

RESEARCH ARTICLE

Predicting epidemic evolution on contact networks from partial observations

Jacopo Bindi¹*, Alfredo Braunstein^{1,2,3}, Luca Dall'Asta^{1,2}

1 Department of Applied Science and Technology, Politecnico di Torino, Torino, Italy, **2** Collegio Carlo Alberto, Moncalieri, Italy, **3** Human Genetics Foundation, Torino, Italy

* These authors contributed equally to this work.

* jacopo.bindi@polito.it



Abstract

The massive employment of computational models in network epidemiology calls for the development of improved inference methods for epidemic forecast. For simple compartment models, such as the Susceptible-Infected-Recovered model, Belief Propagation was proved to be a reliable and efficient method to identify the origin of an observed epidemics. Here we show that the same method can be applied to predict the future evolution of an epidemic outbreak from partial observations at the early stage of the dynamics. The results obtained using Belief Propagation are compared with Monte Carlo direct sampling in the case of SIR model on random (regular and power-law) graphs for different observation methods and on an example of real-world contact network. Belief Propagation gives in general a better prediction that direct sampling, although the quality of the prediction depends on the quantity under study (e.g. marginals of individual states, epidemic size, extinction-time distribution) and on the actual number of observed nodes that are infected before the observation time.

OPEN ACCESS

Citation: Bindi J, Braunstein A, Dall'Asta L (2017) Predicting epidemic evolution on contact networks from partial observations. PLoS ONE 12(4): e0176376. <https://doi.org/10.1371/journal.pone.0176376>

Editor: Yamir Moreno, Universidad de Zaragoza, SPAIN

Received: August 21, 2016

Accepted: April 10, 2017

Published: April 26, 2017

Copyright: © 2017 Bindi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The dataset about sexual contacts studied in the paper was made publicly available by Authors of the Ref. 20-21 and can be found at <http://dx.doi.org/10.1371/journal.pcbi.1001109>.

Funding: The authors acknowledge funding by Italian FIRB Project RBFR10QUW4, ERC grant 267915, and project SIBYL from Fondazione CRT under the initiative “La Ricerca dei Talenti”. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Governments and health-care systems maintain costly surveillance programs to report and monitor over time new infection cases for a variety of diseases, from seasonal influenza to the most dreadful viruses such as Ebola. Although surveillance is at the core of modern epidemiology, the early detection of a disease does not automatically guarantee that the fate of the epidemics will be easy to predict, because of the intrinsic stochasticity of the transmission process and the incompleteness of the accessible information. The two issues are often intertwined because a mathematical description that provides a sufficiently accurate prediction at some spatial/temporal scale could become inadequate at another, due to the lack of sufficiently detailed information. For instance, individual-based stochastic compartment models, such as the Susceptible-Infected-Recovered model, are widely used to describe disease transmission in contact networks, but human interactions have only recently become the object of accurate data mining, and exclusively in small and controlled environments such as schools and hospitals (e.g. by means of the RFID technology) [1–4]. For large-scale epidemic forecast a detailed individual-based description is challenging, therefore researchers and practitioners have resort

Competing interests: The authors have declared that no competing interests exist.

to coarse-grained metapopulation representations integrated with large-scale datasets on human mobility and real-time estimated parameters [5–7].

Beside the difficulty of obtaining accurate data on human interactions, also the observation of the epidemic progression is usually partial, in particular during the initial stages of an outbreak. For this reason, the ability to use all available information to produce a reliable forecast from early and partial observations is crucial to minimize the impact of a disease, at the same time saving financial resources. Using a simple SIR model, Holme [8, 9] recently showed that even in the ideal situation in which all information about the structure of the interpersonal network is available, the intrinsic stochasticity of the epidemic process makes prediction of relevant quantities, such as the final outbreak size and the extinction time, very difficult. The quality of the prediction depends on the epidemic parameters (transmission and recovery rates) and on the structure of the underlying network. Predicting the evolution of the epidemics becomes then even more difficult when only partial observation of the state of the individuals is provided.

In a recent series of works [10–12], the problem of inferring the origin of an epidemics in individual-based models from partial observations was investigated. Among the different methods proposed, the Belief Propagation (BP) method [10] is not only very reliable and efficient in identifying the origin of an observed epidemics, it also makes possible to easily reconstruct the probability marginals of the individual states at any time, exploiting the causality relations that are generated during the epidemic propagation. Hence, the method can be used to complete the missing information at the time of observation and applied to the problem of epidemic forecasting. In the present paper, we will use BP to predict the evolution of a SIR model on a given network from a partial observation of the states of the nodes in the early stage of the dynamics.

The paper is organized as follows. In Section Inference Models we define direct sampling and Bayesian methods and we introduce the metrics used for validation of the results. Section Results contains a comparison between the prediction obtained using Belief Propagation and Monte Carlo sampling for simulated SIR epidemics on random (regular and power-law) graphs, as well as on a real network of sexual contacts. For these networks, the effectiveness of the methods to predict local (e.g. marginals of individual states) and global (e.g. epidemic size, extinction-time distribution) properties are discussed. Section Methods reports the description of the main techniques employed in this work, in particular the static factor-graph representation of the epidemic process and the Belief Propagation equations used to evaluate the relevant posterior probabilities of the epidemic process given the observations.

Inference models

Prediction from partial observations in the SIR model

We consider a discrete time susceptible-infected-recovered (SIR) epidemic model on a graph $G = (V, E)$ that represents a contact network of $N = |V|$ individuals. At each time step of the dynamics a node $i \in V$ can be in one of three possible states: susceptible (S), infected (I) and recovered (R). The state of a node i at time t is represented by a variable $x_i^t \in \{S, I, R\}$. The stochastic process is defined by a set of parameters $\{\lambda_{ij}, \lambda_{ji}\}_{(i, j) \in E}$ and $\{\mu_i\}_{i \in V}$, such that at each time step an infected node i can infect every susceptible node j in his neighborhood ∂i with a probability λ_{ij} , then recover with probability μ_i (see Sec. Methods for further details on the dynamics). For a given assignment of the infection parameters and a given initial condition $\mathbf{x}^0 = \{x_i^0\}_{i \in V}$, a huge number of different realizations of the stochastic process exists, although some of these outcomes are more likely to occur than others. Epidemic forecasting consists in providing predictions about how much likely some outcomes are in the form of probability

distributions, in particular the probability marginals for the states of individual nodes. In realistic situations, the epidemic forecast is performed at some time after the initial infection event, when a number of infected cases is discovered in the population. The information available is thus usually localized in time and involves only a fraction of the overall population: we assume that at time T_{obs} the state $x_i^{T_{obs}}$ is made available for a set of nodes $i \in V_{obs} \subset V$ and no information about the state is supplied for the nodes not in V_{obs} . In order to focus only on the effects of partial observation, we assume that the structure of the contact network is completely known, and it does not change over time. We remark that, in the case we knew how the network changes over time, we could easily generalize the prediction methods to time-varying networks; unfortunately, the prediction of future contacts in time-varying networks is usually by itself a non-trivial inference problem [13, 14].

Direct sampling

Since the SIR stochastic process is Markovian, when $V_{obs} = V$ (complete observation) the probability of the future states \mathbf{x}^t for $t > T_{obs}$ can be estimated performing a *direct sampling*, that is generating a large number M_s of virtual realizations of the Markov chain from the same initial conditions (a complete observation at T_{obs}) and directly estimating the probability of an event from its relative frequency of occurrence in the experiment. In particular, if we call $x_{i,\ell}^t$ the value of the variable i at time t in the ℓ -th realization of the stochastic process from the same initial conditions, the individual probability marginal $P(x_i^t = X | \mathbf{x}^{T_{obs}})$ can be estimated from the experimental average

$$\hat{P}(x_i^t = X | \mathbf{x}^{T_{obs}}) = \frac{1}{M_s} \sum_{\ell=1}^{M_s} \mathbb{I}[x_{i,\ell}^t = X] \tag{1}$$

that converges as the empirical averages of a Bernoulli variable, i.e it rapidly converges to the correct value with a standard deviation that decreases with the number of trials as $\propto 1/\sqrt{M_s}$.

When $\mathbf{x}^{T_{obs}}$ is only partially known ($V_{obs} \subset V$), the uncertainty about the future evolution of an epidemic state is much larger; for instance, Fig 1 shows five very different evolutions of the epidemic process after the same partial observation. We call \mathbf{y}^{obs} the partial configuration that we observe at T_{obs} , such that $x_i^{T_{obs}} = y_i^{obs}, \forall i \in V_{obs}$. In order to apply the direct sampling method to the case of partial information we first need a way to complete the missing information at T_{obs} . In this work we consider two simple ways to choose the states of unobserved nodes at T_{obs} :

- **random sampling:** given the incomplete observation of the system, the states of unobserved nodes at time T_{obs} are drawn randomly, independently and uniformly with the same probability $1/3$, then direct sampling is performed with such an initial condition.
- **density sampling:** given the incomplete observation of the system, the fraction of observed nodes in each state $X \in \{S, I, R\}$ at time T_{obs} is used as an empirical probability to assign, independently and uniformly at random, the state of the unobserved nodes. Direct sampling is then employed to predict future states. The method can be generalized to include dependence on node attributes, such as the degree, by assigning to the unobserved nodes a state with a probability computed from the knowledge of the states of observed nodes with the same attributes.

We remark that unlike the case of complete information, the estimators obtained from these methods through direct sampling have non-zero bias.

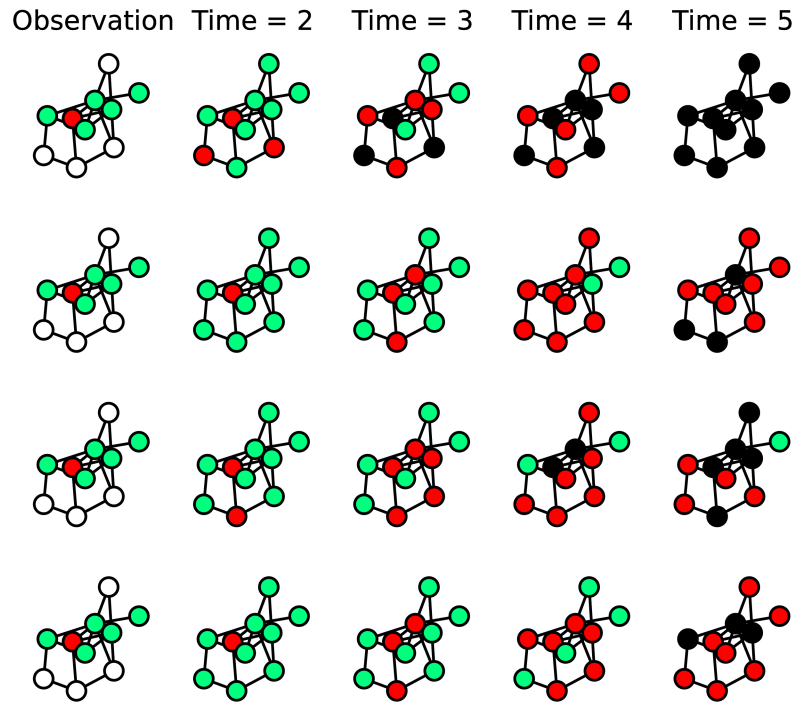


Fig 1. Each line represents a different realization for the SIR epidemic process given the (same) incomplete observation of the initial condition. Configurations in the leftmost column represent the observed state of the system, the other columns represent the time evolution of the epidemic process in that specific realization. Nodes colors: Green = Susceptible, Red = Infected, Black = Recovered, White = Unobserved.

<https://doi.org/10.1371/journal.pone.0176376.g001>

A bayesian approach

The *posterior* probability of a configuration \mathbf{x}^t at time t given an observation \mathbf{y}^{obs} at time T_{obs} can be written as

$$P(\mathbf{x}^t | \mathbf{y}^{obs}) = \frac{P(\mathbf{x}^t, \mathbf{y}^{obs})}{P(\mathbf{y}^{obs})} \propto P(\mathbf{x}^t, \mathbf{y}^{obs} | \mathbf{x}^0) P(\mathbf{x}^0), \quad (2)$$

where in the last expression we neglected the *a priori* probability $P(\mathbf{y}^{obs})$ of the observed state that only acts as a normalization constant in our analysis, while $P(\mathbf{x}^0)$ is the *prior* on the initial conditions. In this Bayesian approach, the prediction of the epidemic evolution after T_{obs} requires to compute the joint probability distribution $P(\mathbf{x}^t, \mathbf{y}^{obs} | \mathbf{x}^0)$ of the states at the observation time and at some later time given the initial condition. In principle, this quantity could be evaluated experimentally, by taking into account all possible realizations compatible with the constraints imposed by the dynamics and the observation and discarding the others. However, the number of possible epidemic trajectories of length t grows as 3^{tN} , making this brute-force approach computationally unfeasible even for small systems and very early observations. An approximate sampling method, that we call **Similarity Sampling**, is inspired by the Soft-Margin algorithm recently put forward in [15] to infer epidemic origins. The Similarity Sampling method consists in evaluating $P(\mathbf{x}^t, \mathbf{y}^{obs} | \mathbf{x}^0)$ by computing an empirical histogram over a large number of realizations of the epidemic process, each of them contributing with a probability weight that reflects the similarity to the actually observed states at T_{obs} . Every node in the

set of infected and removed nodes at the time of observation T_{obs} is used as single seed for a given number of realizations. We include unobserved nodes with at least one not susceptible neighbour. We assume to know the initial time of the epidemics within a ΔT_0 of time steps. Therefore we consider realizations with origin in a range $[-\Delta T_0, \Delta T_0]$. The similarity between a generic realization $\hat{\mathbf{x}}$ and the real one \mathbf{x} is measured by computing the Jaccard similarity function

