

# Dynamic message-passing equations for models with unidirectional dynamics

Andrey Y. Lokhov,<sup>1,\*</sup> Marc Mézard,<sup>1,2</sup> and Lenka Zdeborová<sup>3</sup>

<sup>1</sup>*Université Paris-Sud/CNRS, LPTMS, UMR8626, Bât. 100, 91405 Orsay, France*

<sup>2</sup>*Ecole normale supérieure - PSL Research University, 45 rue d'Ulm, 75005 Paris, France*

<sup>3</sup>*Institut de Physique Théorique, IPhT, CEA Saclay and CNRS URA 2306, 91191 Gif-sur-Yvette, France*

(Dated: January 15, 2015)

Understanding and quantifying the dynamics of disordered out-of-equilibrium models is an important problem in many branches of science. Using the dynamic cavity method on time trajectories, we construct a general procedure for deriving the dynamic message-passing equations for a large class of models with unidirectional dynamics, which includes the zero-temperature random field Ising model, the susceptible-infected-recovered model, and rumor spreading models. We show that unidirectionality of the dynamics is the key ingredient that makes the problem solvable. These equations are applicable to single instances of the corresponding problems with arbitrary initial conditions, and are asymptotically exact for problems defined on locally tree-like graphs. When applied to real-world networks, they generically provide a good analytic approximation of the real dynamics.

PACS numbers: 02.50.-r,89.20.-a,64.60.aq

## I. INTRODUCTION

Over the past decade, there has been a growing interest in building analytical tools for the study of out-of-equilibrium dynamics in disordered problems defined on heterogeneous networks. A particular attention has been devoted to the study of cascading and avalanche processes, in the cases where the dynamics is not a relaxation dynamics related to a Hamiltonian, but instead is characterized by a set of stochastic transition rules. Examples of such processes include epidemic spreading [1–4], propagation of information and innovations in social media [5–9], dynamics of magnetic and glassy systems [10, 11], communication protocols, such as gossip algorithms and peer-to-peer file sharing on computer networks [12, 13], activation cascades in biological and neural networks [14, 15], and news updates in financial markets [16, 17]. A common property shared by these processes is the unidirectional nature of the corresponding dynamics: once an elementary constituent of the system under the influence of its neighbors undergoes a transition to a certain state, it can never return to the previous one.

Although the properties of diluted disordered systems have been intensively investigated over the past several years, there is still no well-established tractable method for solving the corresponding dynamics in the general case. One category of problems that has recently attracted a lot of attention is the case of out-of-equilibrium dynamic processes on sparse graphs [18–21]. Methods which are developed in this context can also be used as so-

phisticated mean-field-type approximations for problems defined on general graphs. The generating functional analysis techniques [22], the dynamical replica analysis [23, 24] and the cavity method [25, 26] have been recently used for the construction of a general approach in terms of time-trajectories of variables. However, the general dynamics remains intractable in this formalism except for only a few time steps: the solution of the corresponding equations takes in general a number of operations that grows exponentially with the duration of the process one wants to study. In a few special cases, some progress has been recently made by a number of authors who were able to write, using cavity-like arguments, tractable asymptotically-exact mean-field dynamic equations for several models defined on locally tree-like graphs, such as the zero-temperature random field Ising model (RFIM) [11], the susceptible-infected-recovered (SIR) model [27–31], and the threshold models [32, 33]. All these models share a common property: they describe a unidirectional dynamics involving one transition to the active state; the derivation of the corresponding equations is typically based on identifying correct dynamic variables that are required to obtain the closed-form expressions. These examples lead to the hypothesis that the microscopic irreversibility of the dynamics is a key property that makes it possible to derive such equations [30, 34]. However, in general it is very difficult to guess the right dynamic variables that should be used in the dynamic equations for more complicated models, involving a larger number of states and several non-trivial transitions. Probably, the simplest model of this kind is the so-called rumor spreading model [35–37], which is a three-state dynamic model with two neighbors-dependent transitions.

