POLITECNICO DI TORINO
Repository ISTITUZIONALE

Small-coupling dynamic cavity: A Bayesian mean-field framework for epidemic inference

Original
Small-coupling dynamic cavity: A Bayesian mean-field framework for epidemic inference / Braunstein, A., Catania, G., Dall'Asta, L., Mariani, M., Mazza, F., Tarabolo, M.. - In: PHYSICAL REVIEW RESEARCH. - ISSN 2643-1564. - ELETTRONICO. - 7:2(2025), pp. 1-25. [10.1103/physrevresearch.7.023089]

Availability:
This version is available at: 11583/3000313 since: 2025-05-20T12:34:14Z

Publisher:
APS

Published
DOI:10.1103/physrevresearch.7.023089

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

25 November 2026

Small-coupling dynamic cavity: A Bayesian mean-field framework for epidemic inference

Alfredo Braunstein,[1,2,3] Giovanni Catania,[4] Luca Dall'Asta,[1,2,3] Matteo Mariani,[1,5] Fabio Mazza,[1,6] and Mattia Tarabolo[1,*]

[1]*Institute of Condensed Matter Physics and Complex Systems, Department of Applied Science and Technology, Politecnico di Torino, C.so Duca degli Abruzzi 24, I-10129 Torino, Italy*

[2]*Italian Institute for Genomic Medicine (IIGM) and Candiolo Cancer Institute IRCCS, str. prov. 142, km 3.95, I-10060 Candiolo (TO), Italy*

[3]*INFN, Turin Via Pietro Giuria 1, I-10125 Turin, Italy*

[4]*Departamento de Física Téorica, Universidad Complutense, 28040 Madrid, Spain*

[5]*Department of Neurobiology, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA*

[6]*Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via Ponzio 34/5, I-20133 Milano, Italy*

(Received 19 December 2023; accepted 26 March 2025; published 25 April 2025)

A novel generalized mean-field approximation, called the small-coupling dynamic cavity (SCDC) method, for epidemic inference and risk assessment is presented. The method is developed within a fully Bayesian framework and accounts for noncausal effects generated by the presence of observations. It is based on a graphical model representation of the epidemic stochastic process and utilizes dynamic cavity equations to derive a set of self-consistent equations for probability marginals defined on the edges of the contact graph. By performing a small-coupling expansion, a pair of time-dependent cavity messages is obtained, which capture the probability of individual infection and the conditioning power of observations. In its efficient formulation, the computational cost per iteration of the SCDC algorithm is linear in the duration of the epidemic dynamics. While the method is derived for the susceptible-infected (SI) model, it is straightforwardly applicable to many other Markovian epidemic processes, including recurrent ones. This linear complexity is particularly advantageous for recurrent epidemic processes, where inference methods are typically exponentially complex in the duration of the epidemic dynamics. The method exhibits high accuracy in assessing individual risk, as demonstrated by tests on the SI model applied to various classes of synthetic contact networks, where it performs on par with belief propagation techniques and generally exceeds the performance of individual-based mean-field methods. Additionally, the method was applied to recurrent epidemic models, where it showed interesting performance even for relatively large values of the infection probability, highlighting its versatility and effectiveness in challenging scenarios. Although convergence issues may arise due to long-range correlations in contact graphs, the estimated marginal probabilities remain sufficiently accurate for reliable risk estimation. Future work includes extending the method to non-Markovian recurrent epidemic models and investigating the role of second-order terms in the small-coupling expansion of the observation-reweighted dynamic cavity equations.

DOI: 10.1103/PhysRevResearch.7.023089

I. INTRODUCTION

In the past decade, the increasing availability of detailed epidemiological data and high-accuracy contact-network datasets triggered the study of individual-based epidemic inference problems, which aim to predict the marginal probabilities of infection for individual nodes in a network, given partial observations of the epidemic, such as the states of specific individuals at certain times. The interest has been further stimulated during the COVID-19 pandemic, by the possibility of performing massive epidemic surveillance and digital contact tracing via smartphone applications [1,2]. A variety of computational methods were proposed for tackling this class of inference problems, such as heuristic algorithms based on network centrality measures [3,4], generalized mean-field approximations [5], Monte Carlo methods [6], and machine learning techniques exploiting tailored architectures of autoregressive and graph neural networks [7–11]. The leading techniques for epidemic inference adopts a Bayesian framework, with a simple individual-based epidemic model as prior distribution and the sparse observation of positive/negative test results as evidence [12]. Unfortunately, the associated inference task includes the formidable computational difficulty of computing averages on a complicated high dimensional discrete distribution for which no tractable analytic solution is known. Arguably, generalized mean-field methods and their associated algorithms, are best able to balance the computational demand of accurately estimating these observables with affordable calculation times. In particular, belief propagation (BP) [13–16] has proven to be extremely effective in estimating local marginal probabilities of the posterior distribution, reconstructing the infection state of unobserved individuals,

*Contact author: mattia.tarabolo@polito.it

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

and identifying patient zero and contagion channels. Furthermore, when integrated into the framework of digital contact tracing for COVID-19, the BP-based algorithms have been shown to provide a better assessment of individual risk and improve the mitigation impact of nonpharmaceutical intervention strategies, outperforming competing methods for various epidemic inference problems defined on contact networks [2,17]. BP-based algorithms have also been used to optimally deploy resources for constraining the epidemic spreading to hit a subset of target nodes at specific times [18], similarly to the constraints imposed by observations in inference. The belief propagation approach to spatiotemporal epidemic trajectories can be classified as a generalized mean-field method because the epidemic trajectories of the neighbors of a given individual are assumed to be conditionally independent. This hypothesis is correct when the dynamical process takes place on a contact network without cycles, but the method proved to be very effective also on contact networks with cycles. It should be noted that the dynamic cavity (DC) approach [19,20] turns out to be equivalent to BP in the case of pure time-forward epidemic dynamics without observations. Moreover, for dynamic models with nonrecurrent individual states, in the case of pure time-forward epidemic dynamics without observations, the BP/DC approach simplifies into a dynamic message passing technique that has been extensively used to study spreading processes on networks [21,22].

Despite the apparent success of the BP approach, simpler and faster methods may be preferable, in particular in views of the potential need of real-time calculation for large communities (with thousands to millions of individuals). Moreover, BP is not easily generalizable to recurrent dynamics (such as SIS and SIRS), in which even the trajectory of a single individual can potentially take an exponential number of states. In this regime, further assumptions are required. One possibility is a tensorial approximation of distributions of such trajectories, which can however result in relatively heavy computational requirements [23]. In this context, simpler individual-based mean-field (IBMF) methods may turn out to be of interest. Also known as quenched mean-field and N-intertwined models, these methods simply assume that the simultaneous states of neighboring nodes are statistically independent, and have been shown to provide reasonably accurate approximations to time-forward epidemic dynamics [24–27]. Recently, the individual-based mean-field method has been applied to propose a straightforward inference approach that heuristically incorporates observations of individual states [2]. While this method does not explicitly rely on a Bayesian framework, its simplicity and computational efficiency make it a practical tool for epidemic risk assessment, yielding promising results in relevant scenarios.

In this paper we propose small-coupling dynamic cavity (SCDC), a generalized mean-field approximation for Bayesian epidemic inference and risk assessment. The method is introduced to address the limitations of BP in handling recurrent epidemic processes and to ensure uniform computational complexity for both recurrent and nonrecurrent models. Moreover, SCDC offers a simple and interpretable approximation of BP, where the messages have a clear physical interpretation, making it particularly appealing for a better understanding of the inference algorithm. The method exploits the improved uncorrelation assumption of BP/DC compared to simpler mean-field approaches, while largely preserving the analytic simplicity and computational efficiency of IBMF. The starting point of the method is a graphical model representation of the epidemic stochastic process that allows for a convenient derivation of a set of dynamic cavity equations for functional probability marginals defined on the edges of the contact graph. In this representation, the cavity probability marginal on a directed edge from individual i to individual j is a function of two quantities: the trajectory of the individual state of i in the absence of interactions with j and a conjugate external field acting on i (which replaces the effect of the missing interaction terms in the cavity graph). By performing an expansion of the dynamic cavity equations for weak infection probabilities and truncating the expansion at the first order, a set of self-consistent equations for the average of these two quantities can be obtained. Unlike other individual-based mean-field approximations, the proposed method is conceived as an approximation of a fully Bayesian framework. While both SCDC and IBMF account for noncausal effects generated by the presence of observations, the Bayesian foundation of SCDC makes it a promising basis for future developments, such as improved approximations or theoretical studies of epidemiological inference. These noncausal effects arise because incorporating observations constrains the dynamics to trajectories compatible with the observed data, necessitating the propagation of information backward in time, which in our formulation is encoded through the conjugate fields. In the absence of observations, the conjugate fields responsible for noncausal dynamics vanish, and the individual-based mean-field method for time-forward dynamics can be recovered. For simplicity, we initially present the SCDC in the case of the susceptible-infected (SI) model. Using an efficient formulation based on a transfer-matrix technique, the SCDC method is then extended to general Markovian epidemic processes, including individual recovery, latency, and recurrent infection (e.g., SIR, SEIR, SIS, and SIRS models). Although SCDC is expected to be less accurate than BP (on nonrecurrent models on which it can be applied), for which it represents a sort of weak-infectivity approximation for the distribution of individual trajectories, we show through extensive simulations on realistic examples how this gap is often negligible.

The manuscript is organized as follows: Section II presents the SCDC method and its derivation on the SI model; Sec. III discusses a general efficient formulation of the algorithm that is easy to generalize to other epidemic models (further discussed in Sec. IV); results are presented in Sec. V, concerning both estimates of epidemic outbreaks in the absence of observations and individual risk assessment from partial observations, on both irreversible and recurrent compartmental models; finally, Sec. VI draws the conclusions and highlights future directions to be investigated.

II. METHODOLOGY

A. Definition of the stochastic epidemic model and observations

For simplicity, we will present the proposed method on one of the simplest epidemic models, namely the discrete-time stochastic SI model. The generalization to more general

epidemic models is discussed in Sec. IV. It is thus considered the dynamics of the SI model on a population of N individuals over a temporal window of T time steps (which, for the sake of concreteness, will be called "days"). The daily contacts are directly encoded in the set of parameters specifying the infection transmission, with λ_{ij}^t being the infection probability along the directed edge from individual i to individual j at time t; conversely, we set $\lambda_{ij}^t = \lambda_{ji}^t = 0$ if i and j are not in contact at time t. The epidemic state of the population at time t is represented by a binary vector $\boldsymbol{x}^t = (x_1^t, \ldots, x_N^t)$, with $x_i^t = S$ (respectively, $x_i^t = I$) meaning that i is a susceptible (respectively, infected) individual at time t. We include in the model a small self-infection probability ε_i^t. In general, self-infections can model scenarios where the contact network is only partially known, in such a way to avoid incompatibility between the inference model and observations arising from unknown transmission channels; however, a small self-infection can also be useful to heal pathological cases that can arise in the presence of noise-less observations (more details in Appendix F). The epidemic model is assumed to be Markovian, i.e., the time evolution of the probability $p_t(\boldsymbol{x}^t)$ that the population is in state $\boldsymbol{x}^t$ at time t is given in terms of the following master equation:

$$p_{t+1}(\boldsymbol{x}^{t+1}) = \sum_{\boldsymbol{x}^t \in \{S,I\}^N} W(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t) p_t(\boldsymbol{x}^t), \tag{1}$$

with transition rates factorized over nodes $W(\boldsymbol{x}^{t+1}|\boldsymbol{x}^t) = \prod_i W_i(x_i^{t+1}|\boldsymbol{x}^t)$, where

$$W_i\big(x_i^{t+1} = S|\boldsymbol{x}^t\big) = \delta_{x_i^t,S}\alpha_i^t e^{h_i^t}, \tag{2}$$

$$\begin{aligned} W_i\big(x_i^{t+1} = I|\boldsymbol{x}^t\big) &= \delta_{x_i^t,I} + \delta_{x_i^t,S}\big[1 - \alpha_i^t e^{h_i^t}\big] \\ &= 1 - \delta_{x_i^t,S}\alpha_i^t e^{h_i^t}, \end{aligned} \tag{3}$$

where $\delta_{x,y}$ denotes the Kronecker symbol. For notational simplicity, we introduced the probability of not being self-infected $\alpha_i^t = 1 - \varepsilon_i^t$ and the local field $h_i^t = \sum_{j=1}^N \nu_{ji}^t \delta_{x_j^t,I}$, where $\nu_{ji}^t = \log(1 - \lambda_{ji}^t)$, such that $e^{h_i^t} = \prod_{j=1}^N (1 - \lambda_{ji}^t \delta_{x_j^t,I})$ is the probability of not being infected by the neighbors. Then the probability of the epidemic spreading $\boldsymbol{x} = \{\boldsymbol{x}^0, \boldsymbol{x}^1, \ldots, \boldsymbol{x}^T\}$ can be written, under the Markov assumption, in the following form:

$$p(\boldsymbol{X}) = \prod_{t=0}^{T} p_t(\boldsymbol{x}^t) = \prod_i p_0\big(x_i^0\big) \prod_{t=0}^{T-1} W_i\big(x_i^{t+1}|\boldsymbol{x}^t\big), \tag{4}$$

where, at the initial time $t = 0$, the individuals are assumed to be independent and identically distributed according to $p_0(\boldsymbol{x}^0)$. The likelihood of the model can be defined by a set $\mathcal{O} = \{\boldsymbol{O}_i\}_{i=1,\ldots,N}$ of site-independent observations. We assume that each node i can be observed multiple times across different time steps. Observations are factorized over time as $\boldsymbol{O}_i = (O_i^0, \ldots, O_i^T)$, where $O_i^t = \varnothing$ denotes the absence of observation at time t. A more general scenario admits uncertainty on the outcome of the tests, the latter being eventually quantified by false positive rate η^+ and/or false negative rates η^-. The likelihood of an observation O_i^t on node i at time t can be written as

$$p\big(O_i^t|x_i^t\big) = \begin{cases} (1-\eta^+)\delta_{x_i^t,S} + \eta^-\delta_{x_i^t,I} & \text{if } O_i^t = S, \\ \eta^+\,\delta_{x_i^t,S} + (1-\eta^-)\delta_{x_i^t,I} & \text{if } O_i^t = I, \\ 1 & \text{if } O_i^t = \varnothing. \end{cases} \tag{5}$$

The likelihood over the full set of observations $\mathcal{O}$ can be written as $p(\mathcal{O}|\boldsymbol{X}) = \prod_{i=1}^N \prod_{t=0}^T p(O_i^t|x_i^T)$. In the case of perfectly accurate tests, in which $\eta^+ = \eta^- = 0$, the effect of the observations is to enforce the dynamical trajectories to be compatible with the observed states.

The posterior probability of the trajectory $\boldsymbol{X}$ can be expressed using Bayes' theorem as follows:

$$p(\boldsymbol{X}|\mathcal{O}) = \frac{1}{p(\mathcal{O})} p(\boldsymbol{X}) p(\mathcal{O}|\boldsymbol{X}) \tag{6a}$$

$$\propto \prod_{i=1}^{N} p_0\big(x_i^0\big) p\big(O_i^T|x_i^T\big) \times \prod_{t=0}^{T-1} W_i\big(x_i^{t+1}|\boldsymbol{x}^t\big) p\big(O_i^t|x_i^t\big). \tag{6b}$$

The Bayesian inference problem consists in evaluating marginals of the posterior distribution $p(\boldsymbol{X}|\mathcal{O})$, such as the quantity $p(x_i^t = x|\mathcal{O})$ representing the posterior probability that individual i is in state $x \in \{S, I\}$ at time t given the set of the available observations $\mathcal{O}$. The posterior distribution is, in general, intractable but it is the starting point for the derivation of approximate inference methods.

B. Dynamic cavity equations for the SI model with observations

The posterior probability in Eq. (6b) can be interpreted as a graphical model for dynamical trajectories defined on the contact network. In this context, a BP approach was proposed in Refs. [13,14], and has since been employed for epidemic inference. This approach involves defining messages between neighboring nodes that depend on their pairwise dynamical trajectories, allowing short loops induced by the dynamical constraints in the factor graph to be disentangled.

In this work, we propose an alternative formulation based on a cavity argument, which yields a set of equations for the marginal probability of a pair of variable-field trajectories on each node in the cavity graph. This formulation, referred to as the *observation-reweighted DC equations*, provides a rigorous approach to describing the posterior distribution on sparse graphs. Starting from the posterior probability in Eq. (6b), the derivation (see Appendix A and Refs. [2,13,14]) leads to the following set of equations for the cavity messages $\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\mathcal{O})$:

$$\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\mathcal{O}) \propto p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \sum_{\boldsymbol{x}_{\partial i\setminus j}} e^{\hat{S}_{i\setminus j}} \times \prod_{k\in\partial i\setminus j} \hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\mathcal{O}), \tag{7}$$

where $\partial i \setminus j$ is the set of neighboring indices of i except for j, $\sum_{\boldsymbol{x}_{\partial i\setminus j}} = \prod_{k\in\partial i\setminus j} \prod_{t=0}^T \sum_{x_k^t\in\{S,I\}}$, and $\hat{S}_{i\setminus j}$ is the cavity dynamical action of the spreading dynamics under study, which for

the SI model can be written as

$$\hat{S}_{i\backslash j} = \sum_{t=0}^{T-1} \log\left[\alpha_i^t e^{s_i^t + \sum_{k\in\partial i\backslash j} \nu_{ki}^t \delta_{x_k^t,I}} \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta_{x_i^{t+1},I}\right]. \tag{8}$$

For the sake of brevity, we introduced the compact representations $\boldsymbol{\nu}_{ik}\boldsymbol{x}_i = (\nu_{ik}^0\delta_{x_i^0,I}, \ldots, \nu_{ik}^T\delta_{x_i^T,I})$ and $p(\boldsymbol{O}_i|\boldsymbol{x}_i) = \prod_{t=0}^T p(O_i^t|x_i^t)$. The terms $\hat{c}_{i\backslash j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\boldsymbol{\mathcal{O}})$ are the cavity messages, depending on the trajectory $\boldsymbol{x}_i$ of individual i and the local field trajectory $\boldsymbol{s}_i$, which acts as a proxy for the trajectory of the missing neighboring node j (more precisely, $\boldsymbol{s}_i \propto \boldsymbol{\nu}_{ji}\boldsymbol{x}_j$).

The observation-reweighted dynamic cavity equations can be interpreted as a message-passing procedure, analogous to BP, where messages are iteratively exchanged between nodes until convergence. Importantly, as in BP, the normalization of the messages at each step is not critical and can be adjusted after convergence. Additionally, these equations recover the BP messages through a change of variables, with $\boldsymbol{s}_i$ effectively replacing the explicit dependence on the neighboring node's trajectory. A detailed discussion of this equivalence is provided in Appendices A and B.

While the dynamic cavity equations with variable-field trajectories were originally proposed only for time-forward binary spin dynamics [19,20], those in Eq. (7) also account for probabilistic reweighting due to the observations.