$$\phi(\hat{\mathbf{x}}^{T_{obs}}, \mathbf{y}^{obs}) = \frac{|\mathcal{S}_{I+R}(\hat{\mathbf{x}}^{T_{obs}}) \cap \mathcal{S}_{I+R}(\mathbf{y}^{obs})|}{|\mathcal{S}_{I+R}(\hat{\mathbf{x}}^{T_{obs}}) \cup \mathcal{S}_{I+R}(\mathbf{y}^{obs})|}, \tag{3}$$

where $\mathcal{S}_{I+R}(\hat{\mathbf{x}}^{T_{obs}})$ is the set of infected and recovered individuals that are observed in the configuration $\hat{\mathbf{x}}^{T_{obs}}$. The weight function considered is a gaussian $\propto e^{-(\phi(\hat{\mathbf{x}}^{T_{obs}}, \mathbf{y}^{obs})-1)^2/a^2}$, where a is a free parameter. Then the individual marginal probability computed by Similarity Sampling reads:

$$\hat{P}(x_i^t = X | \mathbf{y}^{obs}) = \frac{1}{Z} \sum_{\ell=1}^{M_s} \mathbb{I}[\hat{x}_{i,\ell}^t = X] e^{-(\phi(\hat{\mathbf{x}}_{\ell}^{T_{obs}}, \mathbf{y}^{obs})-1)^2/a^2}, \tag{4}$$

where $X \in \{S, I, R\}$ and the factor of normalisation is $Z = \sum_{\ell=1}^{M_s} e^{-(\phi(\hat{\mathbf{x}}_{\ell}^{T_{obs}}, \mathbf{y}^{obs})-1)^2/a^2}$. Note that the same result can be achieved by normalizing $\hat{P}(x_i^t = X | \mathbf{y}^{obs})$ after having computed the sum. According to [15], for a fixed value of a , we consider a number M_s of realizations such that $\max(|P_{M_s}(x_i) - P_{\frac{M_s}{2}}(x_i)|) < 0.1$, i.e. the maximum of the differences between individual marginals after M_s and $M_s/2$ realizations is smaller than 0.1. For the problem of source inference, Antulov-Fantulin et al [15] applied the convergence criterion to the marginal probability to be infected at time $t = 0$. Instead, in the case of predicting the evolution of the disease spreading, we check the convergence of the marginal probability of being in any state at any time step. Given the larger number of conditions, reaching the convergence is computationally more demanding. In all results of the present paper we initially set $a = 0.125$. If the convergence criterion is not met for $M_s \leq 8 \times 10^5$ we use $a = 0.5$. The latter value guarantees the convergence for any instance. However, it is possible to consider the marginal probabilities computed using $a = 0.125$ even if the convergence criterion has not been met. In this case, results provided by Similarity Sampling may slightly vary. Although the method could provide a much more accurate estimate of the individual probability marginals than random and density sampling methods, such an accuracy is usually obtained through fine-tuning of the parameters and requires a very high computational power beyond the aim of this work.

Following the recent work by some of the authors [10, 11] we develop here a different approach that consists in addressing the joint probability distribution $P(\mathbf{x}^t, \mathbf{y}^{obs} | \mathbf{x}^0)$ as a probabilistic graphical model defined on a static representation of the dynamical trajectories. When the underlying contact network is a tree, the factor graph on which the graphical model is defined can be also reduced to a tree, and the joint probability distribution can be computed exactly by solving a set of local fixed-point equations known as **Belief Propagation** (BP) equations. On more general graphs, the BP equations can be considered as a heuristic algorithm that, under some decorrelation assumptions for the variables, provides a good approximation of the real probability distribution [16]. The BP equations for the quantity $P(\mathbf{x}^0 | \mathbf{y}^{obs})$, representing the posterior probability of the initial configuration given an observation at a later time, were derived in Refs. [10, 11]. The BP equations for the more complex graphical model in Eq (2) are discussed in detail in Sec. Methods. We stress that, with the BP approach, the prediction of the future evolution of the epidemics passes through the inference of the

(unobserved) dynamical states prior to the time of observation and the reconstruction of the causal relations developed in the dynamics.

Validation metrics

We used the different inference models under study to compute, for every node i , the marginal probabilities $P(x_i^t = X | \mathbf{y}^{obs})$ with $X \in \{S, I, R\}$. These quantities are then used in the binary classification problem of discriminating whether a node has been infected at a time $t' \leq t$ or not, that turns out to be a relevant measure to quantify the performances of the different prediction methods. In order to do that, we rank the nodes in decreasing order of magnitude of the probabilities $P(x_i^t = I | \mathbf{y}^{obs}) + P(x_i^t = R | \mathbf{y}^{obs})$ and build a Receiver Operating Characteristic (ROC) curve [10, 11]. Starting from the origin of the axes, the ROC curve is obtained from the ordered list of nodes by moving upward by one unit whenever a node is correctly classified as already infected at time t (true positive) or rightward in case it is not (false positive). The **area under the ROC curve** (AUC) expresses the probability that a randomly chosen node that was infected before time t is actually ranked higher in terms of the corresponding probability marginal than a randomly chosen susceptible one. When the ranking is equal to the real one, the area under the ROC is 1, whereas a completely random ordering gives an area equal to 0.5.

The area under the ROC curve gives indication of the fraction of the correctly classified nodes, but it does not depend much on the actual values of the marginal probabilities. The latter ones have instead a direct effect on a global quantity of crucial importance, the size of the epidemic outbreak, i.e. the number of nodes reached by the infection. The **average epidemic size** at time t can be expressed as function of the local marginals as [17]

$$size(t) = \frac{1}{N} \sum_i [P(x_i^t = I | \mathbf{y}^{obs}) + P(x_i^t = R | \mathbf{y}^{obs})]. \tag{5}$$

The **extinction time distribution** is another relevant global quantity, that cannot be directly computed from the knowledge of the individual probability marginals, and whose characterization on given network structures is a major issue in epidemic studies [18]. In particular, we are interested in the posterior probability distribution $P(T_{ext} | \mathbf{x}^{T_{obs}})$ that the discrete-time epidemic process dies out at time T_{ext} when it is conditioned on the (possibly partial) observation of the state $\mathbf{x}^{T_{obs}}$ at time T_{obs} .

A crucial point of the BP algorithm is that it is very convenient for computing local quantities, such as marginal probability laws for the single variables or pair-correlations. Some global quantities, such as the average epidemic size, can be directly computed from the knowledge of the local probability marginals. Interestingly, the quantity $P(T_{ext} | \mathbf{x}^{T_{obs}})$ can be expressed as the difference between two terms involving the free energies of the related graphical models when the epidemics are constrained to vanish before time T_{ext} and $T_{ext} - 1$, respectively. In Sec. Methods we show that such free energy can be efficiently computed, in the Bethe approximation, as the sum of local terms by means of the BP equations.

Results

Results for individual node classification and epidemic size

Random regular graphs. A first set of results for random regular graphs of size $N = 1000$ nodes and degree $k = 4$ is displayed in Fig 2 and corresponds to the observation of a fraction of 10% of the nodes chosen uniformly at random at $T_{obs} = 3$. Fig 2 displays (a) the average values of AUC and (b) the average epidemic size as function of the time steps $t > T_{obs}$ for different prediction methods: random sampling (green), density sampling (blue), Similarity Sampling

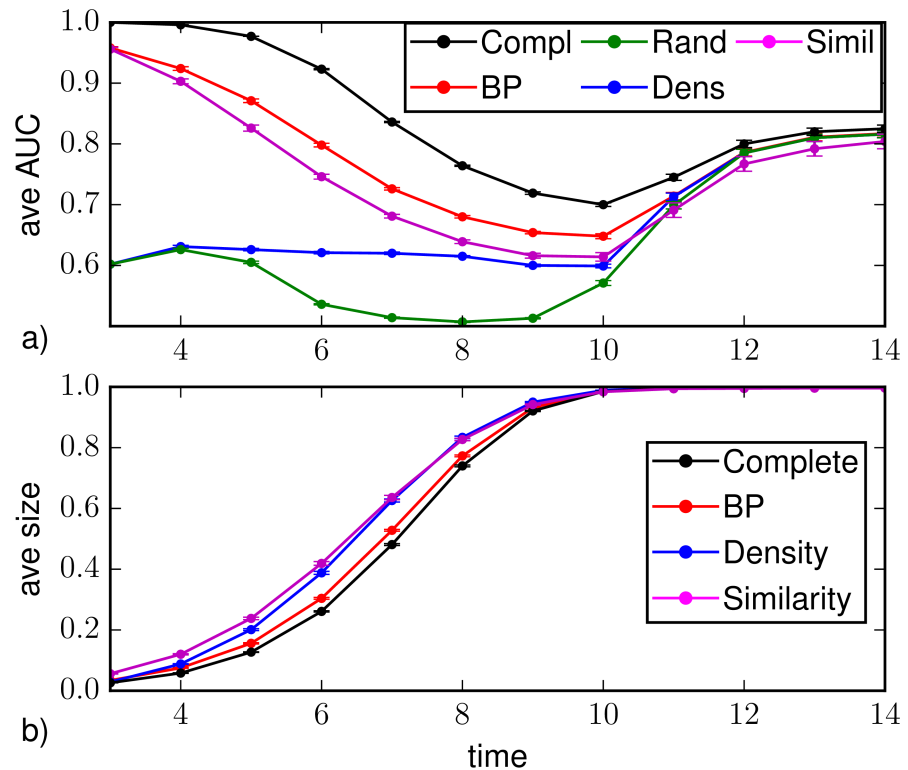


Fig 2. a) Area under the ROC curve as function of the time $t > T_{obs} = 3$ on a random regular graph of $N = 1000$ nodes and average degree $k = 4$. The average is computed over $M_o = 10^3$ epidemic realizations (with homogeneous parameters $\lambda = 0.7, \mu = 0.5$); the vertical bars represent the standard error of the mean. The prediction is obtained after the observation at T_{obs} of a 10%-fraction of the nodes chosen uniformly at random (*random observation*). b) Predicted average epidemic size on random regular graphs ($N = 1000, k = 4, \lambda_{ij} = 0.7, \mu_i = 0.5$) as function of time for a random observation of 10% of the nodes at $T_{obs} = 3$. The inference methods used are direct sampling with complete observation (black), random sampling (green), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red).

<https://doi.org/10.1371/journal.pone.0176376.g002>

(magenta) and Belief Propagation (red). As a reference we also plot results from direct sampling with complete observation (black). The average values are computed on $M_o = 10^3$ instances of observations at the same time T_{obs} , obtained from independent (and so possibly different) realizations of the epidemic propagations. For each observation, the direct sampling algorithms are performed on $M_s = 2.5 \cdot 10^5$ realizations of virtual epidemic processes. We set the error on the initial time $\Delta T_0 = 1$. The Similarity Sampling method seldom converges in a number of realizations $M_s = 8 \cdot 10^5$ when $a = 0.125$, therefore most of the results are obtained using $a = 0.5$.

Due to the intrinsic stochasticity of the SIR model, node classification is not perfect (AUC values smaller than 1) even in case of direct sampling with complete observation. In fact, the corresponding AUC values rapidly decrease after the observation time and recover only at late times since the epidemics dies out and almost all nodes are either in R or S states. If we interpret the average values of the area under the ROC as a proxy for the epidemic predictability, in the case of a complete observation (best-case scenario) the behavior observed is compatible with the effects due to epidemic heterogeneity reported in [5]. The most interesting region corresponds to intermediate times, when the predictability of the process is the lowest. Fig 2a shows that the Belief Propagation technique with partial observation gives values of averaged AUC that are closer to those from complete observation than the other methods. BP and

Similarity Sampling perform largely better in the first stage after the observation, corresponding to the exponential outbreak phase [19]. In particular Similarity Sampling gives an AUC value similar to BP at the time of observation, but a lower AUC value in the subsequent time steps.

Fig 2b shows that density sampling strongly overestimates the average epidemic size with respect to results from complete observation; this is probably an effect of the homogeneous deployment over the graph of infected nodes used to complete the information, that favors a larger epidemic spreading. Disregarding existing correlations between the 90% of the nodes, this scheme could lead to the overestimation of the probability of being infected—in a similar way to mean field approximations. In Similarity Sampling the overestimation of the epidemic size is due both to this procedure to set the seeds of the M_s virtual epidemic realizations and to the approximation on the initial time. Belief Propagation also slightly overestimates the epidemic size, but we think this is essentially due to the fact that in most of the instances the algorithm does not properly converge to the correct marginals.