In this paper, we develop a systematic procedure for deriving the dynamic message-passing (DMP) equations for general models with unidirectional dynamics and ar-

---

\*Electronic address: andrey.lokhov@lptms.u-psud.fr

bitrary number of states. They allow one to estimate the marginal probabilities of each variable at each time on a given network of contacts, using a number of operations that is polynomial both in the size of the network and in the duration of the dynamic process. These equations are applicable to single-instance problems with arbitrary initial conditions, they are asymptotically exact on locally tree-like networks, and typically provide a good approximation for real-world networks. The DMP equations are derived using the cavity method, also known in different fields as the belief propagation (BP), or the sum-product algorithm [26, 38], starting from a BP equation that takes as variables the time trajectories of nodes. Despite the similarity with the BP equations that need to be iterated until convergence, the iteration time in the DMP equations corresponds to the physical time. We show that the unidirectional nature of the dynamics is indeed a crucial element that makes the problem solvable. More precisely, the time trajectories in these models can be fully parametrized with only a few flipping times, leading to a significant simplification of the corresponding dynamic BP formulation. As a result, these BP equations can be rewritten in terms of closed-form DMP equations with a computational complexity which turns out to be reduced from an exponential in the duration of the process to a polynomial. This simplification occurs thanks to the use of dynamic variables that appear naturally to be the weighted sums of messages of the BP equations on trajectories.

The structure of the paper is as follows. In section II we present a systematic framework for treating dynamic problems on locally tree-like graphs, writing a general dynamic belief propagation equation for node trajectories in time. Then, in section III we introduce an important class of considered models with unidirectional dynamics. In the next sections IV-VI we construct a general procedure for the derivation of DMP equations for these models, thus obtaining the single-instance form of the mean-field equations that for some cases have already appeared in the literature, and obtaining new equations for more complicated models. Finally, we provide supporting numerical results and discuss the possible applications of our approach.

## II. DYNAMIC BELIEF PROPAGATION

Belief propagation, or the cavity method, is an iterative method that allows one to estimate efficiently marginal probability distributions in graphical models. It has been proven to be very successful in some applications, e.g. error-correcting codes [39], Bayesian networks [40], and optimization problems [41]. BP makes use of the assumption that the marginal probabilities (called messages) defined on an auxiliary cavity graph (a graph with a removed node) are uncorrelated. This assumption is obviously exact if the underlying network is a tree, in

other cases it is an approximation of the mean-field type (for more details, see [26]). The solution to the BP equations can often be obtained by iterating the equations until convergence. A formulation of this algorithm for static problems is given for consistency in Appendix A.

Our motivation to seek a generalization of the cavity method for dynamic problems has been inspired by the success of belief propagation in static problems. The main idea behind the dynamic belief propagation is to write usual cavity equations using the time trajectories of nodes as variables. This idea has been exploited in a number of previous works on the dynamics of disordered systems [18–21].

Consider a graph  $G = (V, E)$ , defined by a vertex set  $V$  and a set of edges  $E$ . In the dynamic setting, each vertex  $i \in V$  is characterized by a variable, taking the value  $\sigma_i^t$  at time  $t$ . We assume that the set of possible values of  $\sigma_i^t$  is of size  $K$ . We consider a generic dynamic process defined in a discrete-time parallel dynamics and described by a local transition probability  $w_i(\sigma_i^{t+1} | \{\sigma_j^t\}_{j \in \partial i})$  that a node  $i$  takes value  $\sigma_i^{t+1}$  at time  $t+1$  given the values  $\{\sigma_j^t\}$  of its neighbors at time  $t$ . If we denote by  $\vec{\sigma}_i = (\sigma_i^0, \dots, \sigma_i^T)$  the trajectory of variable  $i$  at times  $t = 0, \dots, T$ , where  $T$  is the stopping time, the joint probability distribution of the trajectories  $P(\{\vec{\sigma}_i\}_{i \in V})$  can be written as follows:

$$P(\{\vec{\sigma}_i\}_{i \in V}) = \prod_{i \in V} \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_j^t\}_{j \in \partial i}) P_0, \quad (1)$$

where  $P_0 \equiv P(\{\sigma_i^0\}_{i \in V})$  is the distribution of variables at initial time.