Finally, completing the cavity and computing the total marginal over i gives the posterior marginal probability of one-site trajectories

$$c_i(\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}) = \frac{p_0(x_i^0)}{\mathcal{Z}_i(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) \sum_{\boldsymbol{x}_{\partial i}} e^{S_i} \prod_{k\in\partial i} \hat{c}_{k\backslash i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}), \tag{9}$$

where the summation runs over the trajectories of all neighbors of i, i.e., $\boldsymbol{x}_{\partial i} = \{\boldsymbol{x}_k\}_{k\in\partial i}$. The term S_i represents the local dynamical action, which for the SI model is given by

$$S_i = \sum_{t=0}^{T-1} \log\left[\alpha_i^t e^{\sum_{k\in\partial i} \nu_{ki}^t \delta_{x_k^t,I}} \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta_{x_i^{t+1},I}\right]. \tag{10}$$

The normalization factor is defined as

$$\mathcal{Z}_i(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} p_0(x_i^0) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \sum_{\boldsymbol{x}_{\partial i}} e^{S_i} \times \prod_{k\in\partial i} \hat{c}_{k\backslash i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}). \tag{11}$$

C. Small-coupling expansion

To simplify the cavity equations in Eq. (7), we express the cavity messages in terms of the conjugate field trajectories $\hat{\boldsymbol{h}}_i = (\hat{h}_i^0, \ldots, \hat{h}_i^T)$ by introducing their Fourier transform with respect to $\boldsymbol{s}_i$ (see Appendix B). This yields the following expression for the dynamic cavity equations:

$$c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) = \frac{p_0(x_i^0)}{\mathcal{Z}_{i\backslash j}(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^0} \times \prod_{k\in\partial i\backslash j} \sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k\, c_{k\backslash i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}}) e^{S_{ik}^{\text{int}}}, \tag{12}$$

where $\sum_{\boldsymbol{x}_i} = \prod_{t=0}^T \sum_{x_i^t\in\{S,I\}}$ and $\int D\hat{\boldsymbol{h}}_i = \prod_{t=0}^T \int_{-\infty}^{\infty} d\hat{h}_i^t$. We used the following definition of the Fourier transform and its inverse:

$$c(\boldsymbol{x}, \hat{\boldsymbol{h}}|\boldsymbol{\mathcal{O}}) = \int D\boldsymbol{s} \left(\prod_{t=0}^T \frac{e^{\mathrm{i}s^t\hat{h}^t}}{2\pi}\right) \hat{c}(\boldsymbol{x}, \boldsymbol{s}|\boldsymbol{\mathcal{O}}), \tag{13}$$

$$\hat{c}(\boldsymbol{x}, \boldsymbol{s}|\boldsymbol{\mathcal{O}}) = \int D\hat{\boldsymbol{h}} \left(\prod_{t=0}^T e^{-\mathrm{i}s^t\hat{h}^t}\right) c(\boldsymbol{x}, \hat{\boldsymbol{h}}|\boldsymbol{\mathcal{O}}). \tag{14}$$

The local noninteracting and interacting actions are defined by the epidemic model under study. For the SI model they can be written respectively as

$$S_i^0 = \sum_{t=0}^{T-1} \log\left[\delta(\hat{h}_i^t - \mathrm{i})\alpha_i^t \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta(\hat{h}_i^t)\delta_{x_i^{t+1},I}\right], \tag{15}$$

$$S_{ik}^{\text{int}} = S_{ki}^{\text{int}} = -\mathrm{i}\sum_{t=0}^{T-1} \left(\delta_{x_i^t,I}\nu_{ik}^t\hat{h}_k^t + \delta_{x_k^t,I}\nu_{ki}^t\hat{h}_i^t\right). \tag{16}$$

Here $\delta(x)$ denotes the Dirac δ function. To ensure proper normalization of the cavity messages $c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}})$, we define the normalization factor $\mathcal{Z}_{i\backslash j}(\boldsymbol{\mathcal{O}})$ such that

$$\sum_{\boldsymbol{x}_i} \int D\hat{\boldsymbol{h}}_i\, c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) = 1. \tag{17}$$

This condition ensures that $c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}})$ represents a properly normalized quasi-probability distribution. Consequently, the expectation of any generic function $f(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)$ over the cavity distribution can be computed as

$$\langle f(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)\rangle_{i\backslash j}^{\boldsymbol{\mathcal{O}}} = \sum_{\boldsymbol{x}_i} \int D\hat{\boldsymbol{h}}_i\, c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) f(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i). \tag{18}$$

By applying the inverse Fourier transform to Eq. (17), we find that the normalization condition is equivalent to assuming that the neighboring node j remains in the susceptible state throughout the process in the cavity graph. This corresponds to setting the local field to zero, $\boldsymbol{s}_i = \boldsymbol{0}$

$$\sum_{\boldsymbol{x}_i} \int D\hat{\boldsymbol{h}}_i\, c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} \hat{c}_{i\backslash j}(\boldsymbol{x}_i, \boldsymbol{s}_i = \boldsymbol{0}|\boldsymbol{\mathcal{O}}) = 1. \tag{19}$$

Under the small-coupling approximation, the dynamic cavity method assumes that the absence of node j in the cavity graph implies no influence on the dynamics of node i, as j is treated as always susceptible. In contrast, the original dynamic cavity method [20] explicitly incorporates the fixed trajectories of the missing neighbor j and updates cavity messages through a self-consistent dynamic message-passing algorithm. However, this approach has a computational complexity that can potentially grow exponentially with the number of time steps T. On the one hand, for nonrecurrent dynamics this complexity can be mitigated by reformulating the cavity messages in terms of flipping times [21], offering a more efficient representation. On the other hand, this flipping-time representation is not applicable to recurrent dynamics. Choosing the normalization to be independent of the trajectory of individual

j is highly advantageous for developing an approximation that does not explicitly rely on the epidemic trajectories of both i and j. However, the main drawback of this choice is that some particular epidemic trajectories of individual i, imposed by the observations, can only be explained by the statistical model if a nonzero self-infection probability is introduced. In Appendix F, we provide an example of this potential issue that is simple enough to be discussed analytically and show how the presence of a self-infection probability effectively resolves it.

In the spirit of Plefka's approach [28,29] and high-temperature expansions [30,31], one can perform a formal expansion of the interactive exponential term. Since the argument of the exponential is linear in the parameters ν_{ik}^t and ν_{ki}^t, truncating the expansion at some finite order can be understood as a small-coupling approximation of the dynamic cavity equations, which in the case of epidemic processes corresponds to a small infectivity approximation (i.e., $\lambda_{ij}^t \ll 1$ for all i, j and t). Truncating the Taylor series at the first order, after re-exponentiation of the interacting part we obtain an approximate expression for the cavity messages

$$c_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) \approx \frac{p_0(x_i^0)}{\tilde{\mathcal{Z}}_{i\backslash j}(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^0 + \sum_{k\in\partial i\backslash j} \langle S_{ik}^{\text{int}}\rangle_{k\backslash i}^{\boldsymbol{\mathcal{O}}}} := \tilde{c}_{i\backslash j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}), \tag{20}$$

where also the normalization constant is consistently approximated to $\tilde{\mathcal{Z}}_{i\backslash j}(\boldsymbol{\mathcal{O}})$. The averaged interacting action is a function of both the variable-conjugate field trajectories $(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)$ and of a set of local one-time cavity averages coming from the neighbors, namely $m_{k\backslash i}^t(\boldsymbol{\mathcal{O}}) = \langle \delta_{x_k^t,I}\rangle_{k\backslash i}^{\boldsymbol{\mathcal{O}}}$ and $\mu_{k\backslash i}^t(\boldsymbol{\mathcal{O}}) = \langle -\mathrm{i}\hat{h}_k^t\rangle_{k\backslash i}^{\boldsymbol{\mathcal{O}}}$. For the SI model, this yields

$$\langle S_{ik}^{\text{int}}\rangle_{k\backslash i}^{\boldsymbol{\mathcal{O}}} = \sum_{t=0}^{T-1} \left(\delta_{x_i^t,I}\nu_{ik}^t \mu_{k\backslash i}^t(\boldsymbol{\mathcal{O}}) - \mathrm{i} m_{k\backslash i}^t(\boldsymbol{\mathcal{O}})\nu_{ki}^t \hat{h}_i^t\right). \tag{21}$$

The dynamic cavity equations in Eq. (20), derived under the small-coupling approximation, can be closed self-consistently by evaluating the cavity averages using the approximated distribution functions

$$\begin{aligned} m_{k\backslash i}^t(\boldsymbol{\mathcal{O}}) &\approx \sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k \tilde{c}_{k\backslash i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}}) \delta_{x_k^t,I} \\ &\approx \sum_{\boldsymbol{x}_k} \hat{\tilde{c}}_{k\backslash i}(\boldsymbol{x}_k, \boldsymbol{s}_k = \boldsymbol{0}|\boldsymbol{\mathcal{O}}) \delta_{x_k^t,I}, \end{aligned} \tag{22}$$

$$\begin{aligned} \mu_{k\backslash i}^t(\boldsymbol{\mathcal{O}}) &\approx \sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k \tilde{c}_{k\backslash i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}}) \left(-\mathrm{i}\hat{h}_k^t\right) \\ &\approx \sum_{\boldsymbol{x}_k} \frac{\partial}{\partial s_k^t} \hat{\tilde{c}}_{k\backslash i}(\boldsymbol{x}_k, \boldsymbol{s}_k|\boldsymbol{\mathcal{O}})|_{\boldsymbol{s}_k=\boldsymbol{0}}, \end{aligned} \tag{23}$$

where $\hat{\tilde{c}}_{k\backslash i}(\boldsymbol{x}_k, \boldsymbol{s}_k|\boldsymbol{\mathcal{O}})$ is the approximated cavity message Eq. (7), obtained by the inverse Fourier transform of Eq. (20).

The quantity $m_{k\backslash i}^t$ measures the average probability that node k is infected at time t in the absence of the interaction with i, while $\mu_{k\backslash i}^t$ is of less intuitive interpretation, as it measures the mean of fluctuations around the unperturbed single-site statistics in the cavity graph. A direct calculation of the quantity $\sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k (-\mathrm{i}\hat{h}_k^t) \tilde{c}_{k\backslash i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}} = \varnothing)$ shows that, in the absence of observations, $\mu_{k\backslash i}^t(\boldsymbol{\mathcal{O}} = \varnothing) = 0$. This result, due to the causality of the dynamical process, does not hold anymore when some observations are included.

For notational convenience, we will omit the explicit dependence of all marginals, normalization constants, and cavity averages on the observation set $\boldsymbol{\mathcal{O}}$ in the next sections. Similarly, approximated quantities will no longer be distinguished by the $\tilde{}$ notation, as all references will pertain to their approximated forms unless explicitly stated otherwise.

D. Small-coupling dynamic cavity (SCDC) approximation

By implementing the closure assumption, where the cavity averages $m_{i\backslash j}^t$ and $\mu_{i\backslash j}^t$ are replaced by their expectations over the approximated cavity messages Eq. (20), the small-coupling expansion yields a closed set of dynamic equations for the cavity averages. This becomes evident when the approximated cavity messages Eq. (20) are substituted into the expressions for the cavity averages Eqs. (22) and (23). After integrating out the conjugate fields $\boldsymbol{h}_i$,

$$m_{i\backslash j}^t(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} \delta_{x_i^t,I} \frac{p_0(x_i^0)}{\mathcal{Z}_{i\backslash j}(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_{i\backslash j}^m}, \tag{24}$$

$$\mu_{i\backslash j}^t(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} \frac{p_0(x_i^0)}{\mathcal{Z}_{i\backslash j}(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_{i\backslash j}^\mu}, \tag{25}$$

where we introduced the two terms

$$\begin{aligned} S_{i\backslash j}^m = &\sum_{t=0}^{T-1} \log\left[\alpha_i^t e^{\sum_{k\in\partial i\backslash j} \nu_{ki}^t m_{k\backslash i}^t} \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^t,I}\right) + \delta_{x_i^{t+1},I}\right] \\ &+ \sum_{t=0}^{T-1} \sum_{k\in\partial i\backslash j} \delta_{x_i^t,I} \nu_{ik}^t \mu_{k\backslash i}^t, \end{aligned} \tag{26}$$

$$\begin{aligned} S_{i\backslash j}^\mu = &\sum_{t=0}^{T-1} \log\left[\alpha_i^t \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^t,I}\right)\right] \\ &+ \sum_{t=0}^{T-1} \sum_{k\in\partial i\backslash j} \left(\nu_{ki}^t m_{k\backslash i}^t + \delta_{x_i^t,I} \nu_{ik}^t \mu_{k\backslash i}^t\right). \end{aligned} \tag{27}$$

The normalization constant is chosen to ensure that the time-dependent quantity $m_{i\backslash j}^t$ represents the probability of finding individual i infected at time t in the cavity graph, that is

$$\mathcal{Z}_{i\backslash j}(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} p_0(x_i^0) p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_{i\backslash j}^m}. \tag{28}$$

In addition, the total time-dependent marginal m_i^t of the posterior distribution on the full graph is given by

$$m_i^t(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} \delta_{x_i^t,I} \frac{p_0(x_i^0)}{\mathcal{Z}_i(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^m}, \tag{29}$$

with the quantity S_i^m and the normalization factor $\mathcal{Z}_i$ defined respectively as

$$\begin{aligned} S_i^m = &\sum_{t=0}^{T-1} \log\left[\alpha_i^t e^{\sum_{k\in\partial i} \nu_{ki}^t m_{k\backslash i}^t} \left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^t,I}\right) + \delta_{x_i^{t+1},I}\right] \\ &+ \sum_{t=0}^{T-1} \sum_{k\in\partial i} \delta_{x_i^t,I} \nu_{ik}^t \mu_{k\backslash i}^t, \end{aligned} \tag{30}$$

$$Z_i(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{x}_i} p_0(x_i^0) p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^m}. \tag{31}$$

Equations (24) and (25) represent a set of self-consistent equations defining a noncausal dynamic mean-field approximation that, in what follows, we will refer to as the SCDC method. The dynamical equations are of mean-field type since correlations are neglected, but in the presence of observations, they describe a noncausal dynamical process. Because of the cavity construction, the fundamental unknown of the equations, the one-time cavity marginals $m^t_{i\setminus j}$ and the one-time cavity fields $\mu^t_{i\setminus j}$, are defined by means of local self-consistent conditions, the solution of which can be sought through an iterative message-passing scheme. A computational bottleneck of Eqs. (24) and (25) is represented by the partial trace over single-site trajectories $\boldsymbol{x}_i$, that requires $O(2^T)$ operations, meaning that a complete update of all cavity quantities requires $O(2|E|T2^T)$, where $|E|$ is the total number of nonzero weighted directed edges on the interaction graph. An efficient algorithmic implementation of the SCDC equations is proposed in the next section. It exploits a transfer-matrix approach to perform the trace over the trajectory $\boldsymbol{x}_i$ keeping fixed all quantities $\{m^t_{k\setminus i}\}$ and $\{\mu^t_{k\setminus i}\}$ for all $t = 0, \ldots, T$ and $k \in \partial i \setminus j$, which play the role of the parameters in a "temporal" one-dimensional discrete probabilistic model defined on node i.

III. EFFICIENT FORMULATION OF THE SCDC EQUATIONS

The starting point of the derivation is the cavity normalization constant in Eqs. (24) and (25), that can be written starting from Eq. (28) as

$$\mathcal{Z}_{i\setminus j} = \sum_{\boldsymbol{x}_i} p_0\big(x_i^0\big) \prod_{t=0}^{T-1} M^{i\setminus j}_{x_i^t x_i^{t+1}} p\big(O_i^T | x_i^T\big), \tag{32}$$

where the "transfer matrix" $M^{i\setminus j}_{x^t x^{t+1}}$ is defined as follows:

$$M^{i\setminus j}_{x_i^t x_i^{t+1}} = \begin{pmatrix} M^{i\setminus j}_{t,SS} & M^{i\setminus j}_{t,SI} \\ M^{i\setminus j}_{t,IS} & M^{i\setminus j}_{t,II} \end{pmatrix} = \begin{pmatrix} \alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i} \nu^t_{ki}} p\big(O_i^t | S\big) & \big(1 - \alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}}\big) p\big(O_i^t | S\big) \\ 0 & e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p\big(O_i^t | I\big) \end{pmatrix}. \tag{33}$$

The approximate probability that an individual i is infected at time t in the cavity graph is

$$m^t_{i\setminus j} = \frac{\rho^{i\setminus j}_{\to t}(I)\rho^{i\setminus j}_{t\leftarrow}(I)}{\sum_{x_i^t\in\{S,I\}} \rho^{i\setminus j}_{\to t}\big(x_i^t\big)\rho^{i\setminus j}_{t\leftarrow}\big(x_i^t\big)}. \tag{34}$$

The term $\rho^{i\setminus j}_{\to t}(x_i^t)$ is a cavity forward "temporal" message, obtained recursively for $t = 1, \ldots, T$ as

$$\rho^{i\setminus j}_{\to t}\big(x_i^t\big) = \sum_{x_i^{t-1}\in\{S,I\}} \rho^{i\setminus j}_{\to t-1}\big(x_i^{t-1}\big) M^{i\setminus j}_{x_i^{t-1}x_i^t}, \tag{35}$$

with initial condition $\rho^{i\setminus j}_{\to 0}(x_i^0) = p(x_i^0)$, and $\rho^{i\setminus j}_{t\leftarrow}(x_i^t)$ is a cavity backward "temporal" message, that satisfy for $t = 0, \ldots, T-1$ the recursive equation

$$\rho^{i\setminus j}_{t\leftarrow}\big(x_i^t\big) = \sum_{x_i^{t+1}\in\{S,I\}} \rho^{i\setminus j}_{t+1\leftarrow}\big(x_i^{t+1}\big) M^{i\setminus j}_{x_i^t x_i^{t+1}}, \tag{36}$$

with terminal condition $\rho^{i\setminus j}_{T\leftarrow}(x_i^T) = p(O_i^T | x_i^T)$. The approximate one-time cavity field $\mu^t_{i\setminus j}$ can be computed as follows:

$$\mu^t_{i\setminus j} = \frac{\rho^{i\setminus j}_{\to t}(S) M^{i\setminus j}_{t,SS}\big(\rho^{i\setminus j}_{t+1\leftarrow}(S) - \rho^{i\setminus j}_{t+1\leftarrow}(I)\big)}{\sum_{x_i^t} \rho^{i\setminus j}_{\to t}\big(x_i^t\big)\rho^{i\setminus j}_{t\leftarrow}\big(x_i^t\big)}. \tag{37}$$

The number of operations necessary for the update of a single-time cavity message is now of $O(1)$ when the single-site "temporal" messages $\rho^{i\setminus j}_{\to s}(x_i^s)$ and $\rho^{i\setminus j}_{s\leftarrow}(x_i^s)$ for $s = t, t+1$ are available. The latter quantities are computed by means of time-forward and time-backward update rules from the current set of cavity "temporal" messages and require $O(4T)$ operations. In summary, a complete update of the cavity marginals $m^t_{i\setminus j}$ and cavity fields $\mu^t_{i\setminus j}$, for every directed edge and time step, requires $O(4|E|T)$.

A Julia-based reference implementation of the above algorithm is available at [32]. Notice that from Eq. (37), $\mu^t_{i\setminus j}$ is zero if the time-backward cavity messages at time $t+1$ are uniform. It is shown in Sec. V A and in Appendix C that this condition is satisfied when no observations are present at later times and it leads to a pure time-forward reduction of the SCDC equations. Furthermore, because of the nonrecurrent property of the SI model, it is possible to derive an alternative efficient formulation of the SCDC equations exploiting the infection-time representation: This is explained in detail in Appendix D.

IV. GENERALIZATION TO OTHER EPIDEMIC MODELS

The method generalizes directly to models with a higher number of individual states and transitions. In particular, for the susceptible-infected-susceptible (SIS) and susceptible-infected-removed (SIR) models, there is only one additional transition where an infected individual i can recover at time t with probability r_i^t, with the result that the individual i is either susceptible again (for the SIS model) or in a state of acquired immunity, i.e., $x_i^t = R$ (for the SIR model). A further generalization can be made for the susceptible-infected-removed-susceptible (SIRS), where each recovered individual i can return to the susceptible state at time t due to loss of immunity with probability σ_i^t.

The only difference with the method described in the previous sections is in the expression of the transfer matrix.