The heatplots in Fig 3 display the same set of data classified as function of the number of observed nodes that were infected before the observation time, respectively for density sampling, Similarity Sampling and Belief Propagation. Results for direct sampling with complete observation are presented as a reference. In the case of the average AUC (Fig 3), BP performs

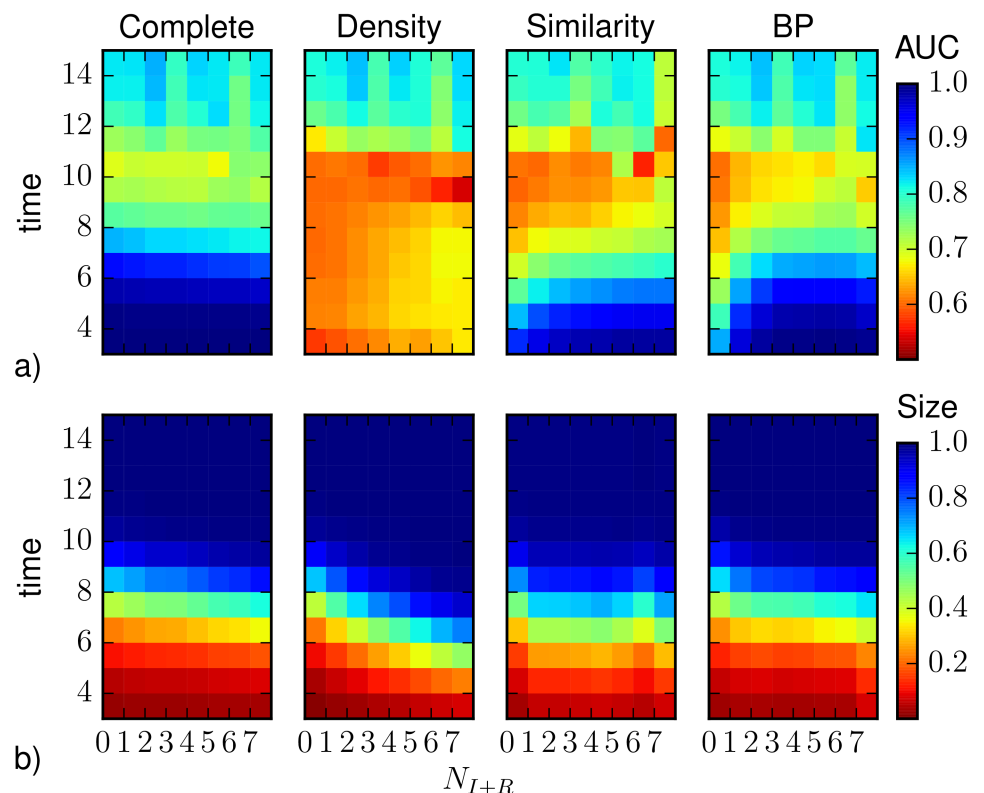


Fig 3. a) The heatplots represent the average AUC as function of time and of the number of observed nodes that were infected before T_{obs} , computed by density sampling, Similarity Sampling, Belief Propagation. b) The average epidemic size predicted by density sampling, Similarity Sampling and Belief Propagation is also shown as function of the number of infected and recovered nodes in the observation. As a reference, in both panels, we plot results obtained, for the same realizations of the SIR process, by direct sampling with complete observation. The horizontal axis refers to the number of infected or recovered nodes present in the 10% observation (also in the case of complete observation).

<https://doi.org/10.1371/journal.pone.0176376.g003>

better than both density sampling and Similarity Sampling in all regimes, in particular the performance is very good in the first steps after the observation, almost independently of the actual number of infected and recovered nodes in the observation. For all methods the results slightly improve when a larger number of nodes reached by the epidemics is observed at T_{obs} . For the average epidemic size, Fig 3b shows that the early-stage prediction by density sampling is negatively affected by the observation of a larger number of infected and recovered nodes. The opposite occurs, though to a lesser extent, for BP: when few infected nodes are observed BP overestimates the epidemic size, the worst prediction by BP giving an average size about 20% larger than that obtained from complete observation. The deviation observed by Similarity Sampling is also more evident when a lower number of infected and recovered nodes is observed, but the overestimation is more homogeneously distributed. Interestingly, the poor performance at large times is localized on realizations in which only a few of the observed nodes already got infected at T_{obs} .

In Fig 4 we show the results for the classification of individual states and the average epidemic size on random regular networks when different values of the epidemic parameters are considered. We include the results obtained considering the marginal probabilities computed by Similarity Sampling with $a = 0.125$ even if the convergence criterion is not met. In general, it provides similar results to the case in which the convergence is required. For $\lambda = 0.7, \mu = 0.5$

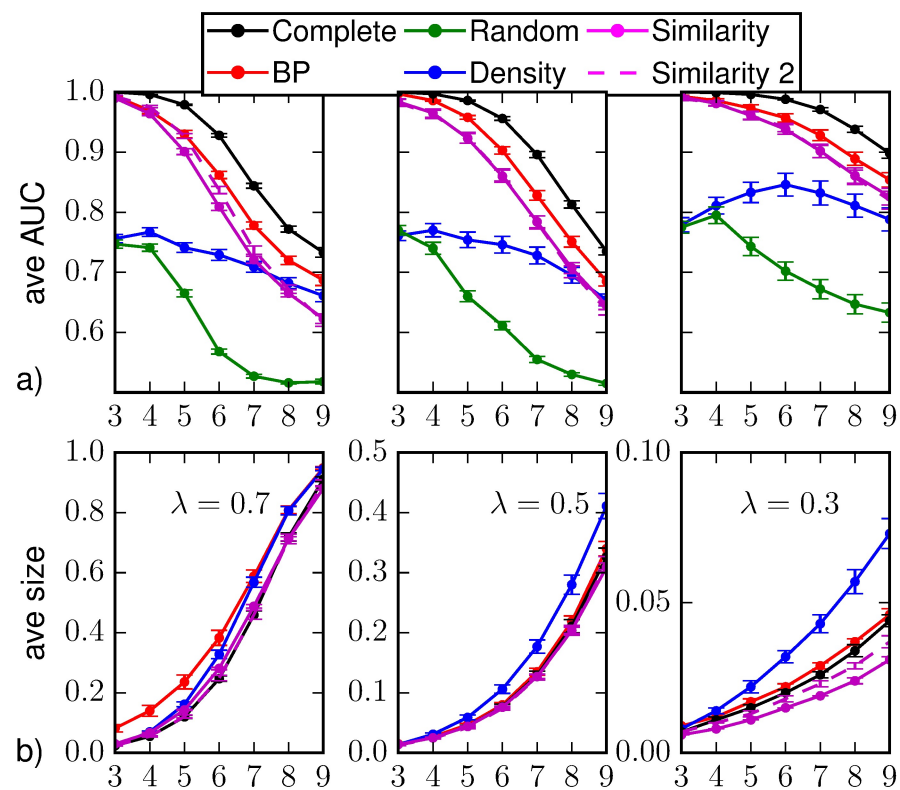


Fig 4. (a) Area under the ROC curve and (b) predicted average epidemic size as function of the time $t > T_{obs} = 3$ on a random regular graph of $N = 1000$ nodes and average degree $k = 4$. The average is computed over $M_o = 50$ epidemic realizations and the prediction is obtained after the observation at T_{obs} of a 30%-fraction of the nodes chosen uniformly at random (*random observation*). The inference methods used are direct sampling with complete observation (black), random sampling (green), density sampling (blue), Similarity Sampling (magenta), Similarity Sampling with $a = 0.125$ (dashed magenta) and Belief Propagation (red). Homogeneous parameters $\lambda = 0.7, \mu = 0.5$, $\lambda = 0.5, \mu = 0.5$ and $\lambda = 0.3, \mu = 0.5$.

<https://doi.org/10.1371/journal.pone.0176376.g004>

it provides slightly better results shortly after T_{obs} , but worse than those obtained by BP. We verified that these results are compatible with the ones obtained considering only the epidemic realizations in which the convergence for $a = 0.125$ is actually reached (additional information on the convergence of Similarity Sampling is provided in Sec. Methods). We argue that this result provides an upper bound to the performances of Similarity Sampling, because these realizations are somehow “simpler” to predict than others for this method.

Barabási-Albert random graph. In the case of heterogeneous graphs, such as those obtained with the Barabási-Albert (BA) growing network model, in addition to the *random observation*, it is interesting to define other observation schemes for the same density of observed nodes:

- *degree-based observation*: nodes are observed in descending order of their degree;
- *local observation*: a connected cluster of observed nodes is generated from a randomly chosen infected node.

We investigated the effect of different observation schemes on random sampling, density sampling, Similarity Sampling and BP.

The results for the average AUC, obtained with observation of 30% of the nodes at $T_{obs} = 3$, are reported in Figs 5 and 6. In the case of complete observation, direct sampling produces

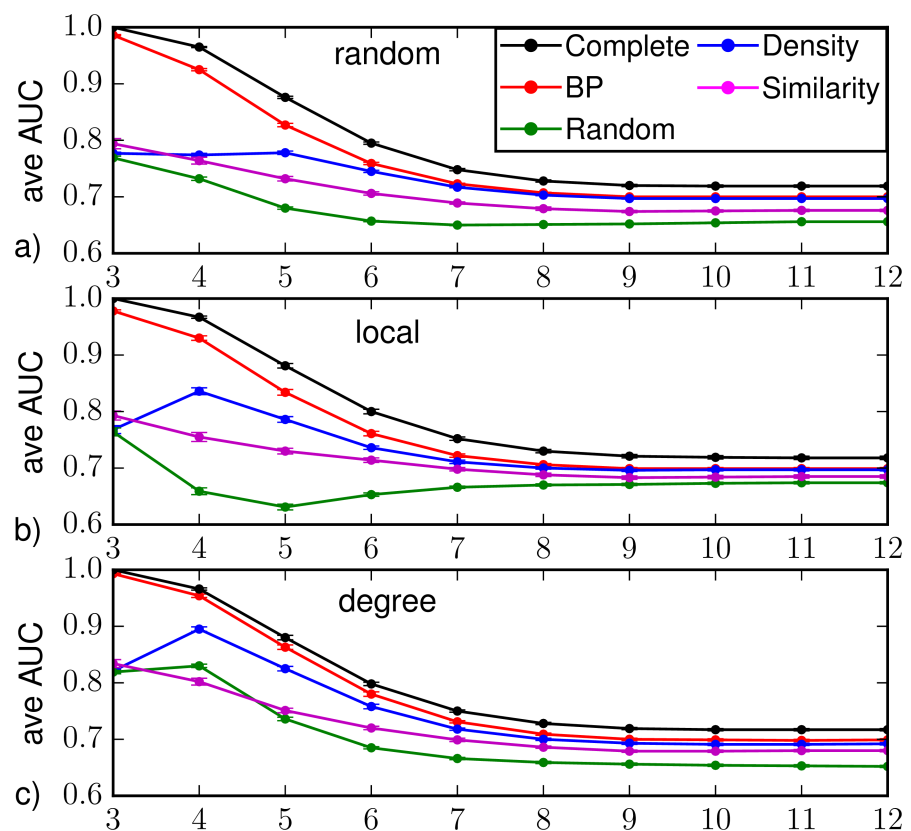


Fig 5. Area under the ROC curve as function of the time $t > T_{obs} = 3$ on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ (with homogeneous epidemic parameters $\lambda = 0.5$, $\mu = 0.6$), in the case of observation of a 30%-fraction of (a) nodes chosen at random uniformly and independently, (b) nodes forming a connected subgraph, (c) the most connected nodes. The average is computed over $M = 201$ epidemic realizations. The inference methods used are direct sampling with complete observation (black), random sampling (green), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red).

<https://doi.org/10.1371/journal.pone.0176376.g005>

monotonically decreasing AUC values for increasing times. The reason is that in finite size networks the parameters chosen give a non-zero probability of finding susceptible nodes in the last stage of the epidemic evolution, then wrong predictions are possible and the AUC remains considerably below one. For random observation, Fig 5a shows that Belief Propagation always gives larger AUC values than the other sampling methods, especially in the first stage of the epidemics, i.e. during the exponential outbreak. The same behavior is found plotting the results as function of the actual number of observed nodes (see heatplots in Fig 6a) that were already infected at the time of observation; in particular, the performances of BP are better when this number is small.

Figs 5b and 6b report results obtained with the observation of a 30%-fraction of nodes forming a connected subgraph. The overall results are very similar to those with random observation, even though density sampling and BP perform slightly better in the time steps immediately after the observation, while Similarity Sampling is slightly worse in the same regime. A degree-based observation is particularly convenient for heterogenous networks. Fig 5c shows that the average values of the ROC area increase in the first stage of the epidemics for all prediction methods, in particular the difference between values obtained by BP and those from direct sampling with complete observation is less than 2%. The results reported in the heatplots (see Fig 6c) are qualitatively similar to those from the other observation schemes, with slightly better prediction performances at early times when the number of infected nodes

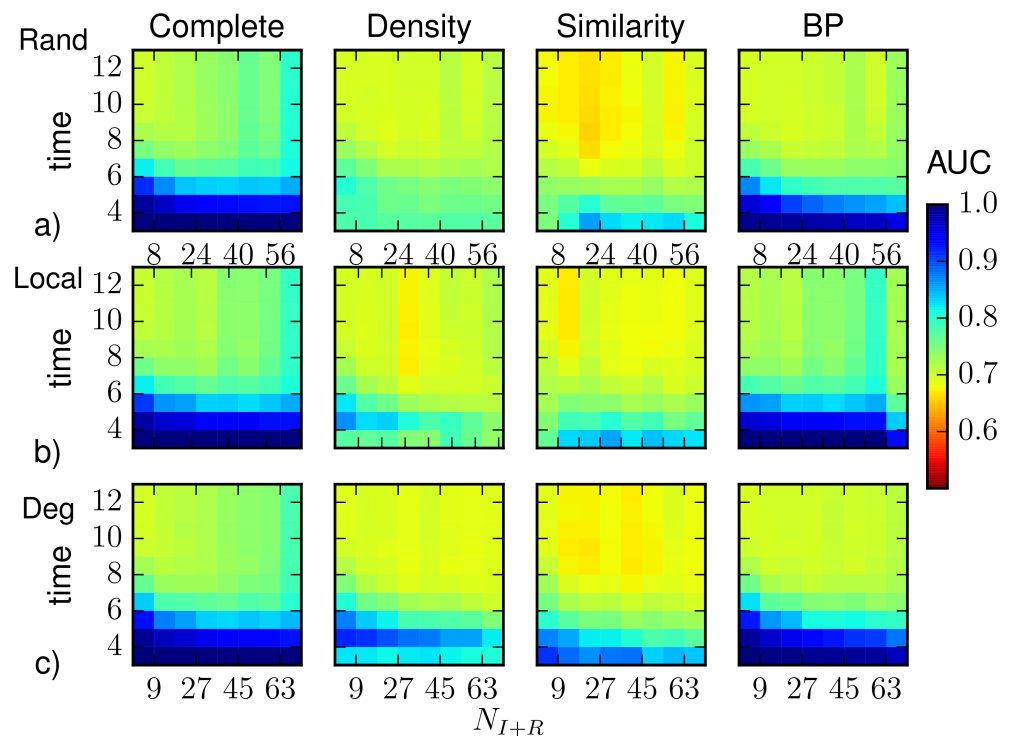


Fig 6. The heatplots represent the average AUC as function of time and of the number of observed nodes that were infected before $T_{obs} = 3$, computed by density sampling, Similarity Sampling, Belief Propagation, on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ with homogeneous parameters $\lambda = 0.5$, $\mu = 0.6$. As a reference, we also plot results obtained, for the same realizations of the SIR process, by direct sampling with complete observation. The prediction is obtained after the observation at T_{obs} of a 30%-fraction of (a) nodes chosen at random uniformly and independently, (b) nodes forming a connected subgraph, (c) the most connected nodes. The horizontal axis refers to the number of infected or recovered nodes present in the 30% observation (also in the case of complete observation).