It is a well-known fact that BP equations are exact for static graphical models when the factor graph is a tree. However, when we consider the factor graph (graph involving check nodes that represent local interactions between variables) of the model defined in (1), in which the variables are time trajectories  $\vec{\sigma}_i$ , it turns out that the factor graph contains many loops, even in the case where  $G$  is a tree, see Fig. 1.

A way to fix this problem consists in exploiting the duality between variables and interactions by putting the variables on the edges. To this purpose we introduce a different representation of the problem, that uses auxiliary variables (time-trajectories)  $\vec{\sigma}_{i \rightarrow j}$  on each directed edge  $(i, j) \in E$ . For a given  $i$ , all the variables  $\vec{\sigma}_{i \rightarrow j}$  are supposed to be copies of the original  $\vec{\sigma}_i$ . They should thus be all equal, and we implement this by adding for each  $i$  an additional constraint  $\vec{\sigma}_{i \rightarrow j} = \vec{\sigma}_{i \rightarrow k}$  for all  $j, k \in \partial i$ . The joint probability distribution (1) of time trajectories can hence be written in terms of these new variables:

$$\begin{aligned} & P(\{\vec{\sigma}_{i \rightarrow j}, \vec{\sigma}_{j \rightarrow i}\}_{(i,j) \in E}) \\ &= \prod_{i \in V} \prod_{t=0}^{T-1} \left[ w_i(\sigma_{i \rightarrow l}^{t+1} | \{\sigma_{k \rightarrow i}^t\}_{k \in \partial i}) \prod_{k \in \partial i \setminus l} \delta_{\sigma_{i \rightarrow l}^t, \sigma_{i \rightarrow k}^t} \right] P_0, \end{aligned} \quad (2)$$

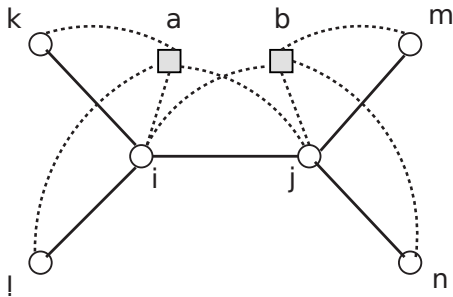


FIG. 1: An example of a factor graph of the graphical model at two nearest times described by the joint probability distribution  $P(\{\vec{\sigma}_i\}_{i \in V})$ . The check node  $a$  represents interaction between the variable  $\sigma_i^{t+1}$  and the variables  $\{\sigma_j^t\}_{j \in \partial i}$  at a previous time step. This factor graph is characterized by systematic short loops.

where  $l$  is any of the variables influenced directly by  $i$ , and  $k \in \partial i \setminus l$  means the set of nodes neighboring node  $i$ , excluding  $l$ . This new form of the probability distribution is described by a factor graph which is very closely related to  $G$ : the new variables  $\vec{\sigma}_{i \rightarrow j}, \vec{\sigma}_{j \rightarrow i}$  live on each edge  $(i, j) \in E$ , and there is a function node (interaction) associated with every vertex  $i \in V$ . If the original graph  $G$  is a tree (respectively, is locally tree-like), the factor graph is a tree (respectively, is locally tree like), see Fig. 2. This crucial property allows to use the BP method in terms of this new description for studying the dynamics, with the guarantee that the resulting equations are exact if  $G$  is a tree.

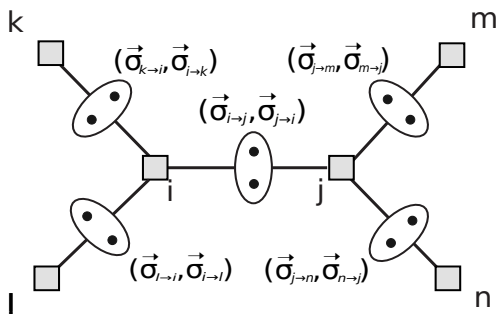


FIG. 2: An example of a factor graph of the graphical model at all times described by the joint probability distribution  $P(\{\vec{\sigma}_{i \rightarrow j}, \vec{\sigma}_{j \rightarrow i}\}_{(i, j) \in E})$ . The check node  $i$  represents interaction between trajectories  $\vec{\sigma}_i$  and  $\{\vec{\sigma}_j\}_{j \in \partial i}$ . This factor graph is characterized by the underlying tree structure if the original graph is a tree.

Let us now write the BP equations. Using the fact that  $\vec{\sigma}_{i \rightarrow j} = \vec{\sigma}_{i \rightarrow k}$  for all  $j, k \in \partial i$ , it is convenient to rename the variables  $\{\vec{\sigma}_{i \rightarrow j}, \vec{\sigma}_{j \rightarrow i}\}_{(i, j) \in E}$  to  $\{\vec{\sigma}_i, \vec{\sigma}_j\}_{(i, j) \in E}$ . The BP equations for the joint probability distribution (2) in

terms of conditional messages  $m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$  read:

$$m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j) = \sum_{\{\vec{\sigma}_k\}_{k \in \partial i \setminus j}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial i \setminus j}, \sigma_j^t) \right] \times P(\{\sigma_i^0\}_{i \in V}) \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\vec{\sigma}_k | \vec{\sigma}_i). \quad (3)$$

Notice that, in general, there appears a normalization constant in front of the BP equation. In our case, thanks to the Markov property of the dynamics, we can explicitly compute the normalization constant, see Appendix B for details.

The message  $m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$  has the meaning of the probability for the trajectory  $\vec{\sigma}_i$  given the trajectory  $\vec{\sigma}_j$  in the transformed cavity graph, where the factor node  $j$  has been removed. We denote the dynamics in the corresponding cavity graph as  $D_j$ . The equation (3) can be iterated until convergence, and the corresponding marginal probability of a time trajectory  $\vec{\sigma}_i$  will be given by

$$m^i(\vec{\sigma}_i) = \sum_{\{\vec{\sigma}_k\}_{k \in \partial i}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial i}) \right] \times P(\{\sigma_i^0\}_{i \in V}) \prod_{k \in \partial i} m^{k \rightarrow i}(\vec{\sigma}_k | \vec{\sigma}_i). \quad (4)$$

Note that in the general case, it takes an exponential number of operations in the duration of the process to solve the equations (3) and (4), since each message has  $K^T$  components, where  $K$  is the number of values that each variable  $\sigma_i^t$  may take, and the sum in (3) is performed over  $K^{T(d_i-1)}$  variables for each node  $i$ , with  $d_i$  being the number of neighbors of  $i$ . However, a crucial simplification occurs for the models with unidirectional dynamics, introduced in the next section.

### III. MODELS WITH UNIDIRECTIONAL DYNAMICS

Let us assume that in the expression for the transition probability  $w_i(\sigma_i^{t+1} | \{\sigma_j^t\}_{j \in \partial i})$ , the value  $\sigma_i^t$  takes one of the  $K$  ordered discrete values that we denote  $\Omega_1, \Omega_2, \dots, \Omega_K$ . We call the dynamic process unidirectional if the node can change its state only in a directed and irreversible way,  $\Omega_1 \rightarrow \Omega_2 \rightarrow \dots \rightarrow \Omega_K$ , and the transition to one of the previous states is forbidden by the dynamic rules.

Among unidirectional processes with  $K = 2$  states ( $\sigma_i^t$  can take one of the two values  $-1 \equiv \downarrow$  or  $1 \equiv \uparrow$ ), one can mention the zero-temperature random field Ising model (zero-temperature RFIM) with homogeneous initial condition, considered in [11]:  $\prod_{i \in V} \delta_{\sigma_i^0, -1} = (\downarrow \downarrow \dots \downarrow)$ . Each spin can flip only if the local field created by its neighbors is positive (precise definitions will be given in the next

section). Once being flipped, the spins in this avalanche dynamics remain in the position  $\uparrow$  for all times, since the local field is a monotone non-decreasing function of time. Therefore, this system has a unidirectional dynamics with two ordered states,  $\downarrow$  and  $\uparrow$ . The model can be generalized to any initial condition if the backward transition from  $\uparrow$  to  $\downarrow$  is explicitly forbidden by the dynamic rules. The situation is different from the standard Glauber dynamics of Ising model with non-zero temperature [42], or the majority dynamics of voters, that switch to one of the alternative opinions according to the majority of their neighbors [21], where each variable is free to flip an arbitrary number of times. Note that the linear threshold model with random thresholds studied in [32, 34] is equivalent to this formulation of the RFIM.

Another example of model with two states and unidirectional dynamics is given by the susceptible-infected (SI) model, in which the node can be in either of two states: susceptible ( $S$ ), or infected ( $I$ ). The propagation of infection on this model occurs due to the pairwise interactions between individuals: for instance, the  $S$  individual can be infected by one of its  $I$  neighbors at each time step, and then remains infected forever. If there exists a recovery mechanism that allows an infected individual to become susceptible again after some time (it corresponds to the so-called susceptible-infected-susceptible, or SIS model, that is used to model the behavior of endemic diseases), the model does not belong to the class of unidirectional models anymore.

Unidirectional dynamic processes with  $K = 3$  states include the susceptible-infected-recovered (SIR) model, an extension of the SI model that is widely used to model epidemic spreading. In this model, the infected node can switch to a recovered ( $R$ ) state with a certain probability at each time step, leading to a depletion of infected agents. Another well known model with unidirectional dynamics and three states is given by the so-called *rumor spreading*, or ignorant-spreader-stifler (ISS) model [4], which describes the propagation of information by spreaders to ignorants that are unaware of rumor, and takes into account the possibility that the spreader can become uninterested in the rumor under the influence of its neighbors. The precise formulation of these models will be given in the section V.

In what follows, we discuss the dynamic message-passing equations for the models listed above, illustrating the general method to derive such equations for other models with arbitrary  $K$ . Typically, these equations would allow us to answer the following question: what is the probability that a certain node  $i$  is in a certain state  $\Omega_a$  at time  $t$ ? For some of these models, the equivalent equations have already appeared in the literature, however, in a form averaged over the ensemble of random graphs and/or over the initial conditions, and not suitable to the algorithmic purposes for single-instance problems. For others, the DMP equations have never been stated previously; the “naive” mean-field equations

that exist for all models are derived under assumptions of homogeneity of transmission probabilities and complete-mixing hypothesis (ignoring the actual topology of the interaction network and assuming that each pair of nodes may interact) that are obviously unrealistic. For each model considered in the following, we will discuss the relation of our DMP equations to those existing in the literature, if any.

#### IV. UNIDIRECTIONAL MODELS WITH $K = 2$ STATES

##### A. Zero-temperature RFIM

The RFIM Hamiltonian reads

$$H = - \sum_{(ij)} J_{ij} \sigma_i \sigma_j - \sum_i (h + h_i) \sigma_i, \quad (5)$$

where  $J_{ij}$  is a non-negative interaction between spins  $i$  and  $j$ ,  $h$  is an external uniform magnetic field, and  $h_i$  is a random magnetic field on site  $i$ , extracted from some probability distribution. At zero temperature spin  $i$  tends to be aligned with its local magnetic field

$$\Delta_i^t = h + h_i + \sum_{k \in \partial i} J_{ki} \sigma_k^t. \quad (6)$$

Consider an initial condition in the form  $P(\{\sigma_i^0\}_{i \in V}) = \prod_{i \in V} \delta_{\sigma_i^0, -1}$  (as mentioned in the previous section, one could choose any initial condition, provided the dynamics is such that the transition from  $\uparrow$  to  $\downarrow$  is forbidden). Define the zero temperature stochastic dynamics respecting the following property: spin  $\sigma_i = -1$  with a positive local field  $\Delta_i$  flips with rate  $1/\tau$ , and does not flip otherwise. Each spin flips at most only once, so the trajectory  $\vec{\sigma}_i$  has a typical form  $|\downarrow_0 \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \downarrow \uparrow_{\tau_i} \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow_{\tau_i}\rangle$  and is in the one-to-one correspondence with the flipping time  $\tau_i$  ( $\tau_i$  is the first time for which  $\sigma_i = 1$ ). If the spin does not flip for all the times  $0, \dots, T-1$ , then by definition we set  $\tau_i = T$  ( $T$  is the stopping time, i.e. the condition  $\tau_i = T$  summarizes all the events that happen after the time  $T$ ).

Using the representation in terms of flipping times, Eq. (3) for  $\tau_i < T$  can be expressed as follows:

$$m^{i \rightarrow j}(\tau_i | \tau_j) = \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} W_{RFIM} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i), \quad (7)$$

where

$$W_{RFIM} = \prod_{t'=0}^{\tau_i-2} \left( 1 - \frac{1}{\tau} \mathbb{1}[\Delta_i^{t'} > 0] \right) \frac{1}{\tau} \mathbb{1}[\Delta_i^{\tau_i-1} > 0]. \quad (8)$$

Here and in what follows we use a convention

$$\prod_{t=a}^{a-\epsilon} (\dots) \equiv 1 \quad (9)$$

for any fixed  $a$  and  $\epsilon > 0$ . Using the fact that messages are properly normalized, we choose

$$m^{i \rightarrow j}(T | \tau_j) = 1 - \sum_{\tau_i=1}^{T-1} m^{i \rightarrow j}(\tau_i | \tau_j). \quad (10)$$

From the conditional messages  $m^{i \rightarrow j}(\tau_i | \tau_j)$ , we can define two quantities:

$$p^{i \rightarrow j}(t) = \sum_{\tau_i > t} m^{i \rightarrow j}(\tau_i | T), \quad (11)$$

$$q^{i \rightarrow j}(t) = \sum_{\tau_i \leq t} m^{i \rightarrow j}(\tau_i | T). \quad (12)$$

These quantities characterize the marginals of the zero-temperature RFIM in the cavity dynamics  $\mathcal{D}_j$ , in which  $\sigma_j^t = -1$  for every  $t$  and never flips, even if  $\Delta_j > 0$ . In this dynamics,  $p^{i \rightarrow j}(t)$  is the probability that spin  $i$  stays in the state  $\sigma_i^t = -1$  at time  $t$ , and  $q^{i \rightarrow j}(t)$  is defined to be the probability that spin  $i$  has already flipped, and hence  $\sigma_i^t = 1$ .

Of course, we will be ultimately interested in writing a closed equation for the marginals in the original dynamics, defined as:

$$p^i(t) = \sum_{\tau_i > t} m^i(\tau_i), \quad (13)$$

$$q^i(t) = 1 - p^i(t), \quad (14)$$

where  $m^i(\tau_i)$  are the marginal probabilities of trajectories (4) in the original dynamics, following the same equation as (7), but with  $\partial i \setminus j$  replaced by  $\partial i$ , in the same way as (4) is related to (3). The DMP equations for this model can be derived, starting from the equation (7), and using the definitions (11), (12), as well as elementary properties of the messages, such as normalization and causality constraints. For details of the derivation see Appendix C. After some algebra, the resulting DMP equations can be shown to take the following form in discretized time notations:

$$\begin{aligned} q^{i \rightarrow j}(t+1) &= \left(1 - \frac{1}{\tau}\right) q^{i \rightarrow j}(t) \\ &+ \frac{1}{\tau} \sum_{\{\sigma_k\}_{k \in \partial i \setminus j}} \mathbb{1} \left[ h + h_i + \sum_{k \in \partial i \setminus j} J_{ki} \sigma_k - J_{ji} > 0 \right] \\ &\times \prod_{k \in \partial i \setminus j: \sigma_k = +1} q^{k \rightarrow i}(t) \prod_{k \in \partial i \setminus j: \sigma_k = -1} [1 - q^{k \rightarrow i}(t)]. \quad (15) \end{aligned}$$

Therefore, the marginal probability for spin  $i$  to be in the state  $+1$  at time  $t+1$  is given by  $q^i(t+1)$ , which can be

computed according to the following expression:

$$\begin{aligned} q^i(t+1) &= \left(1 - \frac{1}{\tau}\right) q^i(t) \\ &+ \frac{1}{\tau} \sum_{\{\sigma_k\}_{k \in \partial i}} \mathbb{1} \left[ h + h_i + \sum_{k \in \partial i} J_{ki} \sigma_k > 0 \right] \\ &\times \prod_{k \in \partial i: \sigma_k = +1} q^{k \rightarrow i}(t) \prod_{k \in \partial i: \sigma_k = -1} [1 - q^{k \rightarrow i}(t)]. \quad (16) \end{aligned}$$

The probability that spin  $i$  is still in the state  $-1$  at time  $t+1$  is then given by  $p^i(t+1) = 1 - q^i(t+1)$ . Note that the DMP equations (15) and (16) can now be run in the real time, starting with initial conditions  $q^i(0) = q^{i \rightarrow j}(0) = 0$  for each node  $i$  and  $j$ ; these equations have a closed self-consistent form, so we no longer need to compute the messages using (7). Note that the computational complexity of the DMP equations for the zero-temperature RFIM is reduced from exponential to linear in time; in the most straightforward implementation, the computation complexity of (15) and (16) is  $\mathcal{O}(N2^c t)$ , where  $c$  is the average degree of the graph.

The averaged form of the DMP equations has been first derived in [11] using a cavity-like argument for the dynamic variables  $q^{i \rightarrow j}(t)$  and  $q^i(t)$ . The derivation, which is close to ours, have been provided in [32], where an equivalent linear threshold model has been investigated. This model has been also studied in a different setting in a form of the voter model in [21].

## B. Generalized SI model

Let us now consider the most general case of a unidirectional dynamic model with two states and pairwise interactions between nodes; each of these independent interactions may lead to a transition to the final state. The definition of the generalized SI model in discrete time can be represented as follows:

$$S(i) + S(j) \xrightarrow{\epsilon_{ji}} I(i) + S(j), \quad (17)$$

$$S(i) + I(j) \xrightarrow{\lambda_{ji}} I(i) + I(j), \quad (18)$$

$$S(i) \xrightarrow{\nu_i} I(i). \quad (19)$$

This diagram represents the dynamic rules at each time step. Here,  $i$  and  $j$  mean two neighboring nodes in the network, and  $\epsilon_{ji}$ ,  $\lambda_{ji}$  and  $\nu_i$  correspond to transition probabilities at each time step.

Again, since there are only two possible states and the dynamics is unidirectional, the time trajectory of a node  $i$   $|S_0 SSSSSSI_{\tau_i} IIIIIIT\rangle$  can be parametrized by a single time  $\tau_i$ , when the spin flips from the state  $S$  to the state  $I$  ( $\tau_i$  is the first time for which  $\sigma_i^{\tau_i} = I$ ). If the node  $i$  is initially in the state  $I$ , we set  $\tau_i = 0$ , and we put

by definition  $\tau_i = T$  if the flipping happens after the observation time  $T$ , or never happens.

Therefore, the dynamic cavity equation (3) takes the following form for the generalized SI model:

$$m^{i \rightarrow j}(\tau_i | \tau_j) = \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} W_{SI} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i), \quad (20)$$

where  $W_{SI}$  is the kernel that resumes the dynamics of the model up to the final time  $T$ , and  $\tau_i < T$ . See Appendix D for details.

The beliefs  $m^i(\tau_i)$ , which can be obtained from (20) in the same way as (4) is obtained from (3), allow one to define the marginal probabilities describing the dynamics of the SI model:

$$P_S^i(t) = \sum_{\tau_i > t} m^i(\tau_i), \quad (21)$$

$$P_I^i(t) = 1 - P_S^i(t). \quad (22)$$

It is also useful to define the marginal probability that node  $i$  is in the state  $S$  at a given time in the cavity graph  $D_j$ , in which the node  $j$  is fixed to the state  $S$  for all times:

$$P_S^{i \rightarrow j}(t) = \sum_{\tau_i > t} m^{i \rightarrow j}(\tau_i | T). \quad (23)$$

After some algebra (see Appendix D for details of derivation), the DMP equations for the generalized SI model take the following form:

$$P_S^{k \rightarrow i}(t) = P_S^k(0)(1 - \epsilon_{ik})^t (1 - \nu_k)^t \prod_{l \in \partial k \setminus i} \theta^{l \rightarrow k}(t), \quad (24)$$

$$\begin{aligned} \theta^{k \rightarrow i}(t) = & \theta^{k \rightarrow i}(t-1) - \epsilon_{ki} \phi_1^{k \rightarrow i}(t-1) \\ & - \lambda_{ki} \phi_2^{k \rightarrow i}(t-1), \end{aligned} \quad (25)$$

$$\begin{aligned} \phi_1^{k \rightarrow i}(t) = & (1 - \epsilon_{ki}) \phi_1^{k \rightarrow i}(t-1) \\ & - (1 - \epsilon_{ki})^t (P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t)), \end{aligned} \quad (26)$$

$$\begin{aligned} \phi_2^{k \rightarrow i}(t) = & (1 - \lambda_{ki}) \phi_2^{k \rightarrow i}(t-1) \\ & + (1 - \epsilon_{ki})^t (P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t)). \end{aligned} \quad (27)$$

The initial conditions are

$$\theta^{i \rightarrow j}(0) = 1, \quad (28)$$

$$\phi_1^{i \rightarrow j}(0) = \delta_{\sigma_i^0, S} = P_S^i(0), \quad (29)$$

$$\phi_2^{i \rightarrow j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0) = 1 - P_S^i(0). \quad (30)$$

As follows from their mathematical definitions, the introduced dynamic variables can be given the following physical interpretations (that follow from their explicit mathematical form):

- $\theta^{k \rightarrow i}(t)$  is the probability that neither of both  $\epsilon$  and  $\lambda$  infection signals has been passed from node  $k$  to node  $i$  up to time  $t$  in the cavity dynamics  $D_i$ ;
- $\phi_1^{k \rightarrow i}(t)$  is the probability that the  $\epsilon$  infection signal has not been passed from node  $k$  to node  $i$  up to time  $t$  in the cavity dynamics  $D_i$  and that  $k$  is in the state  $S$  at time  $t$ ;
- $\phi_2^{k \rightarrow i}(t)$  is the probability that neither of both  $\epsilon$  and  $\lambda$  infection signals has been passed from node  $k$  to node  $i$  up to time  $t$  in the cavity dynamics  $D_i$  and that  $k$  is in the state  $I$  at time  $t$ ;
- $P_S^{k \rightarrow i}(t)$  is the probability that  $k$  is in the state  $S$  at time  $t$  in the cavity dynamics  $D_i$ .

Finally, the marginal probabilities for nodes to be in states  $S$  or  $I$  at time  $t$  are computed via

$$P_S^i(t) = P_S^i(0)(1 - \nu_i)^t \prod_{k \in \partial i} \theta^{k \rightarrow i}(t), \quad (31)$$

$$P_I^i(t) = 1 - P_S^i(t). \quad (32)$$

The exactness of DMP equations for generalized SI model on tree graphs is illustrated in the Fig. 3. Their computational complexity is  $\mathcal{O}(Nct)$ , where  $c$  is the average degree of the graph. Note that the complexity is linear in average degree of the network; this is due to the pairwise nature of the interactions, leading to a factorization of dynamic variables in (24) and (31). The same property will also hold for other pairwise models, considered below. These dynamic equations have never appeared in the literature. However, in some sense this model is a straightforward generalization of the SIR model that is considered further, see the next section.

## V. UNIDIRECTIONAL MODELS WITH $K = 3$ STATES

### A. SIR model

The dynamics of the SIR model is defined in discrete time by infection and recovery probabilities,  $\lambda_{ij}$  and  $\mu_i$ . At each time step, the following dynamics rules apply [4]:

$$S(i) + I(j) \xrