})}}$$

(21)

since the graphs $G_i$ and $G_j$ share all the weights except on the randomly selected edge $(m^*, n^*)$. Next, we see that this ratio is equal to the ratio of transitions rates :

$$\frac{g_\phi(\tau_{m^*, n^*}) g_{\phi}(\rho_{m^*, n^*} | \tau_{m^*, n^*})}{g_\phi(\tau_{m^*, n^*}) g_{\phi}(\rho_{m^*, n^*} | \tau_{m^*, n^*})} = \frac{P(G_j \rightarrow G_i)}{P(G_i \rightarrow G_j)}$$

(22)

and we conclude that the detailed balance holds.

**Ergodicity**

The second thing that needs to be proved is ergodicity. We know, that Markov Chain is ergodic if it is irreducible, aperiodic and positive recurrent. A finite state Markov chain is irreducible and positive recurrent if there’s a finite number $N$ such that any state can be reached from any other state in exactly $N$ steps. To prove this, we consider two generic states $G_i$ and $G_j$ and show that there is a positive probability for reaching state $G_j$ from state $G_i$ in finite number of steps. Any state $G_j$
can be reached from state $G_i$ with the transition path $G_0 \rightarrow G_1 \rightarrow \ldots \rightarrow G_N$, where $G_0 = G_i$ and $G_N = G_j$, where weights are changed link by link. Since number of distinct link in network is finite so it is the number of transitions $N$. As the transitions are independent the probability $P(G_i \rightarrow G_j)$ is equal to the product $\prod_{k=0}^{N-1} P(G_k \rightarrow G_{k+1})$, which is positive since every $P(G_k \rightarrow G_{k+1})$ is positive. Every product is equal to the weights in $G$ come from the distributions $\psi$ and $\phi$. As the probability $P(G_i \rightarrow G_j) > 0$ that means that the Markov Chain is also aperiodic.

**APPENDIX C: INFERENCE VIA SAMPLING**

Now we state the following theorem from Markov Chain theory [49].

**Theorem** Let $\{G_1, \ldots, G_n\}$ be random variables distributed according to the Markov Chain that satisfies the detailed balance and ergodic property with initial condition $G_0$ and let $f(G)$ be any real valued scalar function, then the following holds:

- The probability distribution $G_n$ converges to the stationary distribution:
  $$\lim_{n \to \infty} P(G_n = G) = P(G)$$

- Time average converges to the average over stationary distribution almost surely
  $$\langle f(G) \rangle = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f(G_i) = \sum_G P(G) f(G),$$

where $f(G)$ function is used for estimation of different properties over graph $G$.

We list different estimation functions that can be used with time augmented bond percolation mapping:

- Expected total outbreak size from node $v_i$ can be estimated with the following function:
  $$f(G, v_i) = \sum_j \Theta(\infty - d(v_i, v_j)),$$

where $\Theta(x)$ is the Heaviside step function, which equals to 1 when $x > 0$ i.e. $d(v_i, v_j) < \infty$ and 0 otherwise.

- Expected evolution of number of infected nodes at time $t$ from node $v_i$ can be estimated with the following function:
  $$f(G, v_i, t) = \sum_j \Theta(t - d(v_i, v_j)),$$

where $\Theta(x)$ is the Heaviside step function, which equals to 1 when $x \geq 0$ i.e. $t \geq d(v_i, v_j)$ and 0 otherwise.

- The expected distance time for the spreading to propagate from node $v_i$ to node $v_j$ can be estimated with the following function:
  $$f(G, v_i, v_j) = d_G(v_i, v_j)$$

- The estimation of microscopic configuration e.g. probability of specific realization described by the set of infected nodes $R_\psi$ (observation) from specific source node $v_i$ can be estimated with kernel density estimator [28, 39]:
  $$f(G_j, R_\psi) = \exp\left(\frac{-(\varphi_j - 1)^2}{a^2}\right),$$

where $\varphi_j$ is the similarity function between specific realization $R_\psi$ and estimated realization from weighted graph $G$ from source node $v_i$ and $a$ is the kernel width parameter. The similarity $\varphi_j$ is calculated as
  $$\frac{1}{N} \sum_{k \in e_r} \Theta(t - d(v_i, v_k)) + \frac{1}{N} \sum_{k \in e_r^c} \Theta'(d(v_i, v_k) - t)$$

where $e_r$ denotes the complement i.e. set of non-infected nodes and $N$ is the normalization constant i.e. total number of nodes in a network. Function $\Theta(x)$ equals to 1 when $t \geq d(v_i, v_j)$ and 0 otherwise. Function $\Theta'(x)$ equals to 1 when $d(v_i, v_j) > t$ and 0 otherwise.

- Estimations based on averages of i.i.d. samples have convergence rate $O(n^{-1/2})$, where the Berry-Esseen theorem gives bounds on the constants for convergence. Main advantage of using the kernel estimators [39] is that they have a faster convergence rate. They have physical analogy to using the information potential field to non-parametrically estimate realization probability densities.

- Finally, we do not need to set the Kernel width parameter $a$ in advance. We can choose the parameter $a$ as the infinimum of the set of parameters for which the estimations have converged [28].

**APPENDIX D: SEMI-DIRECTED BOND PERCOLATION**

Previously, the SIR dynamics without memory has been mapped [19] to the semi-directed bond percolation networks. However, still this mapping was only describing the asymptotic dynamics ($t = \infty$). In this section, we formalize this mapping with generalization of transmissibility from the link level to the neighbourhood level in order to make connection with the semi-directed bond percolation. Furthermore, we show that in our mapping we only need to change small adjustments so that it becomes exact mapping with weights.
Newman transmissibility for the link

Let us look at the continuous time SIR model, where infected node transmits the disease to susceptible node at an average rate $\beta$ and infected nodes recover at the constant rate $\gamma$. So the probability of recovering in any short time interval $dt$ is $\gamma dt$ and the probability of transmitting the disease in any short time interval $dt$ is $\beta dt$.

The probability of recovering in any short time interval $dt$ is $\gamma dt$ and the probability that the node is still infected after a $\tau$ time is:

$$
\lim_{dt \to 0} (1 - \gamma dt)^{\pi} = \lim_{\pi \to \infty} (1 + \frac{-\gamma \tau}{\pi})^{\pi} = e^{-\gamma \tau},
$$

and the probability that the node remains infected this long and then recovers in the interval $\tau + dr$ is: $e^{-\gamma \tau} \gamma dr$, which is a standard formulation waiting times is probability density function $\phi(\tau) = e^{-\gamma \tau}$ for the exponential distribution with parameter $\gamma$. Now, we can use the same analogy and say that the waiting time for transmission when there is no recovery to stop it is the probability density function $\psi(t) = \beta e^{-\beta t}$.

Then, the total probability of transmission through edge is:

$$
T = \int_0^\infty e^{-\gamma \tau} dt \int_0^\tau e^{-\beta t} dt = 1 - \frac{\gamma}{\beta + \gamma}.
$$

Exact transmissibility for the neighbourhood

Consider the focal node with $n$ neighbours, the transmissibility for the first neighbourhood (with all correlations) or more precisely the probability $T_{n,k}$ that $k$ out of $n$ nodes get infected prior to the recovery time of central node is:

$$
T_{n,k} = \binom{n}{k} \frac{\gamma}{\beta} \frac{\Gamma(k + 1) \Gamma\left(\frac{\gamma + \beta (n-k)}{\gamma + \beta k}\right)}{\Gamma\left(k + \frac{\gamma + \beta (n-k)}{\beta}\right)}. \tag{25}
$$

Proof

If we know the recovery time $\tau$ of central node, the conditional probability of transmission through any edge is

$$
\int_0^\tau \psi(t) dt = 1 - e^{-\beta t}.
$$

Now, as we have $n$ nodes the probability that $k$ nodes gets infected if the recovery time of central node is $\tau$ is:

$$
P_{(n,k)}|\tau = \binom{n}{k} e^{-\beta t} (n-k) (1 - e^{-\beta t})^k, \tag{26}
$$

where $(1 - e^{-\beta t})^k$ describes that through $k$ transmission passes and through $n-k$ does not pass $e^{-\beta t (n-k)}$ and there are $\binom{n}{k}$ different combinations of $k$ nodes.