<https://doi.org/10.1371/journal.pone.0176376.g006>

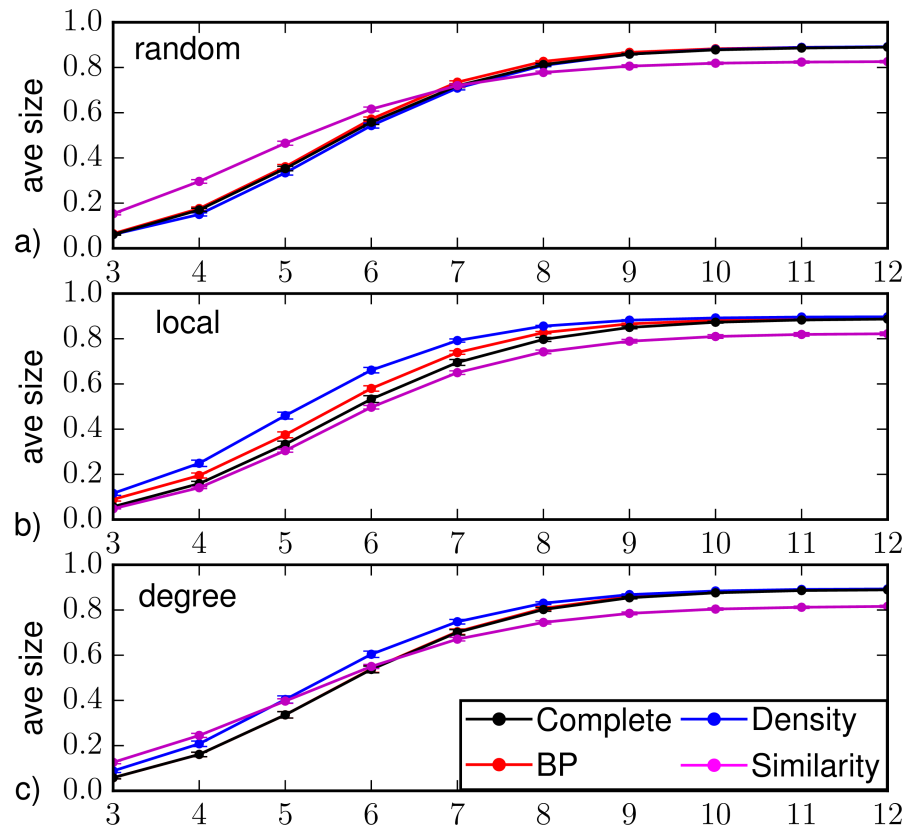


Fig 7. Predicted average epidemic size as function of the time $t > T_{obs} = 3$ on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ (with homogeneous epidemic parameters $\lambda = 0.5$, $\mu = 0.6$), in the case of observation of a 30%-fraction of (a) nodes chosen at random uniformly and independently, (b) nodes forming a connected subgraph, (c) the most connected nodes. The average is computed over $M = 201$ epidemic realizations. The inference methods used are direct sampling with complete observation (black), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red).

<https://doi.org/10.1371/journal.pone.0176376.g007>

in the observation is small. This is possibly due to the fact that these cases correspond to smaller epidemics whose initial evolution is more predictable.

Results on the prediction of the average epidemic size on Barabási-Albert networks are reported in Figs 7 and 8, except for random sampling that strongly overestimates the size in all regimes and observation schemes. With a random observation scheme (see Fig 7a), density sampling and BP provide very accurate prediction along the whole dynamics, while Similarity Sampling provides strong overestimate of the size value at early time and underestimate at late times. Fig 8a suggests that for both density sampling and BP, the accuracy is lower when a small number of infected and recovered nodes is observed. When the number of nodes reached by the infection at T_{obs} is larger, BP performs better than density sampling (4.5% of the nodes larger than the direct sampling with complete observation). The very bad results of Similarity Sampling at late times are mostly due to a very strong underestimation of the average size when the observation contains only few infected/recovered nodes. On the contrary at early times overestimate appears when a large number of infected/recovered nodes is observed.

In Fig 7b we show the prediction of the average epidemic size when the partial observation is performed considering a connected subgraph of 30% of the nodes. In this case all methods

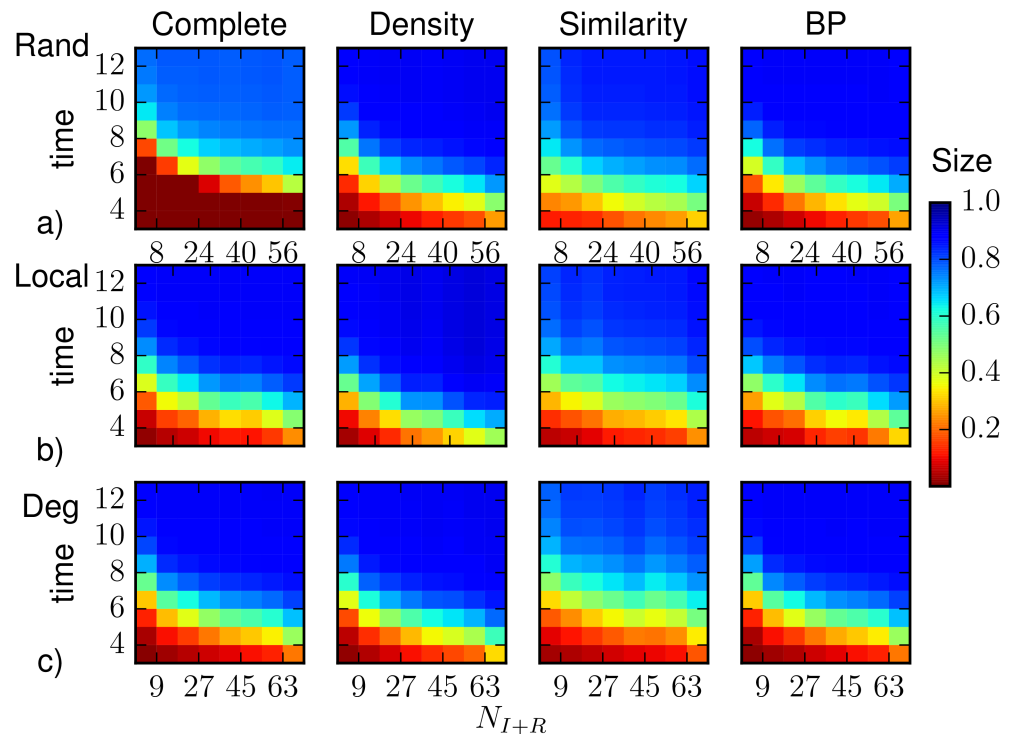


Fig 8. The heatplots represent the average epidemic size as function of time and of the number of observed nodes that were infected before $T_{obs} = 3$, computed by density sampling, Similarity Sampling, Belief Propagation, on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ with homogeneous parameters $\lambda = 0.5$, $\mu = 0.6$. As a reference, we also plot results obtained, for the same realizations of the SIR process, by direct sampling with complete observation. The prediction is obtained after the observation at T_{obs} of a 30%-fraction of (a) nodes chosen at random uniformly and independently, (b) nodes forming a connected subgraph, (c) the most connected nodes. The horizontal axis refers to the number of infected or recovered nodes present in the 30% observation (also in the case of complete observation).

<https://doi.org/10.1371/journal.pone.0176376.g008>

overestimate the epidemic size, with BP performing considerably better than the others. The poor performances of density sampling are expected because it completely neglects the topological information in the observation. For example if infected nodes are surrounded by susceptible ones, the probability of infection for unobserved nodes is lower, but this is not taken into account in the density sampling approach. BP performs instead poorly when there are very few infected nodes in the observed area. This is expected, because in such a situation this method is not able to correctly reconstruct missing information. Finally, Similarity Sampling gives good results for small and intermediate time steps but again it strongly deviates at large times, mostly because of observations with few infected/recovered nodes (see Fig 8b).

We already noticed that Belief Propagation performs very well in the case of a degree-based observation; this is true also for the epidemic size prediction, as shown in Fig 7c. The difference between the average epidemic size predicted by Belief Propagation and the one obtained by direct sampling with complete observation is less than 2% of the nodes in the network. Instead, density sampling overestimates the average epidemic size, especially in the first epidemic outbreak and for a large number of infected and recovered nodes in the observation (see Fig 8c). Density sampling does not make use of the connectivity knowledge, which is a valuable information: an observed highly connected node is more likely to be infected, ignoring this fact leads to assign the same infection probability of the hubs to every node in the

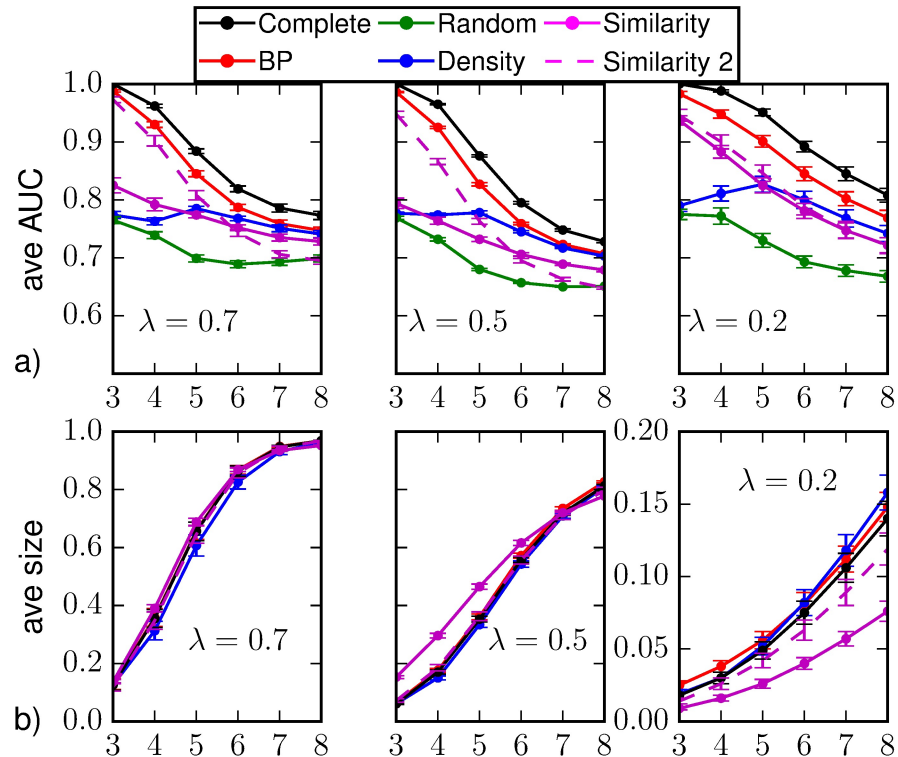


Fig 9. (a) Area under the ROC curve and (b) predicted average epidemic size as function of the time $t > T_{obs} = 3$ on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ in the case of observation of a 30%-fraction of nodes chosen at random uniformly and independently, with homogeneous epidemic parameters $\lambda = 0.7, \mu = 0.5, \lambda = 0.5, \mu = 0.5$, and $\lambda = 0.2, \mu = 0.5$. The average is computed over $M = 50$ epidemic realizations. The inference methods used are direct sampling with complete observation (black), random sampling (green), density sampling (blue), Similarity Sampling (magenta), Similarity Sampling with $a = 0.125$ (dashed magenta) and Belief Propagation (red).

<https://doi.org/10.1371/journal.pone.0176376.g009>

network, leading to larger predicted epidemic sizes. In this respect, one could expect that better results could be obtained simply by introducing a degree-dependence in the infection probability inferred from the observation; nevertheless, preliminary results show no significant improvement in the quality of the prediction.

In Fig 9 we show the results for the classification of individual states and the average epidemic size on BA networks when different values of the epidemic parameters are considered. We include the results obtained considering the marginal probabilities computed by Similarity Sampling with $a = 0.125$ even if the convergence criterion is not met. In this case it provides better results shortly after T_{obs} , but worse predictions at large time. In any case it is still less accurate than BP. We verified that these results are compatible with the ones obtained considering only the epidemic realizations in which the convergence for $a = 0.125$ is actually reached (additional information on the convergence of Similarity Sampling is provided in Sec. Methods). Following the same argument as in Fig 4, we argue that this result provides an upper bound to the performances of Similarity Sampling.