For the SIS model it is now given by

$$M^{i\setminus j}_{x_i^t x_i^{t+1}} = \begin{pmatrix} \alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}} p(O_i^t|S) & \left[1-\alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}}\right] p(O_i^t|S) \\ r_i^t e^{\sum_{k\in\partial i\setminus j} \mu^t_{k\setminus i}\nu^t_{ik}} p(O_i^t|I) & (1-r_i^t) e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p(O_i^t|I) \end{pmatrix}. \tag{38}$$

In the case of the SIR model, each individual i can be in three possible states $x_i^t \in \{S, I, R\}$, but the derivation done for two-state models can be repeated almost straightforwardly (see Appendix E for details). The cavity averages are defined analogously to the SI model through Eqs. (34) and (37) where the quantities $\rho^{i\setminus j}_{\to t}(x_i^t)$ and $\rho^{i\setminus j}_{t\leftarrow}(x_i^t)$ satisfy the set of equations Eqs. (35) and (36), with a 3×3 transfer matrix matrix $M^{i\setminus j}_{x_i^t x_i^{t+1}}$ given by

$$M^{i\setminus j}_{x_i^t x_i^{t+1}} = \begin{pmatrix} \alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}} p(O_i^t|S) & \left[1-\alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}}\right] p(O_i^t|S) & 0 \\ 0 & (1-r_i^t) e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p(O_i^t|I) & r_i^t e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p(O_i^t|I) \\ 0 & 0 & p(O_i^t|R) \end{pmatrix}. \tag{39}$$

The SIRS model differs from the SIR model only for the 3×3 transfer matrix, which is given by

$$M^{i\setminus j}_{x_i^t x_i^{t+1}} = \begin{pmatrix} \alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}} p(O_i^t|S) & \left[1-\alpha_i^t e^{\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}}\right] p(O_i^t|S) & 0 \\ 0 & (1-r_i^t) e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p(O_i^t|I) & r_i^t e^{\sum_{k\in\partial i\setminus j} \nu^t_{ik}\mu^t_{k\setminus i}} p(O_i^t|I) \\ \sigma_i^t p(O_i^t|R) & 0 & (1-\sigma_i^t) p(O_i^t|R) \end{pmatrix}. \tag{40}$$

As long as further compartments are included with transitions being parametrized by individual-based rates, generalization of the above construction follows straightforwardly (e.g., SEIR and SEIRS models).

V. RESULTS

In this section, we provide numerical results to highlight the operation and capabilities of this method. We first analyze the quality of the approximation for time-forward dynamics obtained in the absence of observations; then we effectively demonstrate the role of the cavity fields $\mu^t_{i\setminus j}$ in the presence of observations. The performances of the SCDC method in various instances of epidemic inference for the SI model are presented, both on synthetic and real-world contact networks. Finally, we evaluate the inference performance of the SCDC algorithm on recurrent epidemic models, such as the SIRS, on synthetic contact networks.

A. Time-forward dynamics

The emerging causality-breaking of the SCDC equations is a consequence of the existence of observations at later times, that have to be taken into account in the mathematical model by a flux of information flowing backward in time and conditioning the whole history of the process. This property reflects in the existence of nontrivial values for the one-time cavity fields $\mu^t_{i\setminus j}$. However, when no observation is present it is possible to show that all the cavity fields $\mu^t_{i\setminus j}$ vanish and, consequently, one can recover the usual causal time-forward mean-field dynamics. To prove this, it is convenient to start from a particular form of the update equations for the cavity marginals $m^t_{i\setminus j}$ (see Appendix C for a derivation),

$$m^t_{i\setminus j} = m^{t-1}_{i\setminus j} + \left(1-m^{t-1}_{i\setminus j}\right)\frac{\left(1-g^{t-1}_{i\setminus j}\right)\rho^{i\setminus j}_{t\leftarrow}(I)}{g^{t-1}_{i\setminus j}\,\rho^{i\setminus j}_{t\leftarrow}(S) + \left(1-g^{t-1}_{i\setminus j}\right)\rho^{i\setminus j}_{t\leftarrow}(I)}, \tag{41}$$

where, for brevity of notation, we have defined the term $g^t_{i\setminus j} = \alpha_i^t \exp\left(\sum_{k\in\partial i\setminus j} m^t_{k\setminus i}\nu^t_{ki}\right)$. The messages $\rho^{i\setminus j}_{t\leftarrow}(x_i^t)$ represent the (nonnormalized) backward probability that node i has state $x_i^t \in \{S, I\}$ at time t, given the dynamic constraints and the observations in the future. In the absence of observations (on all nodes at all times $t' \geqslant t$) the backward probability is uniform, i.e., $\rho^{i\setminus j}_{t\leftarrow}(S) = \rho^{i\setminus j}_{t\leftarrow}(I)$, and Eq. (41) reduces to time-forward mean-field equations,

$$m^t_{i\setminus j} = m^{t-1}_{i\setminus j} + \left(1-m^{t-1}_{i\setminus j}\right)\left[1-\alpha_i^{t-1} e^{\sum_{k\in\partial i\setminus j} m^{t-1}_{k\setminus i}\nu^{t-1}_{ki}}\right], \tag{42}$$

and for the total marginals m_i^t

$$m_i^t = m_i^{t-1} + \left(1-m_i^{t-1}\right)\left[1-\alpha_i^{t-1} e^{\sum_{j\in\partial i} m^{t-1}_{j\setminus i}\nu^{t-1}_{ji}}\right]. \tag{43}$$

It is possible to verify numerically that, in the absence of observations, the SCDC algorithm in Eqs. (24) and (25) always converges to the same result obtained by running the time-forward Eqs. (42).

An intuitive form for the discrete-time IBMF dynamics, obtained by assuming independence of individual marginal probabilities of being infected, is given by the equation [26,27]

$$m_i^t = m_i^{t-1} + \left(1-m_i^{t-1}\right)\left[1-\alpha_i^t \prod_{j\in\partial i}\left(1-\lambda_{ji}^{t-1} m_j^{t-1}\right)\right]. \tag{44}$$

A natural extension of Eq. (44) to account for cavity dynamics can be expressed as

$$m_i^t = m_i^{t-1} + \left(1-m_i^{t-1}\right)\left[1-\alpha_i^t \prod_{j\in\partial i}\left(1-\lambda_{ji}^{t-1} m^{t-1}_{j\setminus i}\right)\right], \tag{45}$$

where $m^t_{j\setminus i}$ represents the cavity marginal probability of node j being infected, excluding the influence of node i.

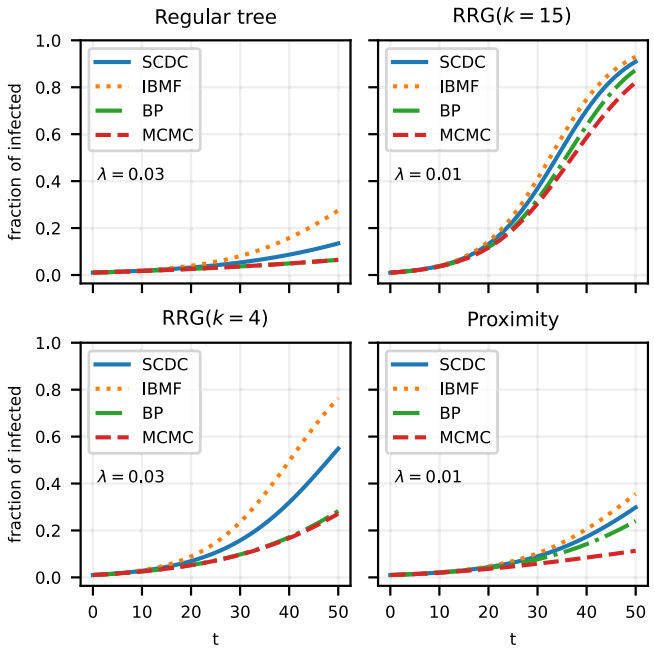


FIG. 1. Forward dynamics with SCDC and other mean-field methods. In each panel, the fraction of infected individuals is shown against the running time of the epidemic, with four different static contact graphs. Comparison is shown between SCDC, IBMF, BP, and Monte Carlo simulations (with 10^4 samples). All the links have the same infection probability λ, whose value is reported inside each panel. From left to right and top to bottom: regular tree with degree $K = 4$ and $N = 485$, RRG with $N = 500$ and degree $K = 4$, RRG with $N = 500$ and $K = 15$, proximity graph with $N = 500$. In all cases, the probability of each individual being infected at time $t = 0$ is set to $\gamma^0 = 5/N$, and the self-infection ε_i^t is set to 0.

This cavity probability evolves over time according to the following equation:

$$m_{i\setminus j}^t = m_{i\setminus j}^{t-1} + \left(1 - m_{i\setminus j}^{t-1}\right)\left[1 - \alpha_i^t \prod_{k\in\partial i\setminus j}\left(1 - \lambda_{ki}^{t-1} m_{k\setminus i}^{t-1}\right)\right]. \tag{46}$$

This cavity formulation provides a more refined perspective by iteratively excluding the feedback loops caused by direct interactions between neighboring nodes, improving the approximation of marginal probabilities in systems with significant correlations. For a densely connected graph (for which $m_{j\setminus i}^t \approx m_j^t$), Eqs. (45) and (44) reduce to the same expression.

Figure 1 illustrates the quality of the approximation obtained using Eqs. (42) and (43) for studying purely time-forward SI dynamics, in the absence of observations. The simulations were performed on well-known classes of static graphs, including regular trees (i.e., trees with degree K), random regular graphs (RRGs, i.e., graphs in which every node has exactly degree K), and proximity random graphs (details provided in Sec. V C). Comparisons are shown between the SCDC, the BP algorithm, and individual-based mean-field equations [IBMF, corresponding to Eqs. (44)]. As a reference, the results obtained with numerical sampling from 10^4 realizations of the exact time-forward Monte Carlo dynamics of the SI model are also reported. All methods based on mean-field approximations tend to overestimate the number of infected individuals in cases where the assumed factorization of probabilities is not exact. The BP algorithm is exact on trees (Bethe Lattice) and very accurate on sparse random graphs, where both SCDC and IBMF instead considerably overestimate the number of infected individuals. The performances of all methods are good on dense random graphs and much worse on graphs with spatial structure, such as proximity random graphs. In all cases under study, the SCDC approximation consistently gives better results than those obtained using IBMF.

We show in Appendix C how the value of λ influences the performance of these methods. The same types of plots as in Fig. 1 are repeated for six different values of λ, ranging from 0.05 to 0.5. In this analysis, we also included the improved cavity version of IBMF, given by Eq. (45), which we refer to as SCDCa. The results show that for small values of λ, SCDC performs better than IBMF in estimating the average fraction of infected individuals on all kind of graphs. Conversely, for large values of λ, IBMF performs slightly better than SCDC on random regular graphs (RRGs) and proximity graphs. However, the approximated version of SCDC, i.e., SCDCa, consistently outperforms both SCDC and IBMF across all cases. Despite its improved accuracy, we chose not to use this approximation for inference tasks, as it lacks sufficient theoretical justification.

B. Effects of backward messages

When at least one observation is included as evidence, the μ cavity fields are nonzero and cause a time-backward propagation of information which changes the probabilistic weight of the epidemic trajectories. To better understand how the method behaves in the presence of observations, we checked on which edges the absolute value of the μ cavity messages is nonvanishing when one or more observations are considered. In particular, we sampled a single epidemic outbreak from a uniform SI model with $\lambda = 0.19$, on a contact graph built by randomly adding some edges between nodes of a tree. Figure 2(b) shows how the fields $\mu_{i\setminus j}^t$ propagate into the contact graph up to three times before the observation. For each edge (i, j) a thick line is plotted if one of the two messages $\mu_{i\setminus j}^t$ or $\mu_{j\setminus i}^t$ is nonvanishing. A thicker line is plotted if both of them are nonvanishing. The plots are shown both for a single observation (top) and two observations (bottom). The plot on the right of Fig. 2(a) shows the inference accuracy of the method for the two cases. We can see that adding an observation greatly increases the inference performance. In particular, the prediction is improved mostly on the branches of the contact tree where the new observation produces the propagation of the cavity fields μ. It is clear that observations lead to the activation of the μ fields, which then propagate back in time away from the observed nodes.

A more intuitive probabilistic interpretation of the role of μ cavity fields in propagating the information obtained from observations can be obtained by monitoring the temporal behavior of the (normalized) time-forward and time-backward

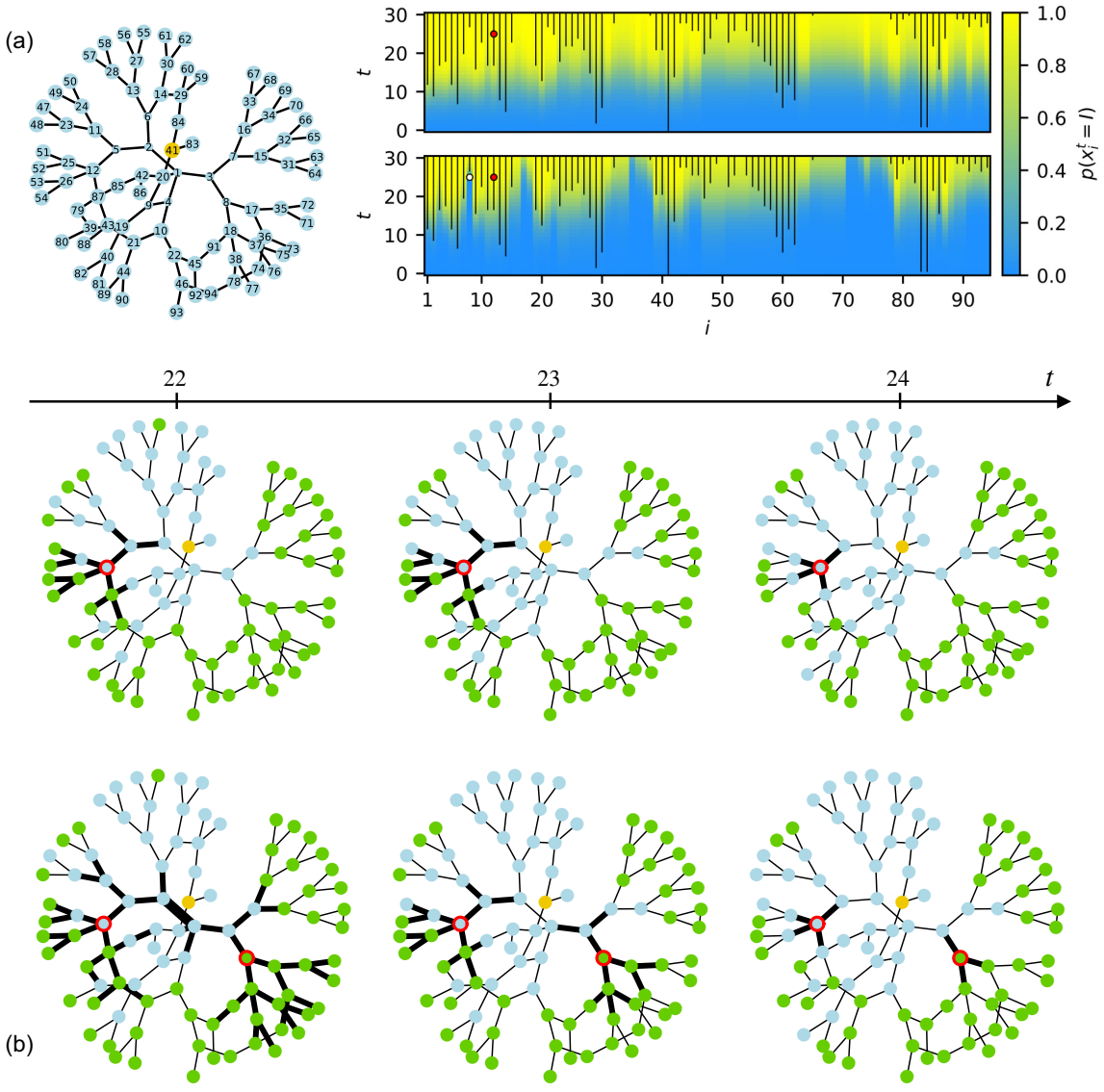


FIG. 2. Behavior of cavity fields $\mu^t_{i\setminus j}$ in the presence of observations. Panel [(a), left]: contact graph obtained randomly adding links to a tree; the epidemic source is shown as a yellow node. Panel [(a), right]: inference performance of SCDC with one observation (top) and two observations (bottom). The two plots show the posterior probability of being infected as a function of time (varying on the y axis) for all nodes of the contact graph [(a), left] in a single epidemic outbreak generated according to a uniform SI model with $\lambda = 0.19$. The same parameter was used to perform the inference. Specifically, the top plot corresponds to one observation on node 12 at time 25, while the bottom one corresponds to two observations on nodes 12 and 8 at time 25. The corresponding state of the observations is represented by colored dots (white $= S$, red $= I$). Black vertical lines mark the true infection periods. (b) Backpropagation through time of the μ cavity fields due to observations. The observed individuals are marked by a red circle. Infected individuals are marked by a blue dot, while susceptible individuals are green. The first line of plots corresponds to one observation, while the second line corresponds to two observations. Time flows from left to right, as shown in the time arrow on the top. Thin edges correspond to vanishing μ fields, thick edges corresponds to one of the two cavity fields $\mu^t_{i\setminus j}$ or $\mu^t_{i\setminus j}$ being nonvanishing, and the thickest edges correspond to both of them being nonvanishing. Observations lead to the activation of μ fields, which propagate back in time away from the observed nodes. The portions of the graph where the μ messages are nonvanishing have better predictions, meaning that the μ fields are fundamental for Bayesian inference.

messages $\tilde{\rho}^{i\setminus j}_{\to t}(x^t_i) = \rho^{i\setminus j}_{\to t}(x^t_i)/Z^{i\setminus j}_{\to t}$ and $\tilde{\rho}^{i\setminus j}_{t\leftarrow} = \rho^{i\setminus j}_{t\leftarrow}(x^t_i)/Z^{i\setminus j}_{t\leftarrow}$, where the normalization factor are respectively defined as

$$Z^{i\setminus j}_{\to t} = \sum_{x^{t-1}_i, x^t_i \in \{S,I\}} \rho^{i\setminus j}_{\to t-1}\left(x^{t-1}_i\right) M^{i\setminus j}_{x^{t-1}_i x^t_i}, \tag{47}$$

$$Z^{i\setminus j}_{t\leftarrow} = \sum_{x^t_i, x^{t+1}_i \in \{S,I\}} \rho^{i\setminus j}_{t+1\leftarrow}\left(x^{t+1}_i\right) M^{i\setminus j}_{x^t_i x^{t+1}_i}. \tag{48}$$

Consider a realization of the SI model taking place on a small tree, as displayed in Fig. 3 (left), in which the root node gets infected at time $t = 10$. An observation of the state of the root node at a time t_o introduces a source of information that affects the temporal behavior of the messages $\rho^{i\setminus j}_{\to t}$ and $\rho^{i\setminus j}_{t\leftarrow}$ for all directed edges (i, j) at all times. In particular, Fig. 3 (top row) shows the normalized time-backward messages $\tilde{\rho}^{i\setminus j}_{t\leftarrow}$ as function of time on a set of edges for $t_o = 8$ (center) and

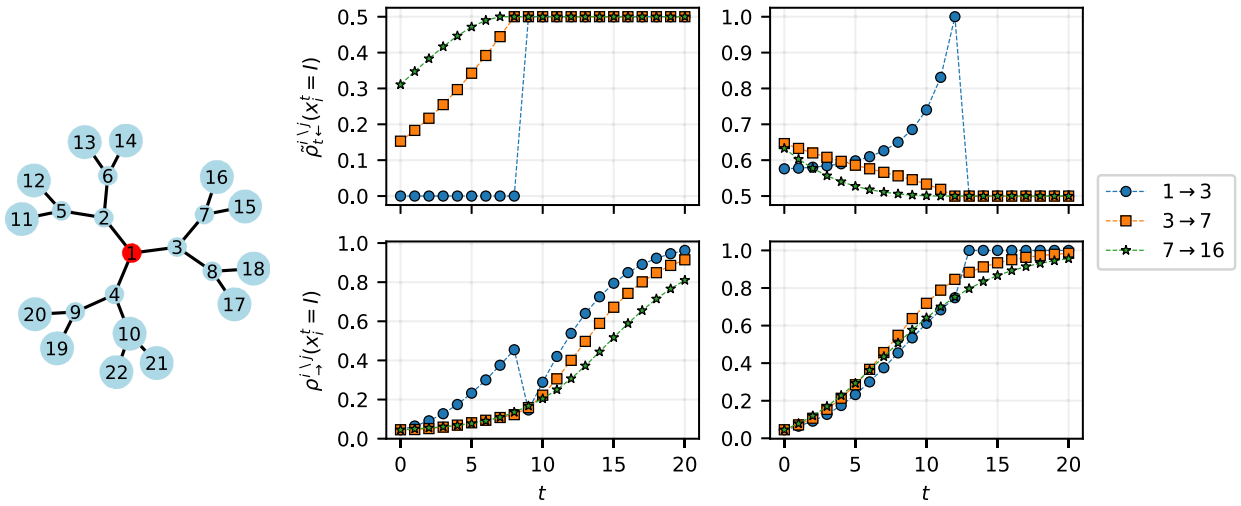


FIG. 3. Time propagation of information generated by observations. Normalized backward (top column) and forward (bottom column) cavity messages $\tilde{\rho}^{i\backslash j}_{t\leftarrow}(x^t_i = I)$ and $\tilde{\rho}^{i\backslash j}_{\rightarrow t}(x^t_i = I)$ plotted over time for different edges of a small tree. The epidemic outbreak is sampled using a uniform SI model with $\lambda = 0.1$. The plot on the left shows the graph. The observed individual is marked red, while the unobserved are blue. The observed node becomes infected at time $t = 10$. The plots on the left column show the messages when observing the central node being susceptible at time $t_o = 8$. The plots on the right column show the messages when observing the central node being infected at time $t_o = 12$.