Then, we need to to get the joint distribution $P_{(n,k),\tau} = P_{(n,k)|\tau} P_\tau$:

$$
P_{(n,k),\tau} = \gamma e^{-\tau} \binom{n}{k} e^{-\beta t} (n-k) (1 - e^{-\beta t})^k. \tag{27}
$$

And finally, we integrate over recovery time $\int_0^{\infty} P_{(n,k),\tau} d\tau$ to loose dependence on $\tau$:

$$
P_{n,k} = \int_0^{\infty} \gamma e^{-\tau} \binom{n}{k} e^{-\beta t (n-k)} (1 - e^{-\beta t})^k d\tau. \tag{28}
$$

After integration we get the following expression:

$$
T_{n,k} = \binom{n}{k} \frac{\gamma}{\beta} \frac{\Gamma(k + 1) \Gamma\left(\frac{\gamma + \beta (n-k)}{\beta}\right)}{\Gamma(k + 1 + \frac{\gamma + \beta (n-k)}{\beta})}. \tag{29}
$$

The $T_{n,k}$ is a valid distribution: (1) $\sum_k T_{n,k} = 1$ and (2) $T_{n,k} \geq 0$ (all arguments to the gamma functions are real and non-negative).

Relation to Newman transmissibility and bond percolation

Now, we show that neighbourhood transmissibility $T_{n,k}$ is equal to Newman transmissibility [18] only for $T_{n=1,k=1}$. But, first let us calculate $T_{n,0}$:

$$
T_{n,0} = \frac{\gamma}{\beta} \frac{\Gamma\left(\frac{n+\beta}{\beta}\right)}{\Gamma\left(\frac{n+\beta}{\beta} + 1\right)} = \frac{\gamma}{\beta} \frac{\Gamma\left(\frac{n+\beta}{\beta} + n\right)}{\Gamma\left(\frac{n+\beta}{\beta} + n + 1\right)} \tag{30}
$$

now by the property of Gamma function $\Gamma(t+1) = t\Gamma(t)$ we obtain the following:

$$
T_{n,0} = \frac{\gamma}{\beta} \frac{\Gamma\left(\frac{n+\beta}{\beta} + n\right)}{\Gamma\left(\frac{n+\beta}{\beta} + n + 1\right)} = \frac{\gamma}{\gamma + n\beta}. \tag{31}
$$

Note, that thus $T_{n=1,k=1} + T_{n=1,k=0} = 1$ (transmission either happens or does not happens) and therefore $T_{n=1,k=1} = 1 - T_{n,0} = 1 - \frac{\gamma}{\gamma + n\beta} = T$.

For general distributions $\phi(\cdot)$ and $\psi(\cdot)$, the generalized transmissibility for the neighbourhood is:

$$
T_{n,k} = \int_0^\infty \phi(\tau) d\tau \binom{n}{k} \left(1 - \Psi(\tau)\right)^{(n-k)} \Psi(\tau)^k, \tag{32}
$$

where $\Psi(\tau) = \int_0^\tau \psi(t) dt$. For $n = k = 1$, we get the standard transmissibility

$$
T_{n=1,k=1} = \int_0^\infty \phi(\tau) d\tau \Psi(\tau) = \int_0^\infty \phi(\tau) d\tau \int_0^\tau \psi(t) dt = T. \tag{33}
$$

Exact mapping: semi-directed bond percolation with weights

Let us recall the mapping from the main paper. In particular, a *time-respecting weighted network* instance $G_k$ is created by taking the input network $G$ and assigning
weights to the edges with the Inverse Smirnov transform\cite{21}:
\[
\rho_{i,j} = \begin{cases} 
\Psi^{-1}(x) : & \Psi^{-1}(x) \leq \Phi^{-1}(y), \\
\infty : & \Psi^{-1}(x) > \Phi^{-1}(y),
\end{cases}
\]

where \(x\) and \(y\) are uniform random numbers \(\in [0,1]\), \(\Phi^{-1}(x)\) and \(\Psi^{-1}(y)\) are inverse functions of the cumulative inter-event distributions: \(\Phi(t) = \int_0^t dt' \phi(t')\) and \(\Psi(t) = \int_0^t dt' \psi(t')\). The quantities \(\Psi^{-1}(y)\) and \(\Phi^{-1}(x)\) respectively represent the samples of the transmission and recovery time obtained with the Inverse Smirnov transform of inter-event distributions.

Note, that in the above mapping the random variables \(x\) and \(y\) were generated per each edge. The exact mapping is obtained by generating a random variable \(y\) for each node that is then shared among each adjacent outgoing edges, where one assumes that each undirected link in the original network can be interpreted as a bi-directional one. Then one performs the aforementioned procedures for both directions. In this way, we can generate weights that are semi-directed and take into the account the conditional dependence between edges. The properties shown in Sec. 1-3 hold similarly for this case.

**APPENDIX E: AIRPORT NETWORK SPREADING**

In order to generate the effective distance matrix on the world airport transportation network, we use the world airport transportation system data\cite{50} \cite{51}. Model of spreading is a simple model of diffusion of the infected agents along the airport transportation network. Each node represents an airport and each edge represents a connection where the diffusion of infected and susceptible travellers happen. From the flux data between nodes, we estimate the corresponding transmission rate per flight \(\beta_{i,j}\) for every link in the network. The rate is now the estimated ratio of number of passengers traveling from node \(v_i\) to node \(v_j\) and the total outgoing traffic i.e. \(\beta_{i,j} = \frac{F_{ij}}{\sum_k F_{i,k}}\). Here, we model a discrete time SIR spreading process, where \(\beta_{i,j}\) is estimated from flux travel data. The recovery rate \(\gamma \approx 0\), which describes rapid pandemic spreading, where infected airport cannot fully get recovered in short time scale. Then the corresponding transmission inter-event density distribution is \(\psi(t) = \sum_{T=0}^{\infty} \delta(t - T)\beta(1 - \beta)^{T-1}\) i.e. geometric distribution - discrete analog of exponential distribution and \(\phi(t) = \delta(t - \infty)\), where \(\delta(t)\) denotes the Dirac delta distribution. Essentially, the weights on the edges are samples of geometric distribution, whose rate parameter is estimated from data.

**ACKNOWLEDGMENTS**

**Acknowledgments.** The authors would like to thank to prof. D. Helbing for constructive feedback and discussions. We also thank A. Lancic, T. Lipić, O. Woolley and K. Caleb for useful comments. The work of N.A.-F. has been funded by the EU Horizon 2020 SoBigData project under grant agreement No. 654024. and in part by the EU Horizon 2020 CIMPLEX project under grant agreement No. 641191. The work of D.T. has been funded by the Croatian Science Foundation IP-2013-11-9623 “Machine learning algorithms for insightful analysis of complex data structures”. K-K. K. acknowledges support by the ERC Grant Momentum (324247).

---

[1] R. M. Anderson and R. M. May, *Infectious diseases in humans* (Oxford University Press, 1992).
[2] Alessandro Vespignani, “Modelling dynamical processes in complex socio-technical systems,” *Nature Physics* 8, 32-39 (2011).
[3] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani, “Epidemic processes in complex networks,” *Rev. Mod. Phys.* 87, 925-979 (2015).
[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports* 424, 175 – 308 (2006).
[5] Daniel T. Gillespie, “A rigorous derivation of the chemical master equation,” *Physica A: Statistical Mechanics and its Applications* 188, 404–425 (1992).
[6] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2012).
[7] James P. Gleeson, “Binary-state dynamics on complex networks: Pair approximation and beyond,” *Phys. Rev. X* 3, 021004 (2013).
[8] C. Castellano and R. Pastor-Satorras, “Thresholds for epidemic spreading in networks.” *Phys. Rev. Lett.* 105, 218701 (2010).
[9] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Phys. Rev. Lett.* 86, 3200–3203 (2001).
[10] Kieran J. Sharkey, “Deterministic epidemic models on contact networks: Correlations and unbiological terms,” *Theoretical Population Biology* 79, 115–129 (2011).
[11] Brian Karrer and M. E. J. Newman, “Message passing approach for general epidemic models,” *Phys. Rev. E* 82, 016101 (2010).
[12] Andrey Y. Lokhov, Marc Mézard, Hiroki Ohta, and Lenka Zdeborová, “Inferring the origin of an epidemic with a dynamic message-passing algorithm,” *Phys. Rev. E* 90, 012801 (2014).
[13] Fabrizio Altarelli, Alfredo Braunstein, Luca Dall’Asta, Alejandro Lage-Castellanos, and Riccardo Zecchina, “Bayesian inference of epidemics on networks via belief propagation,” *Phys. Rev. Lett.* 112, 118701 (2014).
[14] D. Mollison, “Spatial contact models for ecological and epidemic spread,” *J R Stat Soc Series B* 39, 283–326.
We used 20 bin classes in each dimension.
[49] Marc Mezard and Andrea Montanari, *Information, Physics, and Computation* (Oxford University Press, Inc., New York, NY, USA, 2009).

[50] Official Airline Guide, http://www.oag.com.

[51] O. Woolley-Meza, C. Thiemann, D. Grady, J. J. Lee, H. Seebens, B. Blasius, and D. Brockmann, “Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo-ship movements,” *The Europ. Phys. Journal B.* The Eur. Phys. Journal B – Condensed Matter and Complex Systems **84**, 589–600 (2011).