Results for the extinction time distribution

The extinction time distribution is a global feature of the epidemic process, that can strongly depend on the epidemic parameters, the initial conditions and the topological structure of the

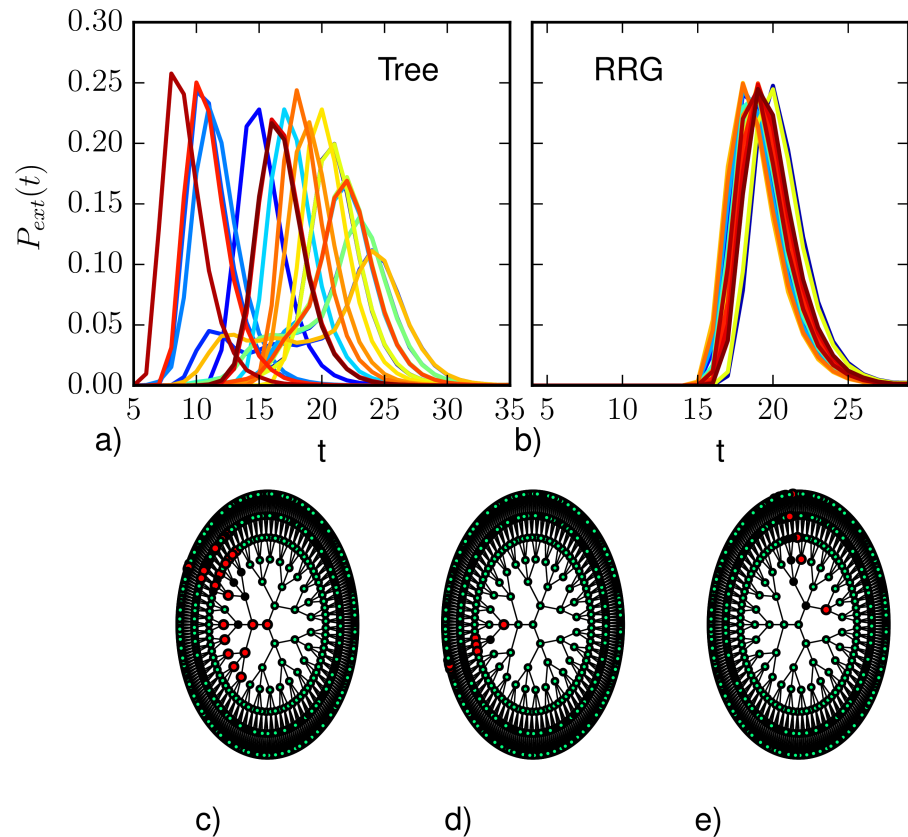


Fig 10. The extinction time distributions for different complete observations: a) on trees with branching ratio $k = 3$ and $N = 1092$ (epidemic parameters $\lambda = 0.7$, $\mu = 0.5$, and observation time $T_{obs} = 5$); b) on random regular graphs of degree 4 and $N = 1000$ nodes (epidemic parameters $\lambda = 0.7$, $\mu = 0.5$ and observation time $T_{obs} = 4$). Panels (c)-(e) illustrate similar realizations of the epidemic process at T_{obs} on a tree graph corresponding to rather different predicted extinction time distributions with maximum value respectively at $T = 21$ (c), $T = 23$ (d), and $T = 16$ (e). Nodes color: Green = Susceptible, Red = Infected, Black = Recovered.

<https://doi.org/10.1371/journal.pone.0176376.g010>

underlying contact network. Here we are interested in predicting the probability distribution for the extinction time when a (possibly partial) observation is provided. Even in the case of complete observation, the results are highly non-trivial, in particular on networks with peculiar topological structure. Fig 10 shows the extinction time distribution $P_{ext}(t) = P(t = T_{ext} | \mathbf{y}^{obs})$ for regular trees (a) and regular random graphs (b). In the case of trees the probability distribution is highly variable: depending on the observation, the width and the maximum value of the distribution can change significantly. Fig 10c–10e show three different realizations at the time of observation T_{obs} . In terms of the number of infected node and their average degree, the snapshots in panels (c) and (e) are similar, but their extinction time probability distributions are rather different ($T_{peak} = 16$ and $T_{peak} = 23$). On the contrary, despite the very different realizations at T_{obs} , snapshots in panels (c) and (d) correspond to similar distributions ($T_{peak} = 23$ and $T_{peak} = 21$). This is due to the arrangement of infected and recovered nodes at the time of observation: the configuration in Fig 10e does not allow to access the root of the tree, so the epidemics is limited and diffusion to other branches of the graph is blocked. In Fig 10c and 10d, instead, the epidemics can spread throughout the graph, causing the distribution to reach a maximum at larger times. The heterogeneity of the extinction time distribution is peculiar of trees and graphs with topological bottlenecks, while random graphs, or graphs with small-

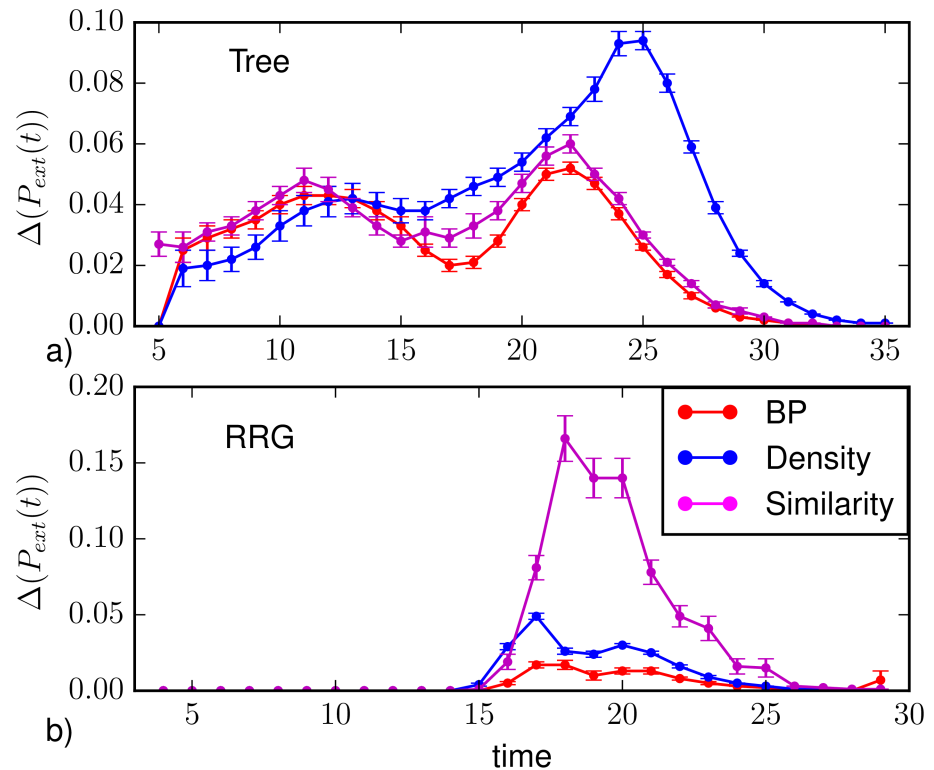


Fig 11. Absolute value of the difference between the extinction time distribution $P_{ext}(t)$ computed from direct sampling with complete information and those calculated with density sampling (blue), BP (red) and Similarity Sampling (magenta). a) On trees of $N = 1092$ nodes, with branching ratio 3 ($\langle k \rangle \approx 2$) and with uniform epidemic parameters $\lambda = 0.7$, $\mu = 0.5$. The partial observation is performed sampling uniformly the state of 10% of the nodes at $T_{obs} = 5$ and averaging over $M_o = 210$ such realizations. b) On random regular graphs of $N = 1000$ nodes and degree $k = 4$ with uniform epidemic parameters $\lambda = 0.7$, $\mu = 0.5$. The partial observation is performed sampling uniformly the state of 30% of the nodes at $T_{obs} = 4$ and averaging over $M_o = 150$ such realizations.

<https://doi.org/10.1371/journal.pone.0176376.g011>

world properties in general, are characterized by very similar distributions for different realizations of the epidemic process (with same epidemic parameters and observation).

The results on the prediction of the extinction time distribution from partial observations is shown in Fig 11. Motivated by the observed strong variability of the extinction time distribution, we first considered the case of regular trees of branching ratio equal to 4 (average degree $\langle k \rangle \approx 2$). The partial observation was obtained sampling randomly the state of 10% of the nodes at $T_{obs} = 5$. Fig 11a displays the average difference between the extinction time distribution predicted using direct sampling with complete observation and that obtained using Belief Propagation (red), density sampling (blue), and Similarity Sampling (magenta). All methods present two regions of higher discrepancy with respect to the prediction with complete observation. As shown by the heatplots in Fig 12b, this is usually due to an underestimation of the probability of extinction in the early stage of propagation and to an overestimation of the probability of extinction at large times. BP is usually able to qualitatively identify the most probable extinction time even when the other methods instead assign more probability mass to much larger times. Heatplots show that the two-peak discrepancy is especially due to observations with few infected and recovered nodes, while the discrepancies between the distributions move mostly at intermediate times when this number is increased. BP performs better than the other methods at every time step, although it presents the same qualitative weaknesses.

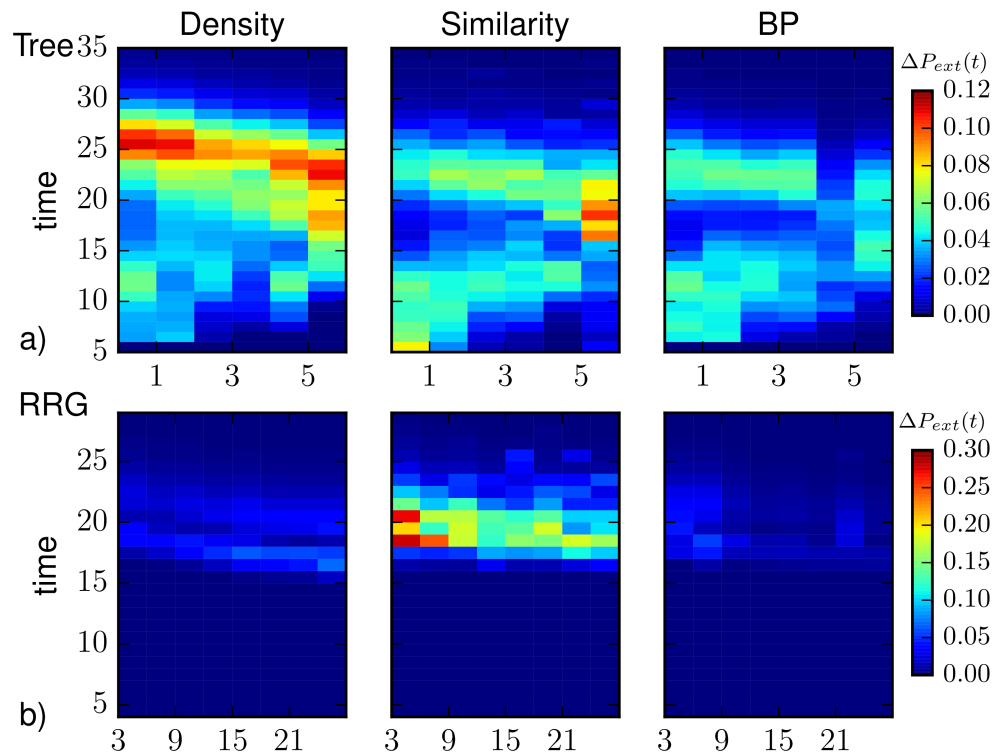


Fig 12. Absolute value of the difference between the extinction time probability distribution $P_{ext}(t)$ computed from direct sampling with complete information and those calculated with density sampling, BP and Similarity Sampling as a function of the number of infected and recovered nodes in the observed subset of nodes. a) On trees of $N = 1092$ nodes, with branching ratio 3 ($\langle k \rangle \approx 2$) and with uniform epidemic parameters $\lambda = 0.7$, $\mu = 0.5$. The partial observation is performed sampling uniformly the state of 10% of the nodes at $T_{obs} = 5$ and averaging over $M_o = 210$ such realizations. b) On random regular graphs of $N = 1000$ nodes and degree $k = 4$ and with uniform epidemic parameters $\lambda = 0.7$, $\mu = 0.5$. The partial observation is performed sampling uniformly the state of 30% of the nodes at $T_{obs} = 4$ and averaging over $M_o = 150$ such realizations.

<https://doi.org/10.1371/journal.pone.0176376.g012>

Interestingly, the Similarity Sampling method overestimates the probability for the epidemics to die out at early time step. In fact, a fraction of the epidemics with a large Similarity index to the observed incomplete snapshot immediately dies out after T_{obs} , leading to an overestimation of the extinction probability at early time steps.

Figs 11b and 12b display the same analysis in the case of random regular graphs of degree $k = 4$ with partial observation of the 30% of the nodes at $T_{obs} = 4$. Although all prediction methods under study are able to reproduce the existence of a unique peak, there are remarkable quantitative differences with the results from direct sampling with complete observation. The BP algorithm provides the best performances, in particular for observations with a large number of infected and recovered nodes. For a low number of infected and recovered nodes, instead, BP gives a larger average difference with respect to density sampling. This effect is mostly due to the non-convergence of the BP algorithm in some instances of the epidemic process, leading an overestimation of the probability of long extinction times. At the time steps close to the peak the Similarity Sampling give a larger average difference with respect to density sampling and BP. The Similarity Sampling gives the largest average difference. The main contribution to the average difference comes when low number of infected and recovered nodes are observed. In this case information provided by the observation is insufficient for Similarity Sampling.

A case study of real contact network

We consider a real network dataset of the sexual encounters of internet-mediated prostitution [20, 21], that was obtained analyzing a Brazilian web community exchanging information between male sex buyers. The original dataset is in the form of a bipartite temporal network, in which an edge between a “sex buyer” A and “sex seller” B is drawn if A posted a comment in a thread about B. The dataset covers the period September 2002 to October 2008 (2,232 days) and 50,185 contacts are recorded between 6,642 sex sellers and 10,106 sex buyers. In our analysis, we do not consider separate classes of vertices and we focus on a sample network comprising a time window between day 1000 and day 1100. The resulting network (SC) has $N = 1293$ nodes, $E = 1571$ edges, average degree $\langle k \rangle \approx 2.4$ and maximum degree $k_{max} = 55$.

We study the predictability of the epidemic evolution on a static projection of the sexual contact network when the observation takes place at times $T_{obs} = 4, 8$ as representatives of early and later time observation. In both cases, density sampling and random sampling make unreliable predictions of the classification of individual states of the nodes (see Fig 13a–13c). For $T_{obs} = 4$, BP gives good results only in the time steps immediately after the observation, then the performances rapidly deteriorate. BP results slightly improve increasing the observation

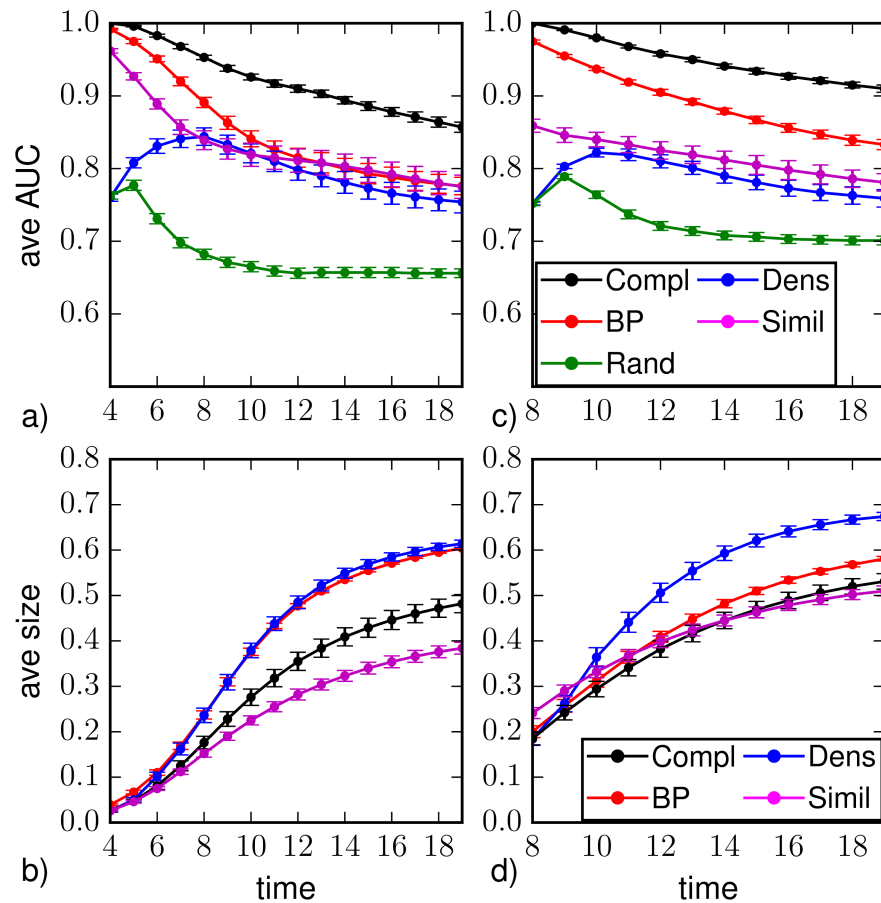


Fig 13. Average area under the ROC curve (a,c) and average epidemic size (b,d) as function to the time $t \geq T_{obs}$ for SIR dynamics ($\lambda = 0.5, \mu = 0.4$) on the SC network. Results are obtained with random sampling (green), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red) from a random observation of 30% of the nodes at $T_{obs} = 4$ (a,b) and $T_{obs} = 8$ (c,d). In all plots direct sampling from a complete observation is shown for comparison (black).

<https://doi.org/10.1371/journal.pone.0176376.g013>

time. Nevertheless BP is better than other methods. For the average epidemic size, Fig 13b shows that Similarity Sampling gives the best prediction at $T_{obs} = 4$ (though underestimating the epidemic size), whereas BP performs as bad as density sampling (and random sampling even worse). BP results improve considerably for $T_{obs} = 8$ while Similarity Sampling turns out to overestimate the epidemic size at early times (Fig 13d).

We remark that results are strongly influenced by the number of infected and recovered nodes in the observation. In this respect, in Fig 14, we repeat all measurements considering observations at $T_{obs} = 4$ containing a number of infected and recovered nodes equal to $N_{I+R} \geq 6$ (corresponding to the 46% of all instances), and at $T_{obs} = 8$ with $N_{I+R} \geq 18$ (75% of all instances). BP performances improve considerably at $T_{obs} = 4$, outperforming all other methods in the case of $T_{obs} = 8$. These results can be better understood if we consider that the network is characterized by a well connected core surrounded by many low degree nodes. When few infected nodes are observed, they typically are low-degree ones and the epidemic process spreads slowly at early time. In this situation Similarity Sampling is facilitated because the trajectories leading to the observed states are a small set. On the contrary, it is less accurate when

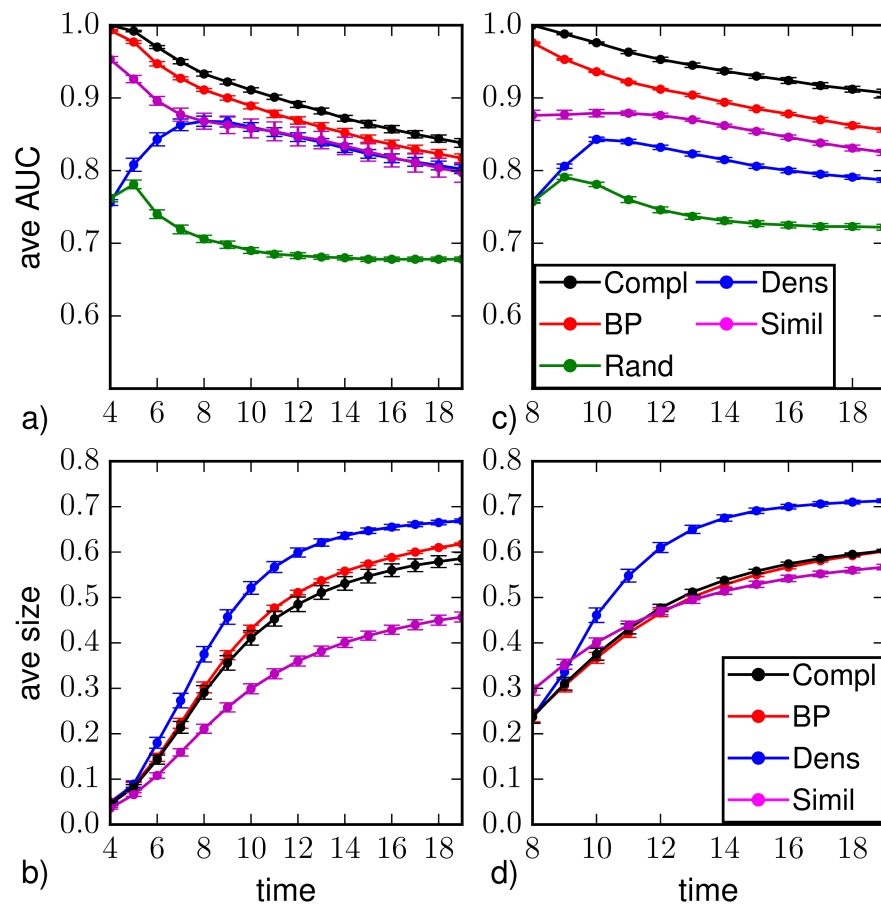


Fig 14. Average area under the ROC curve (a,c) and average epidemic size (b,d) as function to the time $t \geq T_{obs}$ for SIR dynamics ($\lambda = 0.5, \mu = 0.4$) on the SC network. Results are obtained with random sampling (green), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red) from a random observation of 30% of the nodes at $T_{obs} = 4$ (a,b) and $T_{obs} = 8$ (c,d). For $T_{obs} = 8$, only instances with a number of observed infected and recovered nodes $N_{I+R} > 18$ is considered (75% of instances). For $T_{obs} = 4$, only instances with observed infected and recovered nodes $N_{I+R} > 6$ is considered (46% of instances). In all plots direct sampling from a complete observation is shown for comparison (black).

<https://doi.org/10.1371/journal.pone.0176376.g014>

many infected nodes are observed or the observation occurs at later times. However results can likely be improved by an higher computational power. Belief Propagation does not choose the seed among the observed infected and recovered nodes only, it computes the probability of being seed for each node of the network, hence it is more accurate than Similarity Sampling when many infected and recovered nodes are unobserved. When the number of nodes reached by the epidemic spreading at the observation time is small, the effect of the existence of short loops in the network is more important and BP is more likely to overestimate the probability of a node to be infected [22]. It is worth noting that we provide to Similarity Sampling the information about the initial time $t = 0 \pm \Delta T_0$ of the epidemic spreading, on the contrary we don't provide such an information to BP.

We also consider a weighted static projections of the sexual contact network (WSC), in which every existing edge ij is assigned a weight w_{ij} corresponding to the number of contacts between node i and node j during the period under consideration. Then we define the probability that node i infects node j as $\lambda_{ij} = 1 - (1 - \lambda)^{w_{ij}}$. Fig 15 shows results for the average AUC and the average epidemic size. Belief Propagation provides higher values for the AUC than all the other methods at all times, even though AUC decreases with time much faster compared to direct sampling with complete observation. Immediately after the observation BP also provided the best prediction of the average epidemic size, while at late times Similarity Sampling works better.

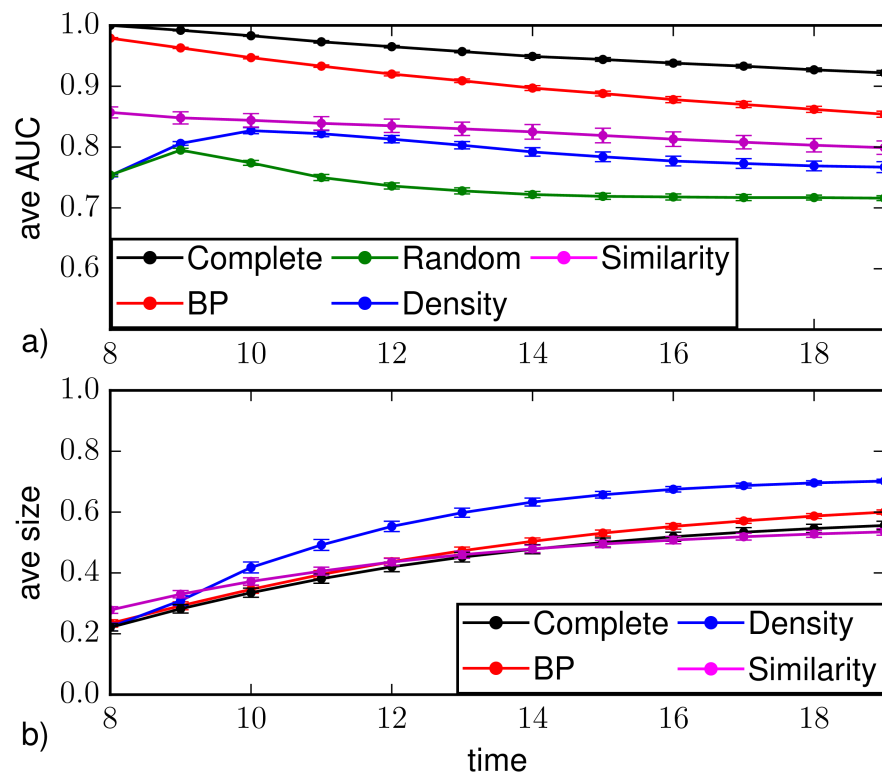


Fig 15. Average area under the ROC curve (a) and average epidemic size (b) as function to the time $t \geq T_{obs}$ for SIR dynamics ($\lambda = 0.5, \mu = 0.4$) on the WSC network. Results are obtained with random sampling (green), density sampling (blue), Similarity Sampling (magenta) and Belief Propagation (red) from a random observation of 30% of the nodes at $T_{obs} = 8$. Direct sampling from a complete observation is shown for comparison (black).

<https://doi.org/10.1371/journal.pone.0176376.g015>

Conclusions

In the present work, we extended the Bayesian Belief Propagation approach to the prediction of the future evolution of an epidemics, providing an efficient distributed algorithm to compute, at any time, the marginal probability of the states of individual nodes in the network. Some global quantities, such as the average epidemic size, can be directly computed as function of the individual marginal probabilities. Here we show that also quantities such as the extinction time distribution, that is intrinsically non local, can be reduced, in the BP approach, to a distributed calculation of local marginals on a locally tree-like factor graph. On random regular graphs and Barabási-Albert networks, the predictions obtained with the BP algorithm are compared with those from other heuristics based on Monte Carlo direct sampling from the same partial observations, while direct sampling with complete information is taken as a reference to assess the quality of the results. We also analyzed a real-world contact network obtained from a Brazilian database of sexual encounters. On random networks, BP provides better prediction than the other methods under study at all time steps. For all methods, the accuracy of the prediction is lower when the actual number of infected and recovered nodes in the observation is small. The errors introduced in the analysis of these configurations can result in a significant distortion of the overall results, in particular in the long time regime, as observed in the case of the average epidemic size measured by Similarity Sampling.

In general, BP is more accurate in the classification of individual marginals than in the estimate of the average epidemic size. A possible reason is that, in some cases, BP equations do not properly converge, resulting in a set of local probability marginals that are slightly different to the correct ones. In particular, convergence issues are mostly due to the presence of small loops in the network, that typically lead to an overestimate of the value of probability marginals by BP. As a consequence, the inaccuracies have little effect on the ranking of individual marginals, on which the ROC classification is based, whereas they are amplified when considering a global quantity such as the average epidemic size. Finally, BP usually approximate the extinction time probability distribution better than other methods.

In the real-world case study, results are affected by the presence of a much higher density of edges with respect to random graphs. BP gives better results for observations at late times. When the observation takes place early, the prediction of all methods is clearly worse, but BP gives the best prediction. The inaccurate prediction by BP is probably due to a combined effect of low level of information in the observation and the existence of many short loops in the network that limits the validity of BP. When considering only observations with a sufficiently large number of infected and recovered nodes, BP results improve considerably with respect to all other methods. An evidence of the role played by short loops comes from the better results generally obtained on the weighted network, because by construction the weighted network is effectively more sparse than the unweighted projection first used. In general, on networks that are not locally tree-like Belief Propagation may show convergence issues that become more significant as the density of the loops increases. In this cases, Belief Propagation likely would not be the recommended method. However, since locally tree networks represent a significant portion of real world and synthetic networks involved in epidemic spreading analysis, we believe that that this work still provides a significant improvement in this field.

In conclusion, BP and Similarity Sampling have advantages and drawbacks depending on the time and type of observation, but BP can be considered the most accurate method to predict both local and global quantities when the underlying network is sparse and when the observation contains a sufficiently large number of infected and recovered nodes. We remark that BP results have been obtained with no knowledge about the initial time of the epidemics, that is another important advantage compared with the other methods.

Methods

The SIR epidemic process

A node $i \in V$ can be in one of the possible states: susceptible (S), infected (I) and recovered (R). At each time step an infected node i can infect each of his neighbor j with a given probability λ_{ij} , then recover with probability μ_i . The state of a node i at time t is represented by a variable $x_i^t \in \{S, I, R\}$. The process is irreversible, so once a node recovered it does not get infected anymore. The Markov chain is described by the following transition probabilities

$$P(x_i^{t+1} = S | \mathbf{x}^t) = \mathbb{I}[x_i^t = S] \prod_{j \in \partial i} (1 - \lambda_{ij} \mathbb{I}[x_j^t = I]) \tag{6}$$

$$P(x_i^{t+1} = I | \mathbf{x}^t) = (1 - \mu_i) \mathbb{I}[x_i^t = I] + \mathbb{I}[x_i^t = S] (1 - \prod_{j \in \partial i} (1 - \lambda_{ij} \mathbb{I}[x_j^t = I])) \tag{7}$$

$$P(x_i^{t+1} = R | \mathbf{x}^t) = \mathbb{I}[x_i^t = R] + \mu_i \mathbb{I}[x_i^t = I]. \tag{8}$$

A realization of the SIR stochastic process is univocally expressed in terms of infection times t_i and recovery times $g_i, \forall i \in V$. Given the initial configuration \mathbf{x}^0 , for each node $i \in V$, a recovery time g_i is randomly drawn according to the distribution $\mathcal{G}_i(g_i) = \mu_i (1 - \mu_i)^{g_i}$ and the infection transmission delays s_{ij} from node i to node j are generated from the conditional distribution

$$\omega(s_{ij} | g_i) = \begin{cases} \lambda_{ij} (1 - \lambda_{ij})^{s_{ij}} & s_{ij} \leq g_i \\ \sum_{s > g_i} \lambda_{ij} (1 - \lambda_{ij})^s & s_{ij} = \infty. \end{cases} \tag{9}$$

Infection times are then given by the deterministic equation

$$t_i = \min_{j \in \partial i} (t_j + s_{ji}) + 1. \tag{10}$$

Factor graph representation

Every realization of the trajectory $(\mathbf{x}^0, \dots, \mathbf{x}^t)$ is in one-to-one correspondence with a static configuration of individual infection times $\mathbf{t} = \{t_i\}_{i \in V}$ and recovery times $\mathbf{g} = \{g_i\}_{i \in V}$. Using this static representation of the epidemic dynamics we can express the posterior probability as

$$P(\mathbf{x}^t | \mathbf{y}^{obs}) \propto \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^0, \mathbf{x}^{T_{obs}}} P(\mathbf{x}^t | \mathbf{t}, \mathbf{g}) P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g}) P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}}) P(\mathbf{t}, \mathbf{g} | \mathbf{x}^0) P(\mathbf{x}^0) \tag{11}$$

where $P(\mathbf{t}, \mathbf{g} | \mathbf{x}^0)$ is the joint probability distribution of infection, $P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}}) = \prod_{i \in V_{obs}} \delta(y_i, x_i^{T_{obs}})$ is the matrix that connects the complete configuration $\mathbf{x}^{T_{obs}}$ to the partially observed one \mathbf{y}^{obs} and recovery times conditioned on the initial configuration \mathbf{x}^0 , and $P(\mathbf{x}^t | \mathbf{t}, \mathbf{g})$ and $P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g})$ are deterministic functions of the set (\mathbf{t}, \mathbf{g}) representing the connection between a (\mathbf{t}, \mathbf{g}) configuration and configurations \mathbf{x}^t and $\mathbf{x}^{T_{obs}}$. Although the sum on the right-hand side of Eq (11) still runs over a possibly huge number (exponentially large in N) of configurations, a representation in which the dynamical relationships between trajectories of neighboring variables is reduced to a set of local constraints on the activation/recovery times is more convenient to develop approximation methods using tools from graphical models and statistical mechanics. By means of Bayes' theorem we compute the posterior probability of a configuration \mathbf{x}^t at time

t given an observation $\mathbf{x}^{T_{obs}}$ at time T_{obs} :

$$\begin{aligned}
 P(\mathbf{x}^t | \mathbf{y}^{obs}) &= \sum_{\mathbf{t}, \mathbf{g}} P(\mathbf{x}^t | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g} | \mathbf{y}^{obs}) \\
 &\propto \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^{T_{obs}}} P(\mathbf{x}^t | \mathbf{t}, \mathbf{g}) P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g}) P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}}) = \\
 &= \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^0, \mathbf{x}^{T_{obs}}} P(\mathbf{x}^t | \mathbf{t}, \mathbf{g}) P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g} | \mathbf{x}^0) P(\mathbf{x}^0) P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}})
 \end{aligned} \tag{12}$$

where $P(\mathbf{x}^0) = \prod_i \gamma_i(x_i^0)$ with

$$\gamma_0(x_i^0) = \gamma \mathbb{I}[x_i^0 = I] + (1 - \gamma) \mathbb{I}[x_i^0 = S] \tag{13}$$

is the factorized prior on the initial condition, and $P(\mathbf{x}^t | \mathbf{t}, \mathbf{g})$ and $P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g})$ are deterministic functions of the set (\mathbf{t}, \mathbf{g}) representing the probability of a configuration \mathbf{x}^t at time t (and, respectively, at time T_{obs}):

$$P(\mathbf{x}^t | \mathbf{t}, \mathbf{g}) = \prod_i \zeta_i^t(t_i, g_i, x_i^t) \tag{14}$$

with

$$\zeta_i^t(t_i, g_i, x_i^t) = \mathbb{I}[x_i^t = S, t < t_i] + \mathbb{I}[x_i^t = I, t_i \leq t \leq (t_i + g_i)] + \mathbb{I}[x_i^t = R, t > t_i + g_i]. \tag{15}$$

The joint probability distribution of infection and recovery times conditioned on the initial configuration reads

$$\begin{aligned}
 P(\mathbf{t}, \mathbf{g} | \mathbf{x}^0) &= \sum_{s_{ij}} P(\mathbf{s} | \mathbf{g}) P(\mathbf{t} | \mathbf{x}^0, \mathbf{g}, \mathbf{s}) P(\mathbf{g}) \\
 &= \sum_{s_{ij}} \prod_{i,j} \omega_{ij}(s_{ij} | g_i) \prod_i \phi_i(t_i, \{t_k, s_{ki}\}_{k \in \partial i}) \mathcal{G}_i(g_i)
 \end{aligned} \tag{16}$$

where

$$\phi_i(t_i, \{t_k, s_{ki}\}_{k \in \partial i}) = \delta(t_i, \mathcal{I}[x_i^0 \neq I] (\min_{k \in \partial i} (t_k + s_{ki}) + 1)). \tag{17}$$

The factor graph representation of a probability distribution is made up of a bipartite graph composed of factor nodes and variable nodes [16]. Each factorized term in Eq (12) is represented by a factor node and each variable of the problem is represented by a variable node. Each factor node is connected to the set of variable nodes involved in the corresponding factorized term. The factor graph of Eq (12) has a loopy structure which can compromise the accuracy of the BP approximation. We can use a factor graph representation that maintains the same topological properties of the original graph in order to guarantee that BP is exact when the underlying graph is a tree. Following [22, 23], we do that by grouping pairs of variable nodes (t_i, t_j) in the same variable node. For each edge (i, j) emerging from node i we introduce a triplet $(t_i^{(j)}, t_j, g_i^j)$, where t_i^j and g_i^j are copies (on j) of the infection time and recovery time of i and the variables $t_{ij} = t_i^{(j)} + s_{ij}$ on which factors ϕ_i depend. Including a constraint that forces copies t_i and g_i to have a common value, we get

$$\psi_i = \delta(t_i, \mathbb{I}[x_i^0 \neq I] (\min_{j \in \partial i} (t_{ij} + 1))) \prod_{j \in \partial i} \delta(t_i^{(j)}, t_i) \delta(g_i^{(j)}, g_i) \tag{18}$$

and

$$\phi_{ij} = \omega_{ij}(t_{ij} - t_i^{(j)} | g_i^{(j)}) \omega_{ji}(t_{ji} - t_j^{(i)} | g_j^{(i)}). \tag{19}$$

In this representation we can write the posterior probability as the following *graphical model*:

$$P(\mathbf{x}^t | \mathbf{y}^{obs}) = \sum_{\mathbf{t}, \mathbf{g}, \mathbf{s}, \mathbf{x}^{T_{obs}}} \prod_{i,j} \omega_{ij} \prod_i \phi_i \mathcal{G}_i \gamma_i \zeta_i^{T_{obs}} \zeta_i^t \prod_{i \in V_{obs}} \delta(y_i, x_i^{T_{obs}}). \tag{20}$$

Belief propagation equations

Given a set $\underline{x} = (x_1, \dots, x_N)$ of random variables with a joint probability distribution

$$M(\underline{z}) = \frac{1}{Z} \prod_a F_a(z_{\partial a}) \tag{21}$$

where $z_{\partial a} \equiv \{x_i | i \in \partial a\}$ is the set of variables involved in the constraint a . Messages are associated with every directed edge on the factor graph and they take values in the space of single-variable probability distributions. The following equations for messages are solved by iteration:

$$\begin{aligned} p_{F_a \rightarrow i}(z_i) &= \frac{1}{Z_{ai}} \sum_{\{z_j | j \in \partial a \setminus i\}} F_a(\{z_i\}_{i \in \partial a}) \prod_{j \in \partial a \setminus i} m_{j \rightarrow F_a}(z_j) \\ m_{i \rightarrow F_a}(z_i) &= \frac{1}{Z_{ia}} \prod_{b \in \partial i \setminus a} p_{F_b \rightarrow i}(z_i) \\ m_i(z_i) &= \frac{1}{Z_i} \prod_{b \in \partial i} p_{F_b \rightarrow i}(z_i). \end{aligned}$$

At the fixed point they provide an approximate value for the variables marginal probability [16]. In our case the factors F_a are $\psi_i, \phi_{ij}, \gamma_i, \zeta_i^{T_{obs}}, \zeta_i^t$ and \mathcal{G}_i and the variables z_i are couples (t_i, g_i) , and triplets $(t_i^{(j)}, g_i^{(j)}, t_{ij}), (t_i, g_i, x_i^{T_{obs}}), (t_i, g_i, x_i^t)$. The explicit form for the update equations of the ψ_i factor nodes is:

$$\begin{aligned} p_{\psi_i \rightarrow j}(t_i^{(j)}, t_{ji}, g_i^{(j)}) &\propto \sum_{g_i, t_i} \sum_{\{t_i^{(k)}, t_{ki}, g_i^{(k)}\}} m_{i \rightarrow \psi_i}(t_i, g_i) \times \\ &\times \prod_{k \in \partial i \setminus j} m_{k \rightarrow \psi_i}(t_i^{(k)}, t_{ki}, g_i^{(k)}) \psi_i(t_i, g_i, \{(t_i^{(k)}, t_{ki}, g_i^{(k)})\}_{k \in \partial i}) \end{aligned} \tag{22}$$

and

$$p_{\psi_i \rightarrow j}(t_i, g_i) \propto \sum_{\{t_i^{(k)}, t_{ki}, g_i^{(k)}\}} \prod_{k \in \partial i \setminus j} m_{k \rightarrow \psi_i}(t_i^{(k)}, t_{ki}, g_i^{(k)}) \psi_i(t_i, g_i, \{(t_i^{(k)}, t_{ki}, g_i^{(k)})\}_{k \in \partial i}). \tag{23}$$

Efficient forms for these update equations are given in [10, 11].

In the factor graph representation, the Bethe free-energy of the graphical model can be expressed as (see also [16])

$$-f = \sum_a f_a + \sum_i f_i - \sum_{(ia)} f_{(ia)} \tag{24}$$

in which the local contributions can be expressed as function of the Belief Propagation messages

$$f_a = \log \left(\sum_{\{z_i: i \in \partial a\}} F_a(\{z_i\}_{i \in \partial a}) \prod_{i \in \partial a} m_{i \rightarrow a}(z_i) \right) \tag{25}$$

$$f_i = \log \left(\sum_{z_i} \prod_{b \in \partial i} P_{F_b \rightarrow i}(z_i) \right) \tag{26}$$

$$f_{(ia)} = \log \left(\sum_{z_i} m_{i \rightarrow a}(z_i) P_{F_b \rightarrow i}(z_i) \right). \tag{27}$$

The extinction-time constraint

The posterior probability $P(T_{ext}|y^{obs})$ of the extinction time from a partial observation at T_{obs} can be written as a difference of posterior probabilities that an epidemic ends within a given

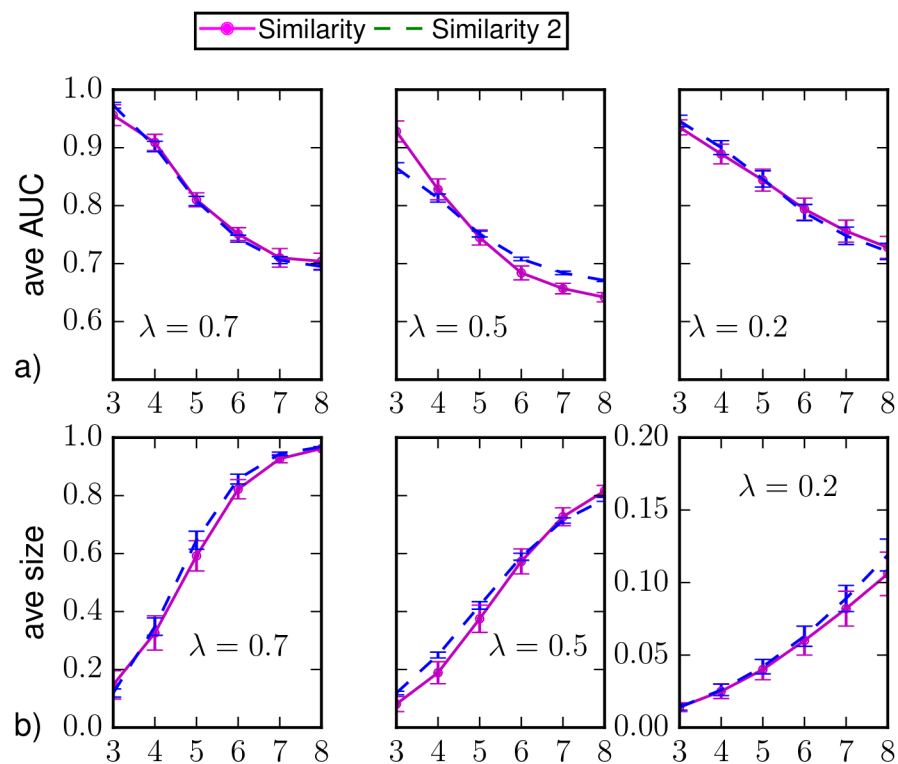


Fig 16. (a) Area under the ROC curve and (b) predicted average epidemic size as function of the time $t > T_{obs} = 3$ on a random regular graph of $N = 1000$ nodes and average degree $k = 4$. The prediction is obtained after the observation at T_{obs} of a 30%-fraction of the nodes chosen uniformly at random (*random observation*). Homogeneous parameters $\lambda = 0.7, \mu = 0.5, \lambda = 0.5, \mu = 0.5$ and $\lambda = 0.3, \mu = 0.5$. The solid magenta line represents the Similarity Sampling for epidemic realizations enabling the convergence for $a = 0.125$ (respectively $M_o = 21, M_o = 21, M_o = 45$); the dashed blue line represents the Similarity Sampling with $a = 0.125$ although convergence is not met (the average is computed over $M_o = 50$ epidemic realizations).

<https://doi.org/10.1371/journal.pone.0176376.g016>

time,

$$P(T_{ext} | \mathbf{y}^{obs}) = P(t_{ext} < T_{ext} | \mathbf{y}^{obs}) - P(t_{ext} < T_{ext} - 1 | \mathbf{y}^{obs}). \tag{28}$$

Using the static representation of dynamical trajectories,

$$\begin{aligned} P(t_{ext} < T_{ext} | \mathbf{y}^{obs}) &= \sum_{\mathbf{t}, \mathbf{g}} P(t_{ext} < T_{ext} | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g} | \mathbf{y}^{obs}) \\ &\propto \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^{T_{obs}}} P(t_{ext} < T_{ext} | \mathbf{t}, \mathbf{g}) P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g}) P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}}) \\ &= \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^0, \mathbf{x}^{T_{obs}}} P(t_{ext} < T_{ext} | \mathbf{t}, \mathbf{g}) P(\mathbf{x}^{T_{obs}} | \mathbf{t}, \mathbf{g}) P(\mathbf{t}, \mathbf{g} | \mathbf{x}^0) P(\mathbf{x}^0) P(\mathbf{y}^{obs} | \mathbf{x}^{T_{obs}}) \\ &= \sum_{\mathbf{t}, \mathbf{g}, \mathbf{x}^0} \mathcal{Q}(\mathbf{y}^{obs}, T_{ext}, \mathbf{t}, \mathbf{g}, \mathbf{x}^0) = Z(T_{ext}, \mathbf{y}^{obs}). \end{aligned} \tag{29}$$

The terms in the latter expression are the same as in Eq (12), with the exception of the following term factorized over the nodes

$$P(t_{ext} < T_{ext} | \mathbf{t}, \mathbf{g}) = \prod_i \mathbb{I}[(t_i + g_i) < T_{ext}] \tag{30}$$

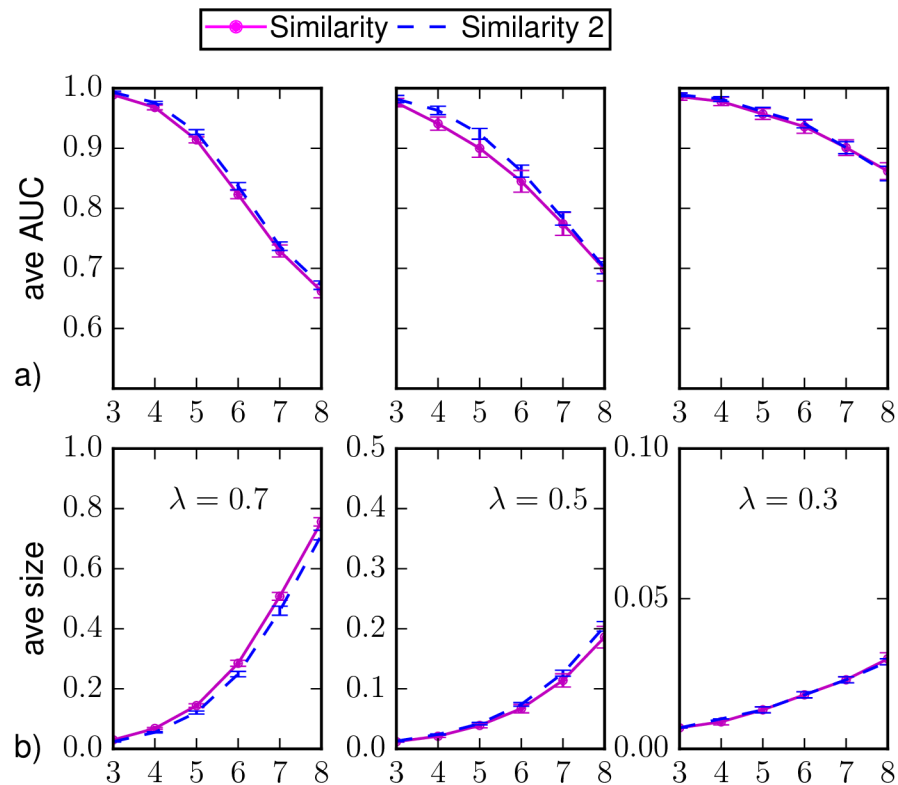


Fig 17. (a) Area under the ROC curve and (b) predicted average epidemic size as function of the time $t > T_{obs} = 3$ on a Barabási-Albert random graph of $N = 1000$ nodes and average degree $\langle k \rangle \approx 4$ in the case of observation of a 30%-fraction of nodes chosen at random uniformly and independently. Homogeneous parameters $\lambda = 0.7, \mu = 0.6, \lambda = 0.5, \mu = 0.6$ and $\lambda = 0.2, \mu = 0.6$. The solid magenta line represents the Similarity Sampling for epidemic realizations enabling the convergence for $a = 0.125$ (respectively $M_o = 18, M_o = 21, M_o = 36$); the dashed blue line represents the Similarity Sampling with $a = 0.125$ although convergence is not met (the average is computed over $M_o = 50$ epidemic realizations).

<https://doi.org/10.1371/journal.pone.0176376.g017>

that constrains the calculation to epidemics that vanish before T_{ext} . In this way, we give null probability to every single site configuration with $t_i + g_i$ larger than T_{ext} (except for $t_i = T_{inf}$ that describes susceptible nodes). The logarithm of the partition function is the free energy of the model, hence

$$-f(T_{ext}, \mathbf{y}^{obs}) = \log Z(T_{ext}, \mathbf{y}^{obs}) = \log P(t_{ext} < T_{ext} | \mathbf{y}^{obs}). \quad (31)$$

In the factor graph representation, the free-energy can be approximated with the Bethe free-energy, that is computed by means of the BP equations.

Similarity Sampling convergence

Figs 16 and 17 display the effects of forcing the convergence criterion with $a = 0.125$ in the case of regular random graphs and Barabasi-Albert graphs, respectively. The solid curves represent results obtained considering only the epidemic realizations for which convergence is met, whereas the dashed curves correspond to results obtained including all realizations (as in the main text). Only in the case of BA graphs with $\lambda = 0.5$ the two curves are not within the error bars. Nevertheless, this difference does not change the qualitative behaviour and the comparison with the BP method (see Figs 4 and 9).

Author Contributions

Conceptualization: JB AB LDA.

Data curation: JB AB LDA.

Formal analysis: JB AB LDA.

Funding acquisition: JB AB LDA.

Investigation: JB AB LDA.

Methodology: JB AB LDA.

Project administration: JB AB LDA.

Resources: JB AB LDA.

Software: JB AB LDA.

Supervision: JB AB LDA.

Validation: JB AB LDA.

Visualization: JB AB LDA.

Writing – original draft: JB AB LDA.

Writing – review & editing: JB AB LDA.

References

1. Barrat A, Cattuto C, Tozzi AE, Vanhems P, Voirin N. Measuring contact patterns with wearable sensors: methods, data characteristics and applications to data-driven simulations of infectious diseases. *Clinical Microbiology and Infection*. 2014; 20(1):10–16. Available from: <http://dx.doi.org/10.1111/1469-0691.12472>. PMID: [24267942](https://pubmed.ncbi.nlm.nih.gov/24267942/)
2. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences*. 2010; 107(51):22020–22025. <https://doi.org/10.1073/pnas.1009094108>

3. Stehlé J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton J, et al. High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. *PLOS ONE*. 2011 08; 6(8):e23176. Available from: <http://dx.doi.org/10.1371/journal.pone.0023176>.
4. Mastrandrea R, Fournet J, Barrat A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PloS one*. 2015; 10(9): e0136497. <https://doi.org/10.1371/journal.pone.0136497> PMID: 26325289
5. Colizza V, Barrat A, Barthélemy M, Vespignani A. The modeling of global epidemics: Stochastic dynamics and predictability. *Bulletin of mathematical biology*. 2006; 68(8):1893–1921. <https://doi.org/10.1007/s11538-006-9077-9> PMID: 17086489
6. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. *BMC medicine*. 2007; 5(1):34. <https://doi.org/10.1186/1741-7015-5-34> PMID: 18031574
7. Belik V, Geisel T, Brockmann D. Natural Human Mobility Patterns and Spatial Spread of Infectious Diseases. *Phys Rev X*. 2011 Aug; 1:011001. Available from: <http://link.aps.org/doi/10.1103/PhysRevX.1.011001>.
8. Holme P, Takaguchi T. Time evolution of predictability of epidemics on networks. *Physical Review E*. 2015; 91(4):042811. <https://doi.org/10.1103/PhysRevE.91.042811>
9. Holme P. Information content of contact-pattern representations and predictability of epidemic outbreaks. *arXiv preprint arXiv:150306583*. 2015;.
10. Altarelli F, Braunstein A, Dall'Asta L, Lage-Castellanos A, Zecchina R. Bayesian Inference of Epidemics on Networks via Belief Propagation. *Phys Rev Lett*. 2014 Mar; 112:118701. Available from: <http://link.aps.org/doi/10.1103/PhysRevLett.112.118701>. PMID: 24702425
11. Altarelli F, Braunstein A, Dall'Asta L, Ingrassio A, Zecchina R. The patient-zero problem with noisy observations. *Journal of Statistical Mechanics: Theory and Experiment*. 2014; 2014(10):P10016. Available from: <http://stacks.iop.org/1742-5468/2014/i=10/a=P10016>.
12. Braunstein A, Ingrassio A. Inference of causality in epidemics on temporal contact networks. *Scientific reports*. 2016; 6:27538. <https://doi.org/10.1038/srep27538> PMID: 27283451
13. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton JF, Vespignani A. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS one*. 2010; 5(7):e11596. <https://doi.org/10.1371/journal.pone.0011596> PMID: 20657651
14. Holme P, Saramäki J. Temporal networks. *Physics Reports*. 2012; 519(3):97–125. *Temporal Networks*. Available from: <http://www.sciencedirect.com/science/article/pii/S0370157312000841>.
15. Antulov-Fantulin N, Lančić A, Šmuc T, Štefančić H, Šikić M. Identification of Patient Zero in Static and Temporal Networks: Robustness and Limitations. *Physical review letters*. 2015; 114(24):248701. <https://doi.org/10.1103/PhysRevLett.114.248701> PMID: 26197016
16. Mezard M, Montanari A. *Information, physics, and computation*. Oxford University Press; 2009.
17. Altarelli F, Braunstein A, Dall'Asta L, Wakeling JR, Zecchina R. Containing Epidemic Outbreaks by Message-Passing Techniques. *Phys Rev X*. 2014 May; 4:021024. Available from: <http://link.aps.org/doi/10.1103/PhysRevX.4.021024>.
18. Holme P. Extinction times of epidemic outbreaks in networks. *PloS one*. 2013; 8(12). <https://doi.org/10.1371/journal.pone.0084429>
19. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic processes in complex networks. *Reviews of modern physics*. 2015; 87(3):925. <https://doi.org/10.1103/RevModPhys.87.925>
20. Rocha LE, Liljeros F, Holme P. Information dynamics shape the sexual networks of Internet-mediated prostitution. *Proceedings of the National Academy of Sciences*. 2010; 107(13):5706–5711. <https://doi.org/10.1073/pnas.0914080107>
21. Rocha LE, Liljeros F, Holme P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comput Biol*. 2011; 7(3):e1001109. <https://doi.org/10.1371/journal.pcbi.1001109> PMID: 21445228
22. Altarelli F, Braunstein A, Dall'Asta L, Zecchina R. Large deviations of cascade processes on graphs. *Physical Review E*. 2013; 87(6):062115. <https://doi.org/10.1103/PhysRevE.87.062115>
23. Altarelli F, Braunstein A, Dall'Asta L, Zecchina R. Optimizing spread dynamics on graphs by message passing. *Journal of Statistical Mechanics: Theory and Experiment*. 2013; 2013(09):P09011. <https://doi.org/10.1088/1742-5468/2013/09/P09011>