$t_o = 12$ (right). The observation of a susceptible node (center) implies that at all times before the observation, the message emerging from that node is exactly zero. Moving away from the observed node the time-backward probability is nonzero (and monotonically increasing with the spatial distance) but monotonically decreasing with time distance from the observation. If instead the root node is observed in the infected state, then the time-backward message shows an instantaneous jump to 1 at the time of observation and gradually decreases at earlier times, as the time-backward probability of being infected for the root node decreases. Moving away from the root, the messages increase as time proceeds backward indicating that surrounding nodes might have caused the infection of the observed one. Clearly, the (normalized) time-backward messages are exactly equal to 0.5 when no information is available, i.e., at later time steps compared to the observation. The normalized time-forward messages $\rho^{i\backslash j}_{\rightarrow t}$, which are shown in Fig. 3 (bottom row) exhibit a similar, though more intuitive, phenomenology. They increase from 0 to 1 with increasing time, to explain the epidemic cascade generated by the infectious individuals at time $t = 0$. The message exiting from the central node collapses to a lower probability as soon as the individual is observed susceptible, after which it keeps increasing due to the forward spread of the dynamics. When the central node is observed infected though, the forward message exiting from it jumps to 1 only after some time-steps.

C. Inference performance for SI model

We consider a typical risk assessment scenario of epidemic inference in which, given the network of contacts and some observations made on an epidemic realization with one initially infected individual, one has to find the probability of each individual being infected at the final time. The simulations of the epidemic realizations according to the SI model are performed using the EpiGen python package [33] on both synthetic and real-world contact networks. The observations are performed on a random subset of the population at the final time. Once the individual probability of being infected at the final time is estimated with an epidemic inference method, the knowledge of the ground truth provided by the corresponding epidemic realization allows to compute a ROC curve of true infected individuals vs. false infected individuals. The area under the ROC curve (AUC) represents an estimate of the probability of correct classification of the individual infection states. The inference through SCDC is carried out using the efficient formulation through transfer matrices, and its performance is evaluated in comparison with two well-established methods for distributed epidemic inference, the simple mean-field (SMF) method [2] and the BP algorithm [2,14]. For both BP and SCDC methods, the individual probability of being infected at the final time is computed from the corresponding total marginals once the message passing algorithm has reached convergence, i.e., the error on the cavity marginals decreases under a predefined tolerance threshold. In some cases, BP and SCDC do not reach convergence in a reasonable number of iterations (a few thousand in the case of the epidemic instances under study). The lack of convergence can be due to a relevant role played by loop structures and long-range correlations. In such cases, the probability marginals are computed taking an average over a sufficiently large number (up to hundreds) of iterations of the message passing update. We analyzed in greater detail how often the two algorithms converge and how convergence affects inference performance, as will be shown later in this paper. The simple mean-field (SMF) inference method introduced in Ref. [2] is instead an inference method based on the IBMF approximation for the SI dynamics in which the information provided by observations of susceptible and infected individuals is taken into account by introducing some specific constraints on the time-forward dynamics (see Ref. [2] for a description in the more general SIR model).

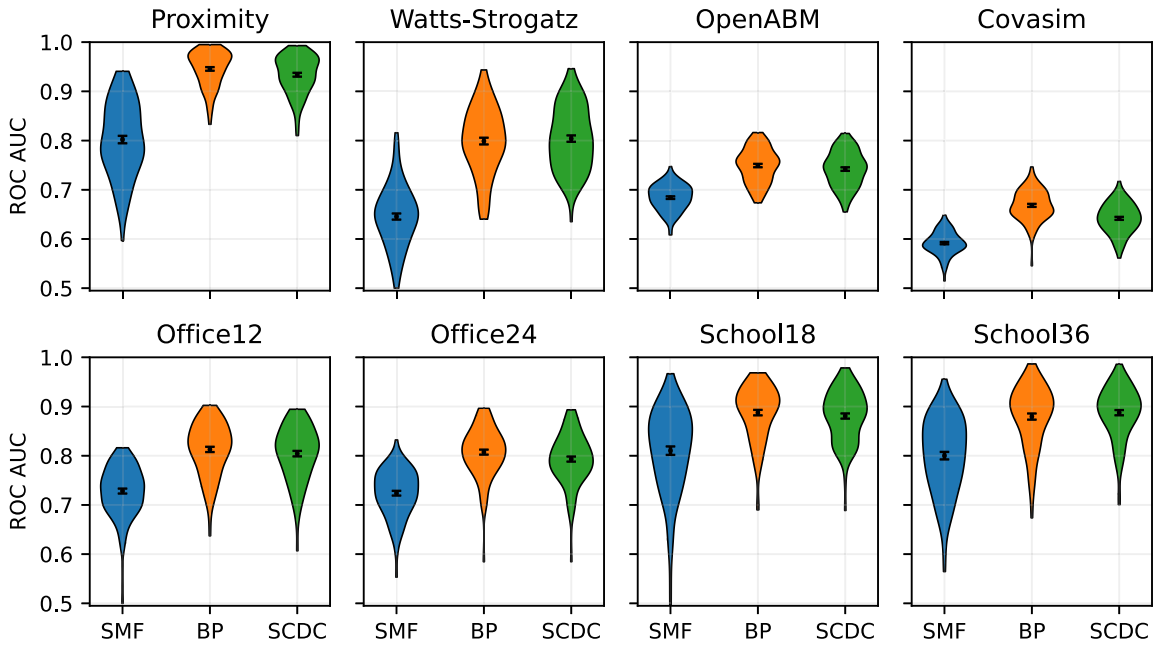


FIG. 4. Epidemic risk assessment of inference methods on SI epidemics. Epidemic risk assessment on different classes of random contact networks. Random contact networks are shown in the top row; from left to right: soft random geometric graphs (Proximity), Watts-Strogatz random networks (Watts-Strogatz), OpenABM contact network (OpenABM), and Covasim contact network (Covasim). Real-world contact networks are shown in the bottom row. Different sizes t_w of time windows were used to aggregate contact data. From left to right: office network from the InVS15 dataset with $T = 12, t_w = 24h$ (Office12), and with $T = 24, t_w = 12h$ (Office24), school network from the Thiers13 dataset with $T = 18, t_w = 6h$ (School18) and with $T = 36, t_w = 3h$ (School36). The performances of SMF, BP, and SCDC are compared in the problem of classifying the infection state of $N - n_{\rm obs}$ unobserved individuals at the last time T of SI epidemic processes when a small number $n_{\rm obs}$ of random observations are provided at the same time T of the epidemic process. The performance is quantified by computing the area under the curve (AUC) for the Receiver operating characteristic (ROC) curves for 100 epidemic realizations for each class of contact network. The violin plots illustrate the distribution of performances. Each plot is overlaid with a black marker indicating the mean value and the associated error of the AUC across epidemic realizations. The width of the violins represents the relative frequency of data points at each value. Other parameters: soft random geometric graphs ($N = 600$, $l_{\rm max} = \sqrt{2.8/N}$, $n_{\rm obs} = 20$, $T = 28$, $\lambda = 0.08$), Watts-Strogatz graphs ($N = 600$, average degree $z = 4$, $p_{rw} = 0.12$, $n_{\rm obs} = 20$, $T = 25$, $\lambda = 0.16$), OpenABM ($N = 2000$, $n_{\rm obs} = 60$, $T = 21$, $\gamma = 0.026$), Covasim ($N = 1000$, $n_{\rm obs} = 40$, $T = 24$, $\gamma = 0.038$), office networks ($N = 219$, $n_{\rm obs} = 15$, $\gamma = 6 \times 10^{-4}$), school networks ($N = 328$, $n_{\rm obs} = 20$, $\gamma = 7 \times 10^{-5}$).

Figure 4 presents the results for various types of graphs. The top row highlights outcomes for synthetic random graphs. The first two panels (from left to right) correspond to two categories of static random networks: Watts-Strogatz graphs [34] and soft random geometric graphs [35]. In the Watts-Strogatz model, the edges of a pristine network with a regular locally connected structure are rewired randomly with a probability $p_{\rm rw}$, leading to the emergence of nontrivial small-world and clustering properties. In soft random geometric graphs, also known as proximity random graphs, individuals are distributed uniformly at random in the unit square, then only pairs at Euclidean distance $l < l_{\rm max}$ are connected with a probability which decays exponentially with l. Both classes of random networks are locally highly structured, with short loops and clusters. The final two panels in the first row refer to synthetic contact networks generated with more realistic agent-based models, the OpenABM-Covid19 [36] and Covasim [37] models. These agent-based models are able to generate realistic contact networks on large populations, by modeling the interactions in households, schools, workplaces and other locations. Some contacts in these networks also change daily to reflect the dynamic nature of real-life interactions. Only the contact network structures generated by these agent-based models over a time horizon of a few weeks is used in the present work, and the epidemic propagations are generated using the standard SI model. In these networks, the link between two individuals i and j is assigned a weight w_{ij}^t, representing the aggregate duration of the contact between i and j in day t. Given an infection rate γ per contact per time unit, the infection probability associated to the contact is then computed as $\lambda_{ij}^t = 1 - e^{-\gamma w_{ij}^t}$. In both cases, when a relatively small number of observations is provided at the last time, the SCDC method is able to outperform SMF and achieves accuracy on par with the BP method.

The same testing framework is also employed to evaluate epidemic inference on two real contact networks, originally presented in Ref. [38], that have been collected with RFID tags in a school (Thiers13 dataset) and in an office environment (InVS15 datasets). The contact data are collected over a period of several days, with a temporal resolution of 20 s, which allows for data aggregation over coarse-grained time windows of a preferred size τ_w. In our study, time windows with size τ_w ranging from 3 h to a day are considered, for a total of T time steps ranging from a minimum of 12 to a

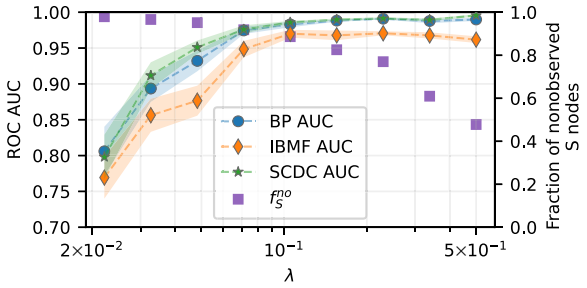


FIG. 5. Effect of λ. AUC comparison of SCDC, BP, and SMF on soft random geometric graphs as a function of the homogeneous infection probability λ. Results are averaged over 100 instances of random networks and epidemic outbreaks with $N = 300$ nodes, $l_{\max} = 1.8$, $T = 15$ days, and 30% of nodes observed at the final time. The ribbon width represents the standard error, showing the variability of the results. The figure also shows the fraction of nonobserved nodes that remain susceptible at the last time, namely f_S^{no}.

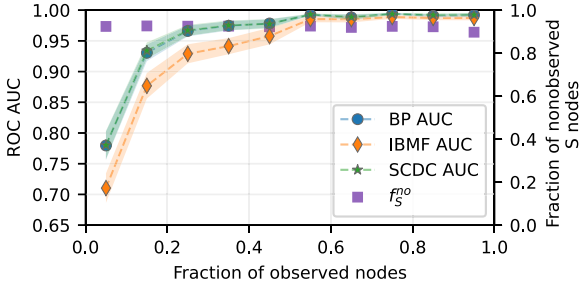


FIG. 6. Effect of the number of observations. AUC comparison of SCDC, BP, and SMF on soft random geometric graphs as a function of the fraction of observed nodes at the last time. Results are averaged over 100 instances of random networks and epidemic outbreaks with $N = 300$ nodes, $l_{\max} = 1.8$, $T = 15$ days, and $\lambda = 0.07$. The ribbon width represents the standard error, showing the variability of the results. The figure also shows the fraction of nonobserved nodes that remain susceptible at the last time, namely f_S^{no}.

maximum of 36 steps. When performing the coarse-graining procedure, the number c_{ij}^t of contacts between i and j occurring in a time window t of size τ_w is computed and used to estimate the infection probability λ_{ij}^t between the two individuals at time step t as $\lambda_{ij}^t = 1 - (1-\gamma)^{c_{ij}^t}$, where γ is a common parameter describing the infectiousness of a single contact. The results of epidemic risk assessment on these real-world contact networks are shown in the second row of Fig. 4, adopting the same metric used in the case of random graphs. Also in this case, for all contact networks under study, the SCDC method has a performance very close to the BP algorithm, and in general superior to the SMF heuristic.

1. Effect of the infection probability and number of observations

In addition to the networks discussed above, we further assess the performance of SCDC, BP, and IBMF on soft random geometric graphs as a function of two critical parameters: the homogeneous infection probability λ and the fraction of observed nodes at the final time.

Figure 5 illustrates the AUC achieved by the three methods as a function of λ, averaged over 100 instances of random contact networks and epidemic outbreaks. The soft random geometric graphs used in this analysis consist of $N = 300$ nodes, with a cutoff distance of $l_{\max} = \sqrt{1.8/N}$. Epidemic outbreaks are simulated over $T = 15$ days, with two randomly selected individuals initially infected at time $t = 0$. At the final time step, a randomly selected fraction of 30% of the nodes is observed. The infection probability λ is varied from 0.02 to 0.5. Surprisingly, the SCDC method shows inference performance comparable to that of BP, even for larger values of λ, despite being a small-coupling approximation to BP. In all cases, BP outperforms the IBMF method in assessing epidemic risk.

Figure 6 evaluates the effect of varying the fraction of observed nodes on the AUC. The setup is identical to that described for Fig. 5, with the infection probability fixed at $\lambda = 0.15$. The fraction of nodes observed at the final time is varied from 5% to 95%. As the fraction of observed nodes increases, the performance of the three methods becomes more comparable. For a smaller fraction of observed nodes, SCDC maintains performance similar to BP, while IBMF performs slightly worse. However, as the observation fraction grows, the differences between the methods diminish, and their performance becomes comparable for sufficiently large fractions of observed nodes.

In some cases, BP and SCDC do not reach convergence within a reasonable number of iterations (a few thousand in the case of the epidemic instances under study). The lack of convergence can be attributed to the relevant role played by loop structures and long-range correlations. In such cases, the probability marginals are computed by taking an average over a sufficiently large number (up to hundreds) of iterations of the message-passing updates. Notably, the instances on which the convergence of BP and SCDC is evaluated are the same as those used in Figs. 5 and 6. Tables I and II present the fraction of converged instances for both BP and SCDC under varying λ and fractions of observed nodes, respectively. While BP consistently shows better convergence rates than SCDC, the convergence rate decreases for both methods as λ increases. Despite this, the inference accuracy of both methods remains comparable, even when the convergence rate of SCDC is significantly lower than that of BP.

D. Inference in recurrent epidemic models

While previous results focus on the quantitative analysis of inference performances on irreversible dynamics, the present subsection aims at illustrating the potential of the SCDC method for epidemic inference on recurrent epidemic models. To do that, we perform a simple analysis inspired by the one already presented in recent work on matrix product belief propagation (MPBP) [23], a novel powerful approximation method for observation-reweighted recurrent dynamics

TABLE I. Convergence rates of BP and SCDC for varying values of λ. The inference setting is the same of Fig. 5.

λ	0.02	0.03	0.05	0.07	0.1	0.15	0.23	0.34	0.5
BP	1.0	0.98	1.0	0.68	0.48	0.2	0.12	0.06	0.0
SCDC	0.7	0.5	0.48	0.2	0.16	0.16	0.16	0.08	0.06

TABLE II. Convergence rates of BP and SCDC for varying fractions of observed nodes. The inference setting is the same of Fig. 6.

F. obs.	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95
BP	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
SCDC	0.8	0.4	0.22	0.28	0.26	0.22	0.38	0.44	0.4	0.5

on graphs. We conducted simulations of a single epidemic outbreak using a SIRS model on an Erdős–Rényi random graph [39] with $N = 100$ nodes and average degree $z = 3$. Figure 7 shows the value of the posterior marginal probability of being infected $p(x_i^t = 0|\mathcal{O})$ inferred by the SCDC method. The color scale indicating these probability values is superimposed on the black bars marking the time intervals of true infections, which enables visual inspection of the inference performance of the method. Notably, the SCDC method rather accurately assigns posterior marginal probabilities that closely align with the observed data, demonstrating its effectiveness even for unobserved nodes or time points that are distant from the observations. Several reinfection events are also correctly captured.

Finally, we performed a quantitative comparison of the inference performance between the SCDC method and MPBP on recurrent epidemic processes, specifically focusing on the SIRS model. To evaluate the reconstruction accuracy of the inference algorithms, we used the fraction of correct MAP predictions. Let $\mathbb{P}(x_i^t)$ denote the inferred probability of node i being in state $x_i^t \in \{S, I, R\}$ at time t, and let X_i^t represent the ground truth, i.e., the true state of node i at time t. The fraction of correct MAP predictions is defined as

$$f_{\text{correct}}^{\text{MAP}} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \delta\left(\arg\max_{x_i^t \in \{S,I,R\}} \mathbb{P}\left(x_i^t\right), X_i^t \right), \tag{49}$$

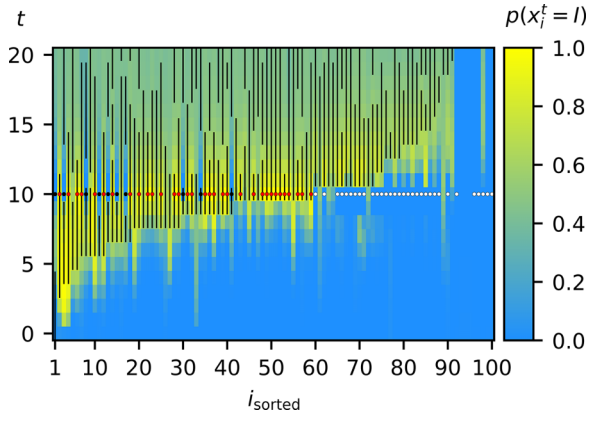


FIG. 7. Inference in recurrent epidemic models. Posterior probability of being infected as a function of time for all nodes of an Erdős–Rényi random graph ($N = 100$ nodes and average degree $z = 3$) in a single epidemic outbreak generated according to a uniform SIRS model. The parameters of the SIRS model are $\lambda = 0.4$, $r = \sigma = 0.15$, for all nodes at all times. The same parameters were used to perform the inference. The 75% of the nodes were randomly observed at the same time 10, and their corresponding state is represented by colored dots (white=S, red=I, black=R). Black vertical lines mark to true infection periods. Node order from left to right reproduces the order of true infection events.

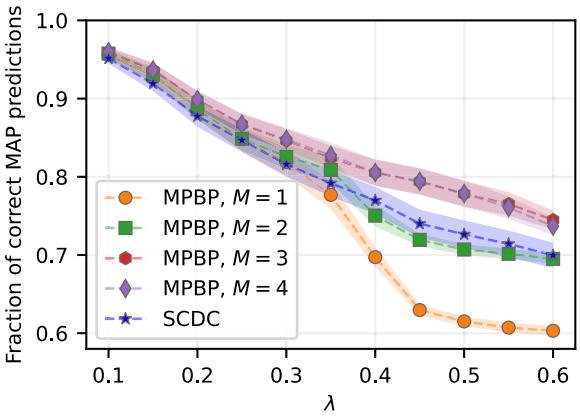


FIG. 8. Epidemic risk assessment on recurrent models. Comparison of the inference performance of the SCDC algorithm and the MPBP algorithm for different bond dimensions $M = 1, 2, 3, 4$ on Erdős–Rényi random graphs with $N = 100$ nodes and average degree $z = 2.5$. The epidemic spread follows an SIRS model with uniform infection probability λ, recovery probability $r_i^t = 0.15$, and immunity-loss probability $\sigma_i^t = 0.15$, for all nodes and times. The duration of the epidemic is $T = 20$ steps, with an initial condition of 2 randomly chosen infected individuals at $t = 0$. At $t = 10$, 75% of the nodes were randomly selected for observation. The averaged fraction of correct maximum-*a-posteriori* (MAP) predictions, computed over 50 instances of random graphs and epidemic outbreaks, is plotted as a function of λ. The ribbon width represents the standard error, showing the variability of the results.

where $\delta(x, y)$ denotes the Kronecker symbol. Figure 8 compares the inference performance of the SCDC algorithm with the MPBP algorithm, evaluated at different bond dimension parameters, $M = 1, 2, 3, 4$. The bond dimension is a parameter of the MPBP method that determines the size of the matrices used in the matrix product approximation of the cavity messages. The comparison is based on SIRS epidemic models simulated on Erdős–Rényi random graphs with $N = 100$ nodes and an average degree $z = 2.5$. The recovery and immunity-loss probabilities are uniform, with $r_i^t = \sigma_i^t = 0.15$ for all nodes and time steps, while the infection probability $\lambda_i^t = \lambda$ is constant across all nodes and time steps within each simulation. For each fixed value of λ, the fraction of correct MAP predictions was averaged over 50 independent realizations of random graphs and epidemic outbreaks. At time step $t = 10$, a randomly chosen fraction of 75% of the nodes was observed, providing partial data for the inference algorithms. The plot shows the average fraction of correct MAP predictions for different values of λ.

The results show that, for small infection probabilities λ, all algorithms perform equivalently achieving high accuracy. In this regime, the slower spread of the epidemic allows for easier inference. However, as λ increases, the epidemic cascades become faster and more challenging to reconstruct, leading to a decline in inference performance. Notably, the SCDC algorithm performance is comparable with that of the MPBP algorithm with bond dimensions $M = 2$, even in this challenging inference regime. Additionally, the MPBP algorithm does not show any improvement in inference performance for $M > 3$, as the results for $M = 4$ are identical to those for

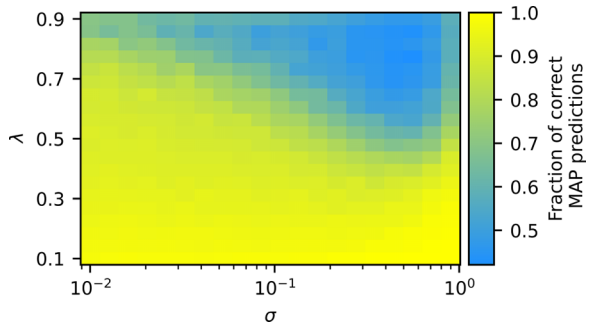


FIG. 9. Effect of λ and σ on epidemic reconstruction. Fraction of correct MAP predictions as a function of the homogeneous infection probability λ and the homogeneous recovery probability σ for a SIS epidemic model on soft random geometric graphs with $N = 300$ nodes. The cutoff distance is set to $l_{\max} = \sqrt{1.8/N}$. The SIS epidemic is simulated over $T = 15$ days, with two randomly chosen nodes infected at time $t = 0$. At time $t = 7$, a randomly selected fraction of 60% of the nodes is observed. The fraction of correct MAP predictions is averaged over 30 independent instances of random graphs and epidemic outbreaks. The color scale represents the fraction of correct MAP predictions, with higher values indicating better reconstruction accuracy.

$M = 3$. This suggests that $M = 3$ represents the maximum inference performance achievable by the MPBP algorithm in this setting.

An important advantage of the SCDC algorithm lies in its computational efficiency: while the computational complexity of MPBP scales as $O(S^6|E|TM^6)$, the complexity of SCDC is only $O(S^2|E|T)$, where S is the number of states in the compartmental epidemic model under study (e.g., $S = 3$ for the SIRS model). Moreover, although MPBP, especially with sufficiently high bond dimensions M, achieves greater accuracy in reconstructing the epidemic spread, the SCDC algorithm provides a simple and interpretable approximation. The terms involved in the SCDC approach, such as m and μ, have a clear and interpretable physical meaning, making the method more accessible for further theoretical development.

As a different analysis, we now consider a SIS epidemic model instead of the previously used SIRS model. In Fig. 9, we evaluate the fraction of correct MAP predictions as a function of both the homogeneous infection probability λ and the homogeneous recovery probability σ. The inference is performed on soft random geometric graphs with $N = 300$ nodes, where the cutoff distance is set to $l_{\max} = \sqrt{1.8/N}$. The SIS epidemics are simulated over $T = 15$ days, with two randomly infected nodes at $t = 0$. At time $t = 7$, a fraction of 60% of the nodes is observed. The fraction of correct MAP predictions was averaged over 30 independent instances of random graphs and epidemic outbreaks. The infection probability is varied from 0.1 to 0.9, while the recovery probability is varied from 0.01 to 0.9. This analysis provides a comparison of the SCDC method's performance under different epidemic dynamics, focusing on how the inference accuracy depends on the infection and recovery probabilities. The results show that the epidemic reconstruction using the SCDC method is generally accurate, except for a small region in the phase space with large values of both λ and σ. This region corresponds to epidemic outbreaks with frequent reinfections, where inference is more challenging due to the complex dynamics of the epidemic spread.

VI. DISCUSSION

The dynamic cavity method is a distributed technique to study discrete-state stochastic processes on graphs, which is exact on trees and often provides very good approximations on sparse graphs. While its original formulation is computationally demanding [19,20], approximations have been introduced [40–45] and, whenever possible, more efficient parametrizations of single dynamical trajectories have been designed [13,14,22]. In the present work, an observation-reweighted version of the dynamic cavity formulation, including individual observations, is devised to model the posterior probability of epidemic processes on contact networks. The formulation exploits a Bayesian approach and is fully equivalent to the belief propagation approach to epidemic trajectories [13,14]. Starting from the reweighted dynamic cavity formulation and exploiting a small-coupling expansion, a novel set of fixed-point equations for a pair of time-dependent cavity messages $m^t_{i\setminus j}$ and $\mu^t_{i\setminus j}$ is obtained. Here, $m^t_{i\setminus j}$ is the approximate probability that individual i is infected at time t in the cavity graph when the interaction with individual j is removed, while $\mu^t_{i\setminus j}$ is a cavity field whose role depends on the presence of observations. In the absence of observations, all cavity fields $\{\mu^t_{i\setminus j}\}$ identically vanish, and the dynamics, expressed solely in terms of cavity probabilities $\{m^t_{i\setminus j}\}$, becomes causal, reducing to a set of generalized mean-field equations. These time-forward equations, tested on random graphs for the SI model, yield higher accuracy compared to the commonly used individual-based mean-field equations (which they reduce to in the regime of low infectiousness and high connectivity) for all the tested settings. They are albeit less accurate than BP, which in the simplified case of nonrecurrent forward dynamics coincides with the dynamic message passing method [21,22]. Simple analyses conducted on the SI model with limited observations demonstrate that the role of the cavity fields $\{\mu^t_{i\setminus j}\}$ is to propagate information about observations to neighboring nodes and subsequently distribute this information throughout the contact network, appropriately tilting the probabilistic weight of the associated dynamic trajectories in view of the presence of observations. The presence of observations renders the epidemic dynamics noncausal, with backward-in-time information flow, as evident from the nonuniform distribution of the normalized backward cavity messages $\tilde{\rho}^{i\setminus j}_{t\leftarrow}(x^t_i)$. The main additional approximation assumed in deriving the SCDC method from the DC equations is the independence of cavity messages from the epidemic trajectory of the removed node. This approximation can introduce inconsistencies with specific trajectories imposed by observations, particularly in regimes with numerous observations, including repeated observations on the same individuals. However, this issue is effectively resolved by introducing a small self-infection probability, which practically eliminated the problem in all applications considered. The SCDC algorithm shows promising effectiveness in assessing the epidemic risk of individuals, with performance comparable to that of BP, of which it is

essentially an approximation, and offering improvements over other heuristic methods based on mean-field approximations. As a fixed-point message passing method, a potential drawback of SCDC lies in its convergence properties. In numerical tests, SCDC experiences convergence problems similar to BP, mainly resulting from long-range correlations generated by loops in the contact graphs. Nevertheless, even in the absence of convergence, the estimated marginal probabilities often remain sufficiently accurate, enabling a reliable estimation of the epidemic risk.

The main advantage of the SCDC method over dynamic cavity and belief propagation for epidemic trajectories lies in its straightforward generalization to epidemic models with multiple states (e.g., SIR, SEIR) and recurrent processes (e.g., SIS, SIRS). Indeed, the fundamental components of the method and the efficient algorithm based on the temporal transfer matrix remain largely unchanged, with only modifications in the matrix dimensions and elements to accommodate the model's increased complexity. Consequently, the SCDC algorithm maintains a linear complexity with respect to the duration of the epidemic process and the number of contacts in the network. We compared SCDC with matrix product belief propagation (MPBP) [23], a novel approximation method for observation-reweighted recurrent dynamics on graphs, based on a matrix product approximation of the cavity messages. Both methods showed comparable inference performance on SIRS models at low infection probabilities. However, as the infection probability increases, the performance gap widens in favor of MPBP. Despite this, SCDC retains a key advantage in terms of computational efficiency and offers a simpler approximation with a more intuitive physical interpretation of the involved parameters. Perhaps surprisingly, this excellent predicting power persists in regimes in which the infectiousness parameter is relatively large, outside the range that was expected due to the assumptions in the derivation, as demonstrated for the inference on SIS models.

The primary limitation of the SCDC method is that its efficient formulation based on the transfer matrix is currently applicable only to Markovian models. Further study is required to develop an efficient algorithm for non-Markovian recurrent epidemic models. Concerning the method itself, another interesting direction for its development involves gaining a better understanding of the role of second-order terms in the small-coupling expansion and developing an improved algorithm that takes them into account. Finally, future directions include the possibility to generalize the approach presented here to other type of dynamical processes on networks, e.g., rumor spreading processes [20,46,47].

ACKNOWLEDGMENTS

Computational resources were provided by the SmartData@PoliTO [48] interdepartmental center on Big Data and Data Science. We thank S. Crotti for suggesting the format of Figs. 2 and 7, similar to what was used in Ref. [23]. G.C. acknowledges support by the Comunidad de Madrid and the Complutense University of Madrid (UCM) through the Atracción de Talento programs (Ref. No. 2019-T1/TIC-13298). F.M. acknowledges support from the project "CODE–Coupling Opinion Dynamics with Epidemics," funded under PNRR Mission 4 "Education and Research"—Component C2—Investment 1.1—Next Generation EU "Fund for National Research Program and Projects of Significant National Interest" PRIN 2022 PNRR, Grant Code P2022AKRZ9, CUP B53D23026080001. A.B. and L.D.A. acknowledge that this study was carried out within the FAIR—Future Artificial Intelligence Research—project and received funding from the European Union NextGenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR)–Missione 4 Componente 2, Investimento 1.3–D.D. Grants No. 1555 11/10/2022 and No. PE00000013). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

DATA AVAILABILITY

The data that support the findings of this article are openly available [32].

APPENDIX A: BELIEF PROPAGATION EQUATIONS FOR THE SI MODEL

The posterior probability in Eq. (6b) can be rewritten in the following factorized form:

$$p(\boldsymbol{X}|\boldsymbol{O}) \propto \prod_{t=0}^{T-1}\prod_{i=1}^{N} f_i^{t+1}\left(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t|\boldsymbol{\mathcal{O}}\right), \tag{A1}$$

where the factors f_i^{t+1} encode both the local transition dynamics and the influence of observations, and are defined as

$$f_i^{t+1}\left(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t|\boldsymbol{\mathcal{O}}\right) = \left(p_0(x_i^0)p(O_i^0|x_i^0)\right)^{\delta_{t,0}} W_i\left(x_i^{t+1}|\boldsymbol{x}^t\right)p\left(O_i^t|x_i^t\right). \tag{A2}$$

Here, $W_i(x_i^{t+1}|\boldsymbol{x}^t)$ represents the transition rate from x_i^t to x_i^{t+1}, which depends solely on the state of the neighbors at time t, denoted as $\boldsymbol{x}_{\partial i}^t$. Although the factor graph associated with this posterior distribution exhibits a substantial number of short loops, the structure can be disentangled by leveraging the interactions between trajectories of neighboring nodes. Specifically, pairs of trajectories $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ that are connected by an edge in the original contact graph can be grouped together. This procedure leads to a reformulation in terms of a new graphical representation, known as the *dual factor graph*.

In the dual factor graph, the variable nodes correspond to pairs of trajectories $(\boldsymbol{x}_i, \boldsymbol{x}_j)$, which are associated with the edges of the original contact graph. The factor nodes, however, correspond to the vertices i of the contact graph. A visual depiction of this transformation is shown in Fig. 10, where panel (b) illustrates the original loopy factor graph, and panel (c) demonstrates the disentangled dual factor graph. The factors in the dual representation are obtained by grouping the contributions over all time steps for each vertex, i.e., $\boldsymbol{f}_i(\boldsymbol{x}_i, \{\boldsymbol{x}_k\}_{k\in\partial i}|\boldsymbol{\mathcal{O}}) = \prod_{t=0}^{T-1} f_i^{t+1}(x_i^{t+1}, \boldsymbol{x}_{\partial i}, x_i^t|\boldsymbol{\mathcal{O}})$. A more comprehensive discussion of this step, including its theoretical implications and practical benefits, can be found in Refs. [2,13,14,22,49].

The BP equations on the dual factor graph are [23]

$$\hat{c}_{i\setminus j}(\boldsymbol{x}_i,\boldsymbol{x}_j|\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{x}_{\partial i\setminus j}}\prod_{t=0}^{T-1} f_i^{t+1}\left(x_i^{t+1},\boldsymbol{x}_{\partial i}^t,x_i^t\middle|\boldsymbol{\mathcal{O}}\right)\prod_{k\in\partial i\setminus j}\hat{c}_{k\setminus i}(\boldsymbol{x}_k,\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}) \tag{A3a}$$

$$\propto p_0\left(x_i^0\right)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\sum_{\boldsymbol{x}_{\partial i\setminus j}}\prod_{t=0}^{T-1} W_i\left(x_i^{t+1}\middle|\boldsymbol{x}^t\right)\prod_{k\in\partial i\setminus j}\hat{c}_{k\setminus i}(\boldsymbol{x}_k,\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}). \tag{A3b}$$

The explicit form of the transition rate is

$$W_i\left(x_i^{t+1}\middle|\boldsymbol{x}^t\right) = \delta_{x_i^{t+1},S}\delta_{x_i^t,S}\alpha_i^t e^{\sum_{k\in\partial i}\nu_{ki}^t\delta_{x_k^t,I}} + \delta_{x_i^{t+1},I}\left(1-\delta_{x_i^t,S}\alpha_i^t e^{\sum_{k\in\partial i}\nu_{ki}^t\delta_{x_k^t,I}}\right) \tag{A4a}$$

$$= \alpha_i^t e^{\sum_{k\in\partial i}\nu_{ki}^t\delta_{x_k^t,I}}\left(\delta_{x_i^{t+1},x_i^t}-\delta_{x_i^{t+1},I}\right)+\delta_{x_i^{t+1},I}. \tag{A4b}$$

The dynamic cavity equations Eq. (7) are therefore obtained from the BP equations with the simple change of variables $s_i^t = \nu_{ji}^t\delta_{x_j^t,I}$. The local fields s_i^t act as proxies for the missing variables x_j^t of the dual factor graph.

APPENDIX B: PATH INTEGRAL DERIVATION OF THE DYNAMIC CAVITY EQUATIONS FOR THE SI MODEL

In this Appendix, we present a detailed derivation of the dynamic cavity equations, Eqs. (7) and (12), which is equivalent to the one proposed in Appendix A. This derivation leverages a path-integral representation of the stochastic epidemic dynamics of the SI model. The approach is based on interpreting the (Markovian) update rules governing the discrete-time stochastic process as a set of dynamical constraints imposed on the degrees of freedom of the system, specifically the binary variables $\{x_i^t\}$. These variables describe the state of each node i at time t, with $x_i^t \in \{S, I\}$ indicating whether the node is susceptible or infected, respectively. The derivation is further grounded in the definition of a dynamic partition function, which encapsulates the stochastic dynamics of the system and incorporates the interplay between the microscopic states of the nodes and their interactions over time.

The dynamic partition function is defined as follows:

$$\mathcal{Z}(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{X}} p(\boldsymbol{X}|\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{X}}\prod_{i=1}^{N} p_0\left(x_i^0\right)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\prod_{t=0}^{T-1} W_i\left(x_i^{t+1}|\boldsymbol{x}^t\right). \tag{B1}$$

To proceed, we explicitly express $W_i(x_i^{t+1}|\boldsymbol{x}^t)$ in terms of the possible transitions:

$$W_i\left(x_i^{t+1}|\boldsymbol{x}^t\right) \propto \delta_{x_i^{t+1},S}W_i(S|\boldsymbol{x}^t) + \delta_{x_i^{t+1},I}W_i(I|\boldsymbol{x}^t), \tag{B2}$$

where $\delta_{x_i^{t+1},S}$ and $\delta_{x_i^{t+1},I}$ are Kronecker δ functions ensuring that only the appropriate transition probabilities contribute to the dynamics. Substituting this into the partition function, we have

$$\mathcal{Z}(\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{X}}\prod_{i=1}^{N} p_0\left(x_i^0\right)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\prod_{t=0}^{T-1}\left[\delta_{x_i^{t+1},S}W_i(S|\boldsymbol{x}^t) + \delta_{x_i^{t+1},I}W_i(I|\boldsymbol{x}^t)\right]. \tag{B3}$$

The transition probabilities $W_i(x_i^{t+1}|\boldsymbol{x}^t)$ depend on the interactions of node i with its neighbors, which are captured through a local external field, h_i^t. Specifically, the field is defined as

$$h_i^t = \sum_j \nu_{ji}^t\delta_{x_j^t,I}, \tag{B4}$$

where ν_{ji}^t represents the interaction strength between nodes j and i at time t. The dependence of h_i^t on the states of neighboring nodes is enforced by introducing a Dirac δ function $\delta(h_i^t - \sum_j \nu_{ji}^t\delta_{x_j^t,I})$. By substituting this definition of the local field into the dynamic partition function, we obtain

$$\mathcal{Z}(\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{X}}\prod_{i=1}^{N} p_0\left(x_i^0\right)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\prod_{t=0}^{T-1}\int_{-\infty}^{\infty} dh_i^t\left[\delta_{x_i^{t+1},S}\delta_{x_i^t,S}\alpha_i^t e^{h_i^t} + \delta_{x_i^{t+1},I}\left(1-\delta_{x_i^t,S}\alpha_i^t e^{h_i^t}\right)\right]\delta\left(h_i^t - \sum_j \nu_{ji}^t\delta_{x_j^t,I}\right), \tag{B5}$$

where α_i^t encodes the self-infection probability, as defined in Sec. II A.

To facilitate computation, we use the integral representation of the Dirac δ function:

$$\delta(h) = \int_{-\infty}^{\infty}\frac{d\hat{h}}{2\pi}e^{i\hat{h}h}, \tag{B6}$$

allowing us to introduce an auxiliary integration variable $\hat{h}_i^t$ for each h_i^t. Substituting this representation into the partition function and integrating out the h_i^t variables introduces new interaction terms in the exponential form. These terms reflect the contributions of the external fields from neighbors and form the basis for deriving the dynamic cavity equations.

$$\mathcal{Z}(\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{X}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t=0}^{T-1} \int_{-\infty}^{\infty} dh_i^t \big[\alpha_i^t e^{h_i^t} \delta_{x_i^t,S}\big(\delta_{x_i^{t+1},S} - \delta_{x_i^{t+1},I}\big) + \delta_{x_i^{t+1},I}\big] \int_{-\infty}^{\infty} \frac{d\hat{h}_i^t}{2\pi} e^{\mathrm{i}\hat{h}_i^t (h_i^t - \sum_j \nu_{ji}^t \delta_{x_j^t,I})} \tag{B7a}$$

$$\propto \sum_{\boldsymbol{X}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t=0}^{T-1} \int_{-\infty}^{\infty} d\hat{h}_i^t \big[\delta\big(\hat{h}_i^t - \mathrm{i}\big)\alpha_i^t e^{h_i^t} \delta_{x_i^t,S}\big(\delta_{x_i^{t+1},S} - \delta_{x_i^{t+1},I}\big) + \delta\big(\hat{h}_i^t\big)\delta_{x_i^{t+1},I}\big] e^{-\mathrm{i}\hat{h}_i^t \sum_j \nu_{ji}^t \delta_{x_j^t,I}} \tag{B7b}$$

$$\propto \sum_{\boldsymbol{X}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t=0}^{T-1} \int_{-\infty}^{\infty} d\hat{h}_i^t \big[\delta\big(\hat{h}_i^t - \mathrm{i}\big)\alpha_i^t e^{h_i^t} \big(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\big) + \delta\big(\hat{h}_i^t\big)\delta_{x_i^{t+1},I}\big] e^{-\mathrm{i}\hat{h}_i^t \sum_j \nu_{ji}^t \delta_{x_j^t,I}} \tag{B7c}$$

$$\propto \sum_{\boldsymbol{X}} \int D\hat{\boldsymbol{H}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t=0}^{T-1} \big[\delta\big(\hat{h}_i^t - \mathrm{i}\big)\alpha_i^t e^{h_i^t} \big(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\big) + \delta\big(\hat{h}_i^t\big)\delta_{x_i^{t+1},I}\big] e^{-\mathrm{i}\hat{h}_i^t \sum_j \nu_{ji}^t \delta_{x_j^t,I}} \tag{B7d}$$

$$\propto \sum_{\boldsymbol{X}} \int D\hat{\boldsymbol{H}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t=0}^{T-1} \big[\delta\big(\hat{h}_i^t - \mathrm{i}\big)\alpha_i^t e^{h_i^t} \big(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\big) + \delta\big(\hat{h}_i^t\big)\delta_{x_i^{t+1},I}\big] \prod_{j>i} e^{-\mathrm{i}(\hat{h}_i^t \nu_{ji}^t \delta_{x_j^t,I} + \hat{h}_j^t \nu_{ij}^t \delta_{x_i^t,I})}, \tag{B7e}$$

where $\int D\hat{\boldsymbol{H}} = \prod_{i=1}^{N} \prod_{t=0}^{T-1} (\int_{-\infty}^{+\infty} d\hat{h}_i^t)$ for shortness of notation. We can simplify the notation by introducing the local noninteracting and interacting actions Eqs. (15) and (16). The dynamical partition function simplifies to

$$\mathcal{Z}(\boldsymbol{\mathcal{O}}) \propto \sum_{\boldsymbol{X}} \int D\hat{\boldsymbol{H}} \prod_{i=1}^{N} p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^0} \prod_{j>i} e^{S_{ij}^{\mathrm{int}}}, \tag{B8}$$

where S_i^0 depends on the local variable-conjugate field trajectory $(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)$ and S_{ij}^{int} depends on the two local variable-conjugate field trajectories $(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)$, $(\boldsymbol{x}_j, \hat{\boldsymbol{h}}_j)$. The probabilistic weight associated with the dynamic partition function $\mathcal{Z}(\boldsymbol{\mathcal{O}})$ is now in a form that can be represented as a graphical model, in which the variable nodes correspond to the spatiotemporal variables $\boldsymbol{x}_i$ and $\hat{\boldsymbol{h}}_i$ and there are two types of factor nodes (see Fig. 10): single-node factors,

$$\phi_i((\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i)) = p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^0}, \tag{B9}$$

and factors involving pairs of variables on neighboring nodes at the same time,

$$\psi_{ij}((\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i), (\boldsymbol{x}_j, \hat{\boldsymbol{h}}_j)) = e^{S_{ij}^{\mathrm{int}}}. \tag{B10}$$

By grouping together single-node variables at all times, i.e., trajectories $(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i) = ((x_i^0, \ldots, x_i^T), (\hat{h}_i^0, \ldots, \hat{h}_i^T))$, the resulting factor graph reproduces the topology of the underlying interaction graph. It should be noted that the choice of variable grouping in this approach disentangles the locally loopy structure of the factor graph associated with the space-time problem. This disentanglement is achieved due to the linear coupling between variables on neighboring nodes, which is obtained by introducing auxiliary local fields $\hat{\boldsymbol{h}}_i$.

According to this graphical model construction, we obtain the *dynamic cavity equations* Eq. (12) as an ansatz for describing the stochastic dynamics associated with the dynamic partition function $\mathcal{Z}(\boldsymbol{\mathcal{O}})$ on a treelike interaction graph,

$$c_{i\setminus j}(\boldsymbol{x}_i, \hat{\boldsymbol{h}}_i|\boldsymbol{\mathcal{O}}) = \frac{p_0\big(x_i^0\big)}{\mathcal{Z}_{i\setminus j}(\boldsymbol{\mathcal{O}})} p(\boldsymbol{O}_i|\boldsymbol{x}_i) e^{S_i^0} \prod_{k\in\partial i\setminus j} \sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k c_{k\setminus i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}}) e^{S_{ik}^{\mathrm{int}}}. \tag{B11}$$

Then, using the Fourier transforms defined as Eqs. (13) the dynamic cavity equations can be written as

$$\frac{\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\boldsymbol{\mathcal{O}})}{p_0\big(x_i^0\big) p(\boldsymbol{O}_i|\boldsymbol{x}_i)} \propto \int D\hat{\boldsymbol{h}}_i \left(\prod_t e^{-\mathrm{i}s_i^t \hat{h}_i^t}\right) e^{S_i^0} \prod_{k\in\partial i\setminus j} \sum_{\boldsymbol{x}_k} \int D\hat{\boldsymbol{h}}_k \int D\boldsymbol{s}_k \left(\prod_t \frac{e^{\mathrm{i}s_k^t \hat{h}_k^t}}{2\pi}\right) \hat{c}_{k\setminus i}(\boldsymbol{x}_k, \hat{\boldsymbol{h}}_k|\boldsymbol{\mathcal{O}}) e^{S_{ik}^{\mathrm{int}}} \tag{B12a}$$

$$\propto \prod_t \int d\hat{h}_i^t \, e^{-\mathrm{i}s_i^t \hat{h}_i^t} \big[\delta\big(\hat{h}_i^t - \mathrm{i}\big)\alpha_i^t e^{h_i^t} \big(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\big) + \delta\big(\hat{h}_i^t\big)\delta_{x_i^{t+1},I}\big]$$

$$\times \prod_{k\in\partial i\setminus j} \left(\prod_t \sum_{x_k^t} \int d\hat{h}_k^t ds_k^t \frac{e^{\mathrm{i}s_k^t \hat{h}_k^t}}{2\pi} e^{-\mathrm{i}(\hat{h}_k^t \nu_{ik}^t \delta_{x_i^t,I} + \hat{h}_i^t \nu_{ki}^t \delta_{x_k^t,I})}\right) \hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{s}_k|\boldsymbol{\mathcal{O}}). \tag{B12b}$$

The expression can be simplified by performing the integrals over the conjugate fields coming from the neighbors $\hat{h}_k^t$ first, then over the neighboring fields s_k^t,

$$\frac{\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\boldsymbol{\mathcal{O}})}{p_0(x_i^0)p(\boldsymbol{O}_i|\boldsymbol{x}_i)} \propto \prod_t \int d\hat{h}_i^t\, e^{-\mathrm{i}s_i^t\hat{h}_i^t}\left[\delta(\hat{h}_i^t - \mathrm{i})\alpha_i^t e^{h_i^t}\left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta(\hat{h}_i^t)\delta_{x_i^{t+1},I}\right]$$

$$\times \prod_{k\in\partial i\setminus j}\left(\prod_t \sum_{x_k^t}\int ds_k^t\, \delta(s_k^t - \nu_{ik}^t\delta_{x_i^t,I}) e^{-\mathrm{i}\hat{h}_i^t\nu_{ki}^t\delta_{x_k^t,I}}\right)\hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{s}_k|\boldsymbol{\mathcal{O}}) \tag{B13a}$$

$$\propto \prod_t \int d\hat{h}_i^t\, e^{-\mathrm{i}s_i^t\hat{h}_i^t}\left[\delta(\hat{h}_i^t - \mathrm{i})\alpha_i^t e^{h_i^t}\left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta(\hat{h}_i^t)\delta_{x_i^{t+1},I}\right]\prod_{k\in\partial i\setminus j}\left(\prod_t\sum_{x_k^t} e^{-\mathrm{i}\hat{h}_i^t\nu_{ki}^t\delta_{x_k^t,I}}\right)\hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}) \tag{B13b}$$

$$\propto \sum_{\boldsymbol{x}_{\partial i\setminus j}}\left(\prod_{k\in\partial i\setminus j}\hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}})\right)\prod_t\int d\hat{h}_i^t\, e^{-\mathrm{i}\hat{h}_i^t(s_i^t + \sum_{k\in\partial i\setminus j}\nu_{ki}^t\delta_{x_k^t,I})}\left[\delta(\hat{h}_i^t - \mathrm{i})\alpha_i^t e^{h_i^t}\left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right) + \delta(\hat{h}_i^t)\delta_{x_i^{t+1},I}\right], \tag{B13c}$$

and finally over the local conjugate fields $\hat{h}_i^t$

$$\frac{\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\boldsymbol{\mathcal{O}})}{p_0(x_i^0)p(\boldsymbol{O}_i|\boldsymbol{x}_i)} \propto \sum_{\boldsymbol{x}_{\partial i\setminus j}}\hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}})\prod_t\left[\delta_{x_i^{t+1},I} + e^{s_i^t + \sum_{k\in\partial i\setminus j}\nu_{ki}^t\delta_{x_k^t,I}}\alpha_i^t e^{h_i^t}\left(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\right)\right]. \tag{B14}$$

By introducing the dynamic cavity action Eq. (8) we obtain the dynamic cavity equations Eq. (7), where cavity messages are now a function of the variable-field trajectories $\boldsymbol{x}_i, \boldsymbol{s}_i$,

$$\hat{c}_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i|\boldsymbol{\mathcal{O}}) \propto p_0(x_i^0)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\sum_{\boldsymbol{x}_{\partial i\setminus j}} e^{\hat{S}_{i\setminus j}}\prod_{k\in\partial i\setminus j}\hat{c}_{k\setminus i}(\boldsymbol{x}_k, \boldsymbol{\nu}_{ik}\boldsymbol{x}_i|\boldsymbol{\mathcal{O}}). \tag{B15}$$

APPENDIX C: REDUCTION TO THE TIME-FORWARD EQUATIONS IN THE ABSENCE OF OBSERVATIONS

A major consequence of the introduction of time-forward messages $\rho_{\to t}^{i\setminus j}$ and time-backward messages $\rho_{t\leftarrow}^{i\setminus j}$ is that, in the absence of observations, it is possible to prove that the quantities $\mu_{i\setminus j}^t$ have to vanish for all edges (i, j) and times t and then recover a purely time-forward dynamics.

Using the definition of $m_{i\setminus j}^t$ by means of the quantities $\rho_{\to t}^{i\setminus j}(x_i^t)$ and $\rho_{t\leftarrow}^{i\setminus j}(x_i^t)$, i.e., Eq (34), but performing the slicing one time step later, we obtain

$$m_{i\setminus j}^t = \frac{1}{Z_{i\setminus j}^t}\sum_{x_i^t, x_i^{t+1}}\rho_{\to t}^{i\setminus j}(x_i^t)\delta_{x_i^t,I}M_{x_i^t x_i^{t+1}}^{i\setminus j}\rho_{t+1\leftarrow}^{i\setminus j}(x_i^{t+1}) \tag{C1a}$$

$$= \frac{1}{Z_{i\setminus j}^t}\rho_{\to t}^{i\setminus j}(I)M_{t,II}^{i\setminus j}\rho_{t+1\leftarrow}^{i\setminus j}(I) \tag{C1b}$$

or slicing one time step earlier,

$$m_{i\setminus j}^t = \frac{1}{Z_{i\setminus j}^t}\sum_{x_i^{t-1}, x_i^t}\rho_{\to t-1}^{i\setminus j}(x_i^{t-1})x_i^t M_{x_i^{t-1}x_i^t}^{i\setminus j}\rho_{t\leftarrow}^{i\setminus j}(x_i^t) \tag{C2a}$$

$$= \frac{1}{Z_{i\setminus j}^t}\left[\rho_{\to t-1}^{i\setminus j}(I)M_{t-1,II}^{i\setminus j}\rho_{t\leftarrow}^{i\setminus j}(I) + \rho_{\to t-1}^{i\setminus j}(S)M_{t-1,SI}^{i\setminus j}\rho_{t\leftarrow}^{i\setminus j}(I)\right], \tag{C2b}$$

where we have introduced the normalization constant $Z_{i\setminus j}^t = \sum_{x_i^t}\rho_{\to t}^{i\setminus j}(x_i^t)\rho_{t\leftarrow}^{i\setminus j}(x_i^t)$. It is straightforward to show that it corresponds to the normalization constant $\mathcal{Z}_{i\setminus j}$, defined in Eqs. (28) and (32). It follows that the normalization is independent of time. Using Eq. (C1b) it is possible to express $m_{i\setminus j}^t$ as function of $m_{i\setminus j}^{t-1}$,

$$m_{i\setminus j}^t = m_{i\setminus j}^{t-1} + \frac{\rho_{\to t-1}^{i\setminus j}(S)M_{t-1,SI}^{i\setminus j}\rho_{t\leftarrow}^{i\setminus j}(1)}{\mathcal{Z}_{i\setminus j}} \tag{C3a}$$

$$= m_{i\setminus j}^{t-1} + \left(1 - m_{i\setminus j}^{t-1}\right)\frac{\left(1 - g_{i\setminus j}^{t-1}\right)\rho_{t\leftarrow}^{i\setminus j}(I)}{\left(1 - g_{i\setminus j}^{t-1}\right)\rho_{t\leftarrow}^{i\setminus j}(I) + g_{i\setminus j}^{t-1}\rho_{t\leftarrow}^{i\setminus j}(S)}. \tag{C3b}$$

As already stressed in the main text, the last expression does not represent a time-forward equation because the quantities $\rho^{i\setminus j}_{t\leftarrow}(x^t_i)$ are computed backward in time from T to step t. Time-forward dynamics is recovered if the two time-backward messages are uniform, which is expected to occur in the absence of observations at later times. To prove this, one can first notice that from Eq. (37), $\mu^t_{i\setminus j}=0$ if the time-backward messages are uniform, i.e., if $\rho^{i\setminus j}_{t+1\leftarrow}(S)=\rho^{i\setminus j}_{t+1\leftarrow}(I)$. In the absence of observations also the inverse implication is true: when the set of messages $\mu^t_{i\setminus j}$ at time t are zero and there is no observation also at time t, then the corresponding time-backward messages $\rho^{i\setminus j}_{t\leftarrow}(x^t_i)$ are also uniform. Let us start from the time $T-1$, because by construction $\mu^T_{i\setminus j}=0$, then using Eq. (37) with the final time condition $\rho^{i\setminus j}_{T\leftarrow}(x^T_i)=p(O^T_i|x^T_i)$ we obtain

$$\mu^t_{i\setminus j}=\frac{1}{\mathcal{Z}_{i\setminus j}}\rho^{i\setminus j}_{\to t}(S)M^{i\setminus j}_{T-1,SS}\big(p(O^T_i|S)-p(O^T_i|I)\big). \tag{C4}$$

If no observation is provided on the final time, then $p(O^T_i|x^T_i)=1$ for $x^T_i=S,I$ and the numerator vanishes, that is $\mu^{T-1}_{i\setminus j}=0$. Moreover,

$$\rho^{i\setminus j}_{T-1\leftarrow}\big(x^{T-1}_i\big)=\sum_{x^T_i}M^{i\setminus j}_{x^{T-1}_i x^T_i}p\big(O^T_i|x^T_i\big) \tag{C5a}$$

$$=M^{i\setminus j}_{x^{T-1}_i S}+M^{i\setminus j}_{x^{T-1}_i I}, \tag{C5b}$$

that is

$$\rho^{i\setminus j}_{T-1\leftarrow}(S)=g^{T-1}_{i\setminus j}p\big(O^{T-1}_i|S\big)+\big(1-g^{T-1}_{i\setminus j}\big)p\big(O^{T-1}_i|S\big)=p\big(O^{T-1}_i|S\big), \tag{C6}$$

$$\rho^{i\setminus j}_{T-1\leftarrow}(I)=e^{\sum_{k\in\partial i\setminus j}\nu^{T-1}_{ik}\mu^{T-1}_{k\setminus i}}p\big(O^{T-1}_i|I\big)=p\big(O^{T-1}_i|I\big), \tag{C7}$$

meaning that $\rho^{i\setminus j}_{T-1\leftarrow}(S)=\rho^{i\setminus j}_{T-1\leftarrow}(I)$ if no observation is included at time $T-1$. In this way, the equality is guaranteed at time $T-1$ and one can proceed by induction. By assuming that, in the absence of observations at times larger than t, the equality is valid for time $t+1$, i.e., $\rho^{i\setminus j}_{t+1\leftarrow}(S)=\rho^{i\setminus j}_{t+1\leftarrow}(I)=\rho^{i\setminus j}_{t+1\leftarrow}$ for all edges (i,j), then one obtains that $\mu^t_{i\setminus j}=0$. Computing the time-backward messages at time t,

$$\rho^{i\setminus j}_{t\leftarrow}\big(x^t_i\big)=\sum_{x^t_i,x^{t+1}_i}M^{i\setminus j}_{x^t_i x^{t+1}_i}\rho^{i\setminus j}_{t+1\leftarrow}\big(x^{t+1}_i\big) \tag{C8a}$$

$$=\big(M^{i\setminus j}_{x^t_i S}+M^{i\setminus j}_{x^t_i I}\big)\rho^{i\setminus j}_{t+1\leftarrow}, \tag{C8b}$$

and using that all $\mu^t_{k\setminus i}$ vanish,

$$\rho^{i\setminus j}_{t\leftarrow}\big(x^t_i\big)\propto\begin{cases}p\big(O^t_i|S\big) & \text{if } x^t_i=S,\\ p\big(O^t_i|I\big) & \text{if } x^t_i=I.\end{cases} \tag{C9}$$

By induction, this is true for every time, as long as no observation is included. Hence, it is possible to conclude that, in the absence of observations, the equations (C3b) for the cavity marginals $m^t_{i\setminus j}$ reduce to the more standard time-forward mean-field equations in Eqs. (42). Figure 11 shows how the value of λ influences the performance of these methods.

APPENDIX D: EFFICIENT IMPLEMENTATION IN THE INFECTION TIME REPRESENTATION

As an alternative to the generic efficient formulation presented in Sec. III in terms of transfer-matrix formalism, the computational complexity of Eqs. (24) and (25) can be reduced from exponential (in the temporal length T) to polynomial exploiting the nonrecurrence of the SI dynamic process—in which only configurations of the type $\boldsymbol{x}_i=(0,\dots,0,1,\dots 1)$ are allowed—by using a simpler representation in terms of infection times. An epidemic trajectory of the SI process can be parameterized by a unique set of integer variables t_i (one for each node) representing the first time at which individual i is infected, and taking values in $t_i\in\{0,\dots,T+1\}$. The case $t_i=0$ corresponds to individual i being originally infected at the initial time, i.e., being a patient-zero of the epidemics. The other special case $t_i=T+1$ models the scenario where individual i never gets infected during the dynamics (which formally corresponds to $t_i=+\infty$). The trajectory $\boldsymbol{x}_i$ can be simply expressed as $x^t_i=S\Theta(t_i-1-t)+I\Theta(t-t_i)$, where $\Theta(x)$ is a Heaviside-step function, with the convention $\Theta(0)=1$. After some algebra, we can rewrite Eqs. (24), (25), and (28) as follows:

$$m^t_{i\setminus j}=\frac{1}{\mathcal{Z}_{i\setminus j}}\sum_{t_i=0}^{t}p(\boldsymbol{O}_i|t_i)\gamma_0(t_i)\left(\prod_{r=0}^{t_i-2}\alpha_i^r e^{\sum_{k\in\partial i\setminus j}m^r_{k\setminus i}\nu^r_{ki}}\right)\big(1-\mathbb{I}_{1\leqslant t_i\leqslant T}\alpha_i^{t_i-1}e^{\sum_{k\in\partial i\setminus j}m^{t_i-1}_{k\setminus i}\nu^{t_i-1}_{ki}}\big)\prod_{s=t_i}^{T-1}e^{\sum_{k\in\partial i\setminus j}\nu^s_{ik}\mu^s_{k\setminus i}}, \tag{D1a}$$

$$\mu^t_{i\setminus j} = \frac{1}{\mathcal{Z}_{i\setminus j}} \sum_{t_i=t+2}^{T+1} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) \left(\prod_{r=0}^{t_i-2} \alpha_i^r e^{\sum_{k\in\partial i\setminus j} m^r_{k\setminus i} v^r_{ki}}\right) \left(1 - \mathbb{I}_{1\leqslant t_i\leqslant T} \alpha_i^{t_i-1} e^{\sum_{k\in\partial i\setminus j} m^{t_i-1}_{k\setminus i} v^{t_i-1}_{ki}}\right) \prod_{s=t_i}^{T-1} e^{\sum_{k\in\partial i\setminus j} v^s_{ik}\mu^s_{k\setminus i}}$$

$$- \frac{\mathbb{I}_{0\leqslant t\leqslant T-1}}{\mathcal{Z}_{i\setminus j}} p(\boldsymbol{O}_i|t+1)\gamma_0(t+1) \left(\prod_{r=0}^{t} \alpha_i^r e^{\sum_{k\in\partial i\setminus j} m^r_{k\setminus i} v^r_{ki}}\right) \prod_{s=t+1}^{T-1} e^{\sum_{k\in\partial i\setminus j} v^s_{ik}\mu^s_{k\setminus i}}, \tag{D1b}$$

$$\mathcal{Z}_{i\setminus j} = \sum_{t_i=0}^{T+1} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) \left(\prod_{r=0}^{t_i-2} \alpha_i^r e^{\sum_{k\in\partial i\setminus j} m^r_{k\setminus i} v^r_{ki}}\right) \left(1 - \mathbb{I}_{1\leqslant t_i\leqslant T} \alpha_i^{t_i-1} e^{\sum_{k\in\partial i\setminus j} m^{t_i-1}_{k\setminus i} v^{t_i-1}_{ki}}\right) \prod_{s=t_i}^{T-1} e^{\sum_{k\in\partial i\setminus j} v^s_{ik}\mu^s_{k\setminus i}}, \tag{D1c}$$

where the function $\gamma_0(t_i)$ is related to the probability of node i being a patient zero, namely

$$\gamma_0(t_i) = \begin{cases} p_0\big(x_i^0 = I\big), & t_i = 0, \\ 1 - p_0\big(x_i^0 = I\big), & t_i > 0. \end{cases} \tag{D2}$$

We have introduced the indicator function $\mathbb{I}_x$ which is equal to one when its argument x is true, and zero otherwise. Notice how all the summations and products w.r.t. the infection times are now linear in T. Analogously, the likelihood term for each observations on node i [Eq. (5)] can be rewritten under this representation as

$$p\big(O_i^t = S\big|t_i\big) = (1-\eta^+)\Theta(t_i - t - 1) + \eta^-\Theta(t - t_i), \tag{D3}$$

$$p\big(O_i^t = I\big|t_i\big) = \eta^+\Theta(t_i - t - 1) + (1-\eta^-)\Theta(t - t_i), \tag{D4}$$

where $p(\boldsymbol{O}_i|t_i) = \prod_{t=0}^{T} p(O_i^t|t_i)$. Analogous expressions w.r.t. Eqs. (D1a) and (D1c) can be derived for the single-node marginal m_i^t and its normalization $\mathcal{Z}_i$. An efficient computational scheme can be attained by updating all the cavity quantities of a fixed node at once, and then performing a random shuffling on the order in which nodes are updated. Intuitively, the speed-up induced by this protocol is that the forward and backward contribution to each cavity message - say, on link (i, j) can be computed by removing the corresponding link contribution from the *site* term i. To clarify this point, let us define the following four quantities, for each node i:

$$\mathcal{R}_i^{\rightarrow t} = \log \alpha_i^t + \sum_{k\in\partial i} m^t_{k\setminus i} v^t_{ki}, \tag{D5}$$

$$\mathcal{R}_i^{t\leftarrow} = \sum_{k\in\partial i} v^t_{ik}\mu^t_{k\setminus i}, \tag{D6}$$

$$K_i^{\rightarrow t} = \sum_{r=0}^{t-2} \mathcal{R}_i^{\rightarrow r}, \tag{D7}$$

$$K_i^{t\leftarrow} = \sum_{s=t}^{T-1} \mathcal{R}_i^{s\leftarrow}. \tag{D8}$$

Analogous definitions for the cavity quantities hold for each edge (i, j), just by letting the above summations run over all the neighbors of node i but j. Equations (D5) and (D6) have also a physical interpretation. For instance, $\mathcal{R}_i^{\rightarrow t}$ is a mean-field approximation for the log-probability of node i not being infected at time t by none of its neighbors. The above definitions Eqs. (D5)–(D8) allow one to write the cavity equations Eqs. (D1a)–(D1c) in a more compact form:

$$m^t_{i\setminus j} = \frac{1}{\mathcal{Z}_{i\setminus j}} \sum_{t_i=0}^{t} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) e^{K^{\rightarrow t_i}_{i\setminus j}} \theta^{\rightarrow t_i - 1}_{i\setminus j} e^{K^{t_i\leftarrow}_{i\setminus j}}, \tag{D9}$$

$$\mu^t_{i\setminus j} = \frac{1}{\mathcal{Z}_{i\setminus j}} \Bigg(\sum_{t_i=t+2}^{T+1} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) e^{K^{\rightarrow t_i}_{i\setminus j}} \theta^{\rightarrow t_i - 1}_{i\setminus j} e^{K^{t_i\leftarrow}_{i\setminus j}} - p(\boldsymbol{O}_i|t+1)\gamma_0(t+1)\mathbb{I}_{0\leqslant t\leqslant T-1} e^{K^{\rightarrow t+2}_{i\setminus j} + K^{t+1\leftarrow}_{i\setminus j}} \Bigg), \tag{D10}$$

$$\mathcal{Z}_{i\setminus j} = \sum_{t_i=0}^{T+1} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) e^{K^{\rightarrow t_i}_{i\setminus j}} \theta^{\rightarrow t_i - 1}_{i\setminus j} e^{K^{t_i\leftarrow}_{i\setminus j}}, \tag{D11}$$

where we have introduced $\theta^{\rightarrow t}_{i\setminus j} = 1 - \mathbb{I}_{1\leqslant t+1\leqslant T} e^{\mathcal{R}^{\rightarrow t}_{i\setminus j}}$.

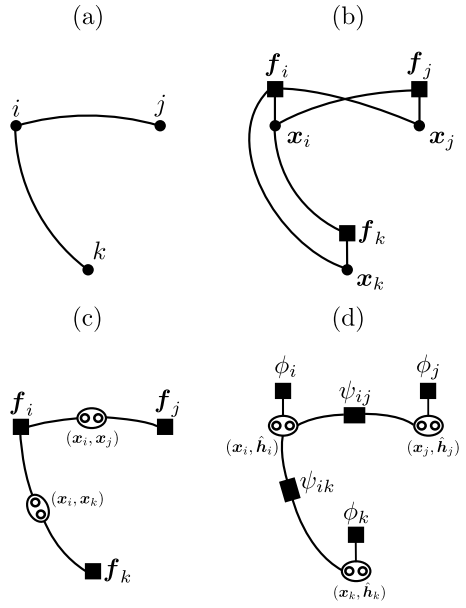


FIG. 10. Factor graph representation for epidemic inference. (a) Original contact graph. (b) Loopy, naive factor graph associated to the graphical model Eq. (A1). (c) Disentangled dual factor graph for the graphical model Eq. (A1). The factor graph maintain the structure of the original contact graph. (c) Disentangled factor graph for the graphical model interpretation of the dynamical partition function in Eq. (B8) with factors as given in Eqs. (B9) and (B10).

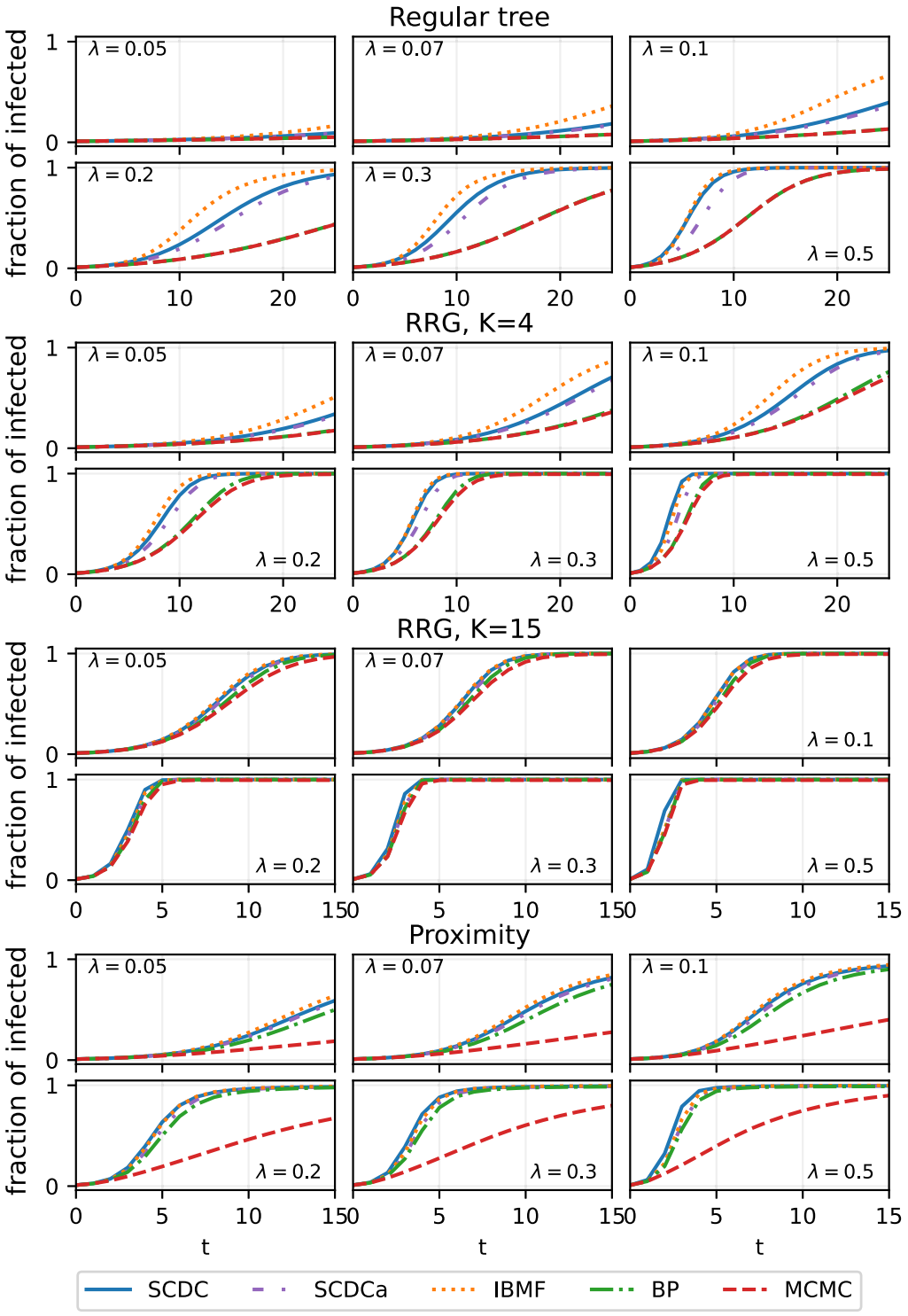


FIG. 11. Forward dynamics with varying λ. Fraction of infected individuals against epidemic time, with four different static contact graphs. From top to bottom: regular tree with degree $K = 4$ and $N = 485$, RRG with $N = 500$ and degree $K = 4$, RRG with $N = 500$ and $K = 15$, proximity graph with $N = 500$. Comparison is shown between SCDC, IBMF, BP, and Monte Carlo simulations (with 10^4 samples). All the links have the same infection probability λ, whose value is reported inside each panel. Six different values of λ are used. In all cases, the probability of each individual being infected at time $t = 0$ is set to $p_0(x_i^0) = 5/N$, and the self-infection ε_i^t is set to 0 for every node i.

At fixed i, the quantities $\mathcal{R}^{\rightarrow t}_{i\backslash j}$, $\mathcal{R}^{t\leftarrow}_{i\backslash j}$, $K^{\rightarrow t}_{i\backslash j}$, $K^{t\leftarrow}_{i\backslash j}$ can be computed for each cavity by removing just one link contribution (i.e., the one corresponding to the link removed in that specific cavity graph), without further $O(|\partial i|)$ computations, namely

$$\mathcal{R}^{\rightarrow t}_{i\backslash j} = \mathcal{R}^{\rightarrow t}_{i} - m^{t}_{j\backslash i} v^{t}_{ji}, \tag{D12}$$

$$\mathcal{R}^{t\leftarrow}_{i\backslash j} = \mathcal{R}^{t\leftarrow}_{i} - v^{t}_{ij} \mu^{t}_{j\backslash i}, \tag{D13}$$

$$K^{\rightarrow t}_{i\backslash j} = K^{\rightarrow t}_{i} - \sum_{r=0}^{t-2} m^{r}_{j\backslash i} v^{r}_{ji}, \tag{D14}$$

$$K^{t\leftarrow}_{i\backslash j} = K^{t\leftarrow}_{i} - \sum_{s=t}^{T-1} v^{s}_{ij} \mu^{s}_{j\backslash i}, \tag{D15}$$

for any $j \in \partial i$. Furthermore, the computation of $K^{\rightarrow t}_{i}$ and $K^{t\leftarrow}_{i}$ can be done recursively, in a forward (respectively, backward) direction w.r.t. time, i.e., by exploiting $K^{\rightarrow t}_{i} = K^{\rightarrow t-1}_{i} + \mathcal{R}^{\rightarrow t-2}_{i}$ and $K^{t\leftarrow}_{i} = K^{t+1\leftarrow}_{i} + \mathcal{R}^{t\leftarrow}_{i}$. Clearly, equivalent relations hold for the cavity quantities $K^{\rightarrow t}_{i\backslash j}$ and $K^{t\leftarrow}_{i\backslash j}$. Using all the above schemes, the overall computational cost to perform a single update of all the cavity quantities for a node i scales as $O(|\partial i|T)$. A further advantage of such a computational scheme is that the update of all the cavities for one node can be performed in parallel, a particularly convenient choice especially when dealing with dense graphs. The convergence criterion can be defined either w.r.t. the cavity messages $\{m^{t}_{i\backslash j}\}$ and/or their conjugates $\{\mu^{t}_{i\backslash j}\}$, or eventually w.r.t. the single-site marginals $\{m^{t}_{i}\}$: the latters can be computed as

$$m^{t}_{i} = \frac{\sum_{t_i=0}^{t} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) e^{K^{\rightarrow}_{i}(t_i)} \theta_i^{t_i-1} e^{K^{\leftarrow}_{i}(t_i)}}{\sum_{t_i=0}^{T+1} p(\boldsymbol{O}_i|t_i)\gamma_0(t_i) e^{K^{\rightarrow}_{i}(t_i)} \theta_i^{t_i-1} e^{K^{\leftarrow}_{i}(t_i)}}, \tag{D16}$$

where the normalization is explicitly shown at the denominator, and θ^{t}_{i} is the straightforward extension of $\theta^{t}_{i\backslash j}$ to the single node case. This expression is equivalent to Eq. (29) of the main text but rewritten using the infection-time representation just discussed.

APPENDIX E: DERIVATION OF SCDC ON THE SIR MODEL

In a SIR model, the possible individual states are $x^{t}_{i} \in \{S, I, R\}$ and the transition probabilities between states are given by

$$W_i\big(x^{t+1}_{i} = S \big| \boldsymbol{x}^{t}\big) = \delta_{x^{t}_{i},S} \alpha^{t}_{i} e^{h^{t}_{i}}, \tag{E1}$$

$$W_i\big(x^{t+1}_{i} = I \big| \boldsymbol{x}^{t}\big) = \delta_{x^{t}_{i},I}\big(1 - r^{t}_{i}\big) + \delta_{x^{t}_{i},S}\big[1 - \alpha^{t}_{i} e^{h^{t}_{i}}\big], \tag{E2}$$

$$W_i\big(x^{t+1}_{i} = R \big| \boldsymbol{x}^{t}\big) = \delta_{x^{t}_{i},R} + \delta_{x^{t}_{i},I} r^{t}_{i}, \tag{E3}$$

where r^{t}_{i} is the recovery probability of individual i at time t and h^{t}_{i} is the usual local field, defined as $h^{t}_{i} = \sum_{j=1}^{N} v^{t}_{ji} \delta_{x^{t}_{j},I}$, where $v^{t}_{ji} = \log(1 - \lambda^{t}_{ji})$, such that $e^{h^{t}_{i}} = \prod_{j=1}^{N}(1 - \lambda^{t}_{ji}\delta_{x^{t}_{j},I})$ is the probability of not being infected by the neighbors. The dynamical partition function of the system can be computed following the same steps of the derivation for the SI model, i.e., introducing the local fields h^{t}_{i} through the integral representation of a Dirac δ function, and integrating over them at all them. After some calculations, we end up with the following expression for the dynamical partition function of the observation reweighted SIR dynamics:

$$\begin{aligned}\mathcal{Z}(\boldsymbol{\mathcal{O}}) = \sum_{\boldsymbol{X}} \int \mathcal{D}\hat{\boldsymbol{H}} \prod_{i} p(x^{0}_{i}) p_0(\boldsymbol{O}_i|\boldsymbol{x}_i) \prod_{t} \Big\{ \delta\big(\hat{h}^{t}_{i}\big)\Big[\delta_{x^{t+1}_{i},I}\big(\delta_{x^{t}_{i},I}\big(1 - r^{t}_{i}\big) + \delta_{x^{t}_{i},S}\big) + \delta_{x^{t+1}_{i},R}\big(\delta_{x^{t}_{i},R} + \delta_{x^{t}_{i},I} r^{t}_{i}\big)\Big] \\ + \delta\big(\hat{h}^{t}_{i} - \mathrm{i}\big)\alpha^{t}_{i}\big[\delta_{x^{t+1}_{i},S}\delta_{x^{t}_{i},S} - \delta_{x^{t+1}_{i},I}\delta_{x^{t}_{i},S}\big]\Big\} \prod_{j>i} e^{\delta_{x^{t}_{j},I} v^{t}_{ji}(-i\hat{h}^{t}_{i}) + \delta_{x^{t}_{i},I} v^{t}_{ij}(-i\hat{h}^{t}_{j})},\end{aligned} \tag{E4}$$

which can in turn be interpreted as a graphical model in a similar way to what was done for the SI model. The dynamic cavity equations for the SIR model with observations have the same form of the SI model [Eq. (12)], where the only difference is in the definition of the sum over the trajectories, which now runs over $\{S, I, R\}$, and in the definition of the local noninteracting action

$$S^{0}_{i} = \sum_{t} \log\Big\{\delta\big(\hat{h}^{t}_{i}\big)\Big[\delta_{x^{t+1}_{i},I}\big(\delta_{x^{t}_{i},I}\big(1 - r^{t}_{i}\big) + \delta_{x^{t}_{i},S}\big) + \delta_{x^{t+1}_{i},R}\big(\delta_{x^{t}_{i},R} + \delta_{x^{t}_{i},I} r^{t}_{i}\big)\Big] + \delta\big(\hat{h}^{t}_{i} - \mathrm{i}\big)\alpha^{t}_{i}\big[\delta_{x^{t+1}_{i},S}\delta_{x^{t}_{i},S} - \delta_{x^{t+1}_{i},I}\delta_{x^{t}_{i},S}\big]\Big\}. \tag{E5}$$

The expansion for small infection rates can be carried out similarly to the SI model, leading to the same expression for the cavity averages Eqs. (24) and (25), where the two term $S^{m}_{i\backslash j}$ and $S^{\mu}_{i\backslash j}$ are now defined as

$$\begin{aligned}S^{m}_{i\backslash j} = \sum_{t} \log\Big\{\delta_{x^{t+1}_{i},I}\big(\delta_{x^{t}_{i},I}\big(1 - r^{t}_{i}\big) + \delta_{x^{t}_{i},S}\big) + \delta_{x^{t+1}_{i},R}\big(\delta_{x^{t}_{i},R} + \delta_{x^{t}_{i},I} r^{t}_{i}\big) \\ + \alpha^{t}_{i}\big(\delta_{x^{t+1}_{i},S}\delta_{x^{t}_{i},S} - \delta_{x^{t+1}_{i},I}\delta_{x^{t}_{i},S}\big) e^{\sum_{k\in\partial i\backslash j} m^{t}_{k\backslash i} v^{t}_{ki}}\Big\} + \sum_{t}\sum_{k\in\partial i\backslash j} \delta_{x^{t}_{i},I} v^{t}_{ik} \mu^{t}_{k\backslash i},\end{aligned} \tag{E6}$$

$$S^{\mu}_{i\backslash j} = \sum_{t} \log\big[\alpha^{t}_{i}\big(\delta_{x^{t+1}_{i},S}\delta_{x^{t}_{i},S} - \delta_{x^{t+1}_{i},I}\delta_{x^{t}_{i},S}\big)\big] + \sum_{t}\sum_{k\in\partial i\backslash j}\big(m^{t}_{k\backslash i} v^{t}_{ki} + \delta_{x^{t}_{i},I} v^{t}_{ik} \mu^{t}_{k\backslash i}\big). \tag{E7}$$

The normalization function is defined as usual by Eq. (28), and ensures normalization of the cavity averages $m^t_{i\setminus j}$, which are therefore interpreted as the probability of node i of being infected at time t when node j is removed from the contact graph. Even with the new definitions of $S^m_{i\setminus j}$ and $S^\mu_{i\setminus j}$, we can use the efficient formulation Eq. (32) in terms of transfer matrices $M_{x^t_i,x^{t+1}_i}$, with the only difference in their expression. The transition matrices are now 3×3, and defined by Eq. (39).

APPENDIX F: EXAMPLE OF NORMALIZATION ISSUE FOR LEAVES OF THE CONTACT GRAPH

The small-coupling expansion requires to assume the normalization $\mathcal{Z}_{i\setminus j}$ in Eq. (28), which sums over all the possible trajectories of node i assuming $\boldsymbol{s}_i = 0$ at every time. The method thus considers all the trajectories in which the cavity node j is always susceptible, and therefore cannot infect node i. In particular, there are situations, such as the one considered in the example below, in which the normalization vanishes, meaning that it is not possible to explain an observed trajectory within the standard SI model. While this could seem pathological, it is worth stressing that the assumption done is necessary to obtain a message-passing algorithm that is independent of the trajectory of node j, a crucial condition to perform the expansion on which the present method is based. It is however possible to ensure that every trajectory of a node i remains feasible, the normalization constant being finite, by slightly modifying the epidemic model introducing a small self-infection probability. In addition to fix the normalization issue, a small value of self-infection probability does not deteriorate the predictive power of the method.

To better illustrate this problem, we consider a leaf node i and its unique neighbor j. In the cavity graph corresponding to the message $c_{i\setminus j}(\boldsymbol{x}_i, \boldsymbol{s}_i)$, node i will appear as an isolated node. As a consequence, it is expected that the approximation behind the SCDC equations cannot explain, within the cavity graph, an infection actually transmitted from node j to node i. Indeed, because of the absence of further neighbors, the normalization term reads, after integration of the conjugate field $\boldsymbol{h}_i$,

$$\mathcal{Z}_{i\setminus j} = \sum_{\boldsymbol{x}_i} p_0\big(x_i^0\big)p(\boldsymbol{O}_i|\boldsymbol{x}_i)\prod_t\Big[\alpha_i^t\big(\delta_{x_i^{t+1},x_i^t} - \delta_{x_i^{t+1},I}\big) + \delta_{x_i^{t+1},I}\Big], \tag{F1}$$

showing that an infection can only be explained by a self-infection event. When $\varepsilon_i^t = 0$ (i.e., $\alpha_i^t = 1$, no self-infections possible), the cavity message admits trajectories for which node i is always susceptible or infected. When a perfect observation (i.e., $\eta^+ = \eta^- = 0$) imply an infection event at some $t \neq 0$, the normalization vanishes, indicating an inconsistency in the model. This is prevented by the existence of a finite self-infection probability. Since it is recommended to operate in the limit of a vanishing self-infection, in the present case it is possible to analytically verify the limiting behavior for the cavity messages $m^t_{i\setminus j}$ and $\mu^t_{i\setminus j}$.

As an example, we suppose that the leaf i is observed to be susceptible at time t_S and then is observed to be infected at time $t_I > t_S$. We consider a uniform self-infection probability $\varepsilon_i^t = \varepsilon$ ($\alpha_i^t = \alpha = 1-\varepsilon$) for any time t and any node i, and a uniform prior probability $p_0(x_i^0 = S) = 1-\gamma$, $p_0(x_i^0 = I) = \gamma$. The forward messages are

$$\rho^{i\setminus j}_{\rightarrow t}(S) = \begin{cases}(1-\gamma)\alpha^t & \text{if } t \leqslant t_I,\\ 0 & \text{if } t > t_I,\end{cases} \tag{F2}$$

$$\rho^{i\setminus j}_{\rightarrow t}(I) = \begin{cases}\gamma + (1-\alpha)(1-\gamma)\sum_{l=0}^{t-1}\alpha^l & \text{if } t \leqslant t_S,\\ (1-\alpha)(1-\gamma)\sum_{l=t_S}^{t-1}\alpha^l & \text{if } t_S < t \leqslant t_I,\\ (1-\alpha)(1-\gamma)\sum_{l=t_S}^{t_I-1}\alpha^l & \text{if } t > t_I,\end{cases} \tag{F3}$$

and the backward messages

$$\rho^{i\setminus j}_{t\leftarrow}(S) = \begin{cases}(1-\alpha)\alpha^{t_S-t}\sum_{l=0}^{t_I-1-t_S}\alpha^l & \text{if } t \leqslant t_S,\\ (1-\alpha)\sum_{l=0}^{t_I-1-t}\alpha^l & \text{if } t_S < t \leqslant t_I,\\ 1 & \text{if } t > t_I,\end{cases} \tag{F4}$$

$$\rho^{i\setminus j}_{t\leftarrow}(I) = \begin{cases}0 & \text{if } t \leqslant t_S,\\ 1 & \text{if } t > t_S.\end{cases} \tag{F5}$$

The normalization factor, taking into account the observations, is given by

$$\mathcal{Z}_{i\setminus j} = (1-\alpha)(1-\gamma)\sum_{l=t_S}^{t_I-1}\alpha^l, \tag{F6}$$

so that the cavity marginal is given by

$$m^t_{i\setminus j} = \begin{cases}0 & \text{if } t \leqslant t_S,\\ \frac{(1-\alpha)(1-\gamma)\sum_{l=t_S}^{t-1}\alpha^l}{(1-\alpha)(1-\gamma)\sum_{l=t_S}^{t_I-1}\alpha^l} & \text{if } t_S < t \leqslant t_I,\\ 1 & \text{if } t > t_I.\end{cases} \tag{F7}$$

In the limit of vanishing self-infection $\varepsilon \to 0$ ($\alpha \to 1$), the cavity marginal takes the simple expression

$$m^t_{i\setminus j} = \begin{cases}0 & \text{if } t \leqslant t_S,\\ \frac{t-t_S}{t_I-t_S} & \text{if } t_S < t \leqslant t_I,\\ 1 & \text{if } t > t_I,\end{cases} \tag{F8}$$

which gives a reasonable probability profile for the node i to be infected in the absence of node i. It is worth stressing that this is not the full marginal m_i^t, which also depends on the messages coming from j to i. The cavity field instead diverges for times t between the two observation times

$$\mu^t_{i\setminus j} = \begin{cases}1 & \text{if } t \leqslant t_S,\\ -\infty & \text{if } t_S < t \leqslant t_I,\\ 0 & \text{if } t > t_I,\end{cases} \tag{F9}$$

which is a clear consequence of having a vanishing normalization factor in the limit of zero self-infection. The divergence of

the cavity field is thus the very nonphysical effect of the inconsistency already discussed. In practice, to avoid divergences triggered by some peculiar combinations of observations, we then implement the algorithm using a cutoff $\mu_c < 0$ on the values of $\mu_{i\setminus j}$, such that the update rule (37) is implemented as follows:

$$\mu^t_{i\setminus j} = \max\left\{\mu_c, \frac{\rho^{i\setminus j}_{\to t}(S) M^{i\setminus j}_{SS}\left(\rho^{i\setminus j}_{t+1\leftarrow}(S) - \rho^{i\setminus j}_{t+1\leftarrow}(I)\right)}{\rho^{i\setminus j}_{\to t}(S)\rho^{i\setminus j}_{t\leftarrow}(S) + \rho^{i\setminus j}_{\to t}(I)\rho^{i\setminus j}_{t\leftarrow}(I)}\right\}. \tag{F10}$$

[1] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, Science **368**, eabb6936 (2020).

[2] A. Baker, I. Biazzo, A. Braunstein, G. Catania, L. Dall'Asta, A. Ingrosso, F. Krzakala, F. Mazza, M. Mézard, A. P. Muntoni, M. Refinetti, S. S. Mannelli, and L. Zdeborová, Epidemic mitigation by statistical inference from contact tracing data, Proc. Natl. Acad. Sci. USA **118**, e2106548118 (2021).

[3] D. Shah and T. Zaman, Detecting sources of computer viruses in networks: Theory and experiment, SIGMETRICS Perform. Eval. Rev. **38**, 203 (2010).

[4] D. Shah and T. Zaman, Rumors in a network: Who's the culprit? IEEE Trans. Inf. Theory **57**, 5163 (2011).

[5] A. Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborová, Inferring the origin of an epidemic with a dynamic message-passing algorithm, Phys. Rev. E **90**, 012801 (2014).

[6] N. Antulov-Fantulin, A. Lančić, T. Šmuc, H. Štefančić, and M. Šikić, Identification of patient zero in static and temporal networks: Robustness and limitations, Phys. Rev. Lett. **114**, 248701 (2015).

[7] I. Biazzo, A. Braunstein, L. Dall'Asta, and F. Mazza, A Bayesian generative neural network framework for epidemic inference problems, Sci. Rep. **12**, 19673 (2022).

[8] A. Braunstein, G. Catania, L. Dall'Asta, M. Mariani, and A. P. Muntoni, Inference in conditioned dynamics through causality restoration, Sci. Rep. **13**, 7350 (2023).

[9] C. Shah, N. Dehmamy, N. Perra, M. Chinazzi, A.-L. Barabási, A. Vespignani, and R. Yu, Finding patient zero: Learning contagion source with graph neural networks, arXiv:2006.11913.

[10] C. W. Tan, P.-D. Yu, S. Chen, and H. V. Poor, DeepTrace: Learning to optimize contact tracing in epidemic networks with graph neural networks, IEEE Trans. Signal Inf. Process. Networks **11**, 97 (2025).

[11] G. Cutura, B. Li, A. Swami, and S. Segarra, Deep demixing: Reconstructing the evolution of epidemics using graph neural networks, in *29th European Signal Processing Conference (EUSIPCO), Dublin, Ireland* (IEEE, Piscataway, NJ, 2021), pp. 2204–2208.

[12] D. Ghio, A. L. M. Aragon, I. Biazzo, and L. Zdeborová, Bayes-optimal inference for spreading processes on random networks, Phys. Rev. E **108**, 044308 (2023).

[13] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Lage-Castellanos, and R. Zecchina, Bayesian inference of epidemics on networks via belief propagation, Phys. Rev. Lett. **112**, 118701 (2014).

[14] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Ingrosso, and R. Zecchina, The patient-zero problem with noisy observations, J. Stat. Mech. (2014) P10016.

[15] A. Braunstein and A. Ingrosso, Inference of causality in epidemics on temporal contact networks, Sci. Rep. **6**, 27538 (2016).

[16] J. Bindi, A. Braunstein, and L. Dall'Asta, Predicting epidemic evolution on contact networks from partial observations, PLoS ONE **12**, e0176376 (2017).

[17] A. P. Muntoni, F. Mazza, A. Braunstein, G. Catania, and L. Dall'Asta, Effectiveness of probabilistic contact tracing in epidemic containment: The role of superspreaders and transmission path reconstruction, PNAS Nexus **3**, pgae377 (2024).

[18] A. Y. Lokhov and D. Saad, Optimal deployment of resources for maximizing impact in spreading processes, Proc. Natl. Acad. Sci. USA **114**, E8138 (2017).

[19] I. Neri and D. Bollé, The cavity approach to parallel dynamics of Ising spins on a graph, J. Stat. Mech. (2009) P08009.

[20] Y. Kanoria and A. Montanari, Majority dynamics on trees and the dynamic cavity method, Ann. Appl. Probab. **21**, 1694 (2011).

[21] A. Y. Lokhov, M. Mézard, and L. Zdeborová, Dynamic message-passing equations for models with unidirectional dynamics, Phys. Rev. E **91**, 012811 (2015).

[22] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Large deviations of cascade processes on graphs, Phys. Rev. E **87**, 062115 (2013).

[23] S. Crotti and A. Braunstein, Matrix product belief propagation for reweighted stochastic dynamics over graphs, Proc. Natl. Acad. Sci. USA **120**, e2307935120 (2023).

[24] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Discrete-time Markov chain approach to contact-based disease spreading in complex networks, Europhys. Lett. **89**, 38009 (2010).

[25] P. Van Mieghem, J. Omic, and R. Kooij, Virus spread in networks, IEEE/ACM Trans. Networking **17**, 1 (2008).

[26] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models*, Interdisciplinary Applied Mathematics (Springer International Publishing, Cham, 2017), Vol. 46.

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

[28] T. Plefka, Convergence condition of the TAP equation for the infinite-ranged Ising spin glass model, J. Phys. A: Math. Gen. **15**, 1971 (1982).

[29] B. Bravi, P. Sollich, and M. Opper, Extended Plefka expansion for stochastic dynamics, J. Phys. A: Math. Theor. **49**, 194003 (2016).

[30] A. Georges and J. S. Yedidia, How to expand around mean-field theory using high-temperature expansions, J. Phys. A: Math. Gen. **24**, 2173 (1991).

[31] A. Maillard, L. Foini, A. L. Castellanos, F. Krzakala, M. Mézard, and L. Zdeborová, High-temperature expansions and message passing algorithms, J. Stat. Mech. (2019) 113301.

[32] M. Tarabolo, SmallCouplingDynamicCavity.jl (2023), https://github.com/Mattiatarabolo/SmallCouplingDynamicCavity.jl.

[33] M. Fabio and indaco biazzo, sibyl-team/epigen: Release epigen (v0.2), Zenodo (2023), https://doi.org/10.5281/zenodo.7852232.

[34] D. J. Watts and S. H. Strogatz, Collective dynamics of "small-world" networks, Nature (London) **393**, 440 (1998).

[35] M. D. Penrose, Connectivity of soft random geometric graphs, Ann. Appl. Probab. **26**, 986 (2016).

[36] R. Hinch, W. J. M. Probert, A. Nurtay, M. Kendall, C. Wymant, M. Hall, K. Lythgoe, A. B. Cruz, L. Zhao, A. Stewart, L. Ferretti, D. Montero, J. Warren, N. Mather, M. Abueg, N. Wu, O. Legat, K. Bentley, T. Mead, K. Van-Vuuren *et al.*, OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing, PLoS Comput. Biol. **17**, e1009146 (2021).

[37] C. C. Kerr, R. M. Stuart, D. Mistry, R. G. Abeysuriya, K. Rosenfeld, G. R. Hart, R. C. Núñez, J. A. Cohen, P. Selvaraj, B. Hagedorn, L. George, M. Jastrzębski, A. S. Izzo, G. Fowler, A. Palmer, D. Delport, N. Scott, S. L. Kelly, C. S. Bennette, B. G. Wagner *et al.*, Covasim: An agent-based model of COVID-19 dynamics and interventions, PLoS Comput. Biol. **17**, e1009149 (2021).

[38] M. Génois and A. Barrat, Can co-location be used as a proxy for face-to-face contacts? EPJ Data Sci. **7**, 11 (2018).

[39] P. Erdős and A. Rényi, On random graphs. I., Publ. Math. Debrecen **6**, 290 (1959).

[40] E. Aurell and H. Mahmoudi, A message-passing scheme for non-equilibrium stationary states, J. Stat. Mech. (2011) P04014.

[41] E. Aurell and H. Mahmoudi, Dynamic mean-field and cavity methods for diluted Ising systems, Phys. Rev. E **85**, 031119 (2012).

[42] G. Del Ferraro and E. Aurell, Dynamic message-passing approach for kinetic spin models with reversible dynamics, Phys. Rev. E **92**, 010102(R) (2015).

[43] E. Aurell, G. Del Ferraro, E. Domínguez, and R. Mulet, Cavity master equation for the continuous time dynamics of discrete-spin models, Phys. Rev. E **95**, 052119 (2017).

[44] E. D. Vázquez, G. Del Ferraro, and F. Ricci-Tersenghi, A simple analytical description of the non-stationary dynamics in Ising spin systems, J. Stat. Mech. (2017) 033303.

[45] E. Ortega, D. Machado, and A. Lage-Castellanos, Dynamics of epidemics from cavity master equations: Susceptible-infectious-susceptible models, Phys. Rev. E **105**, 024308 (2022).

[46] D. J. Daley and D. G. Kendall, Epidemics and rumours, Nature (London) **204**, 1118 (1964).

[47] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).

[48] http://smartdata.polito.it.

[49] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Optimizing spread dynamics on graphs by message passing, J. Stat. Mech. (2013) P09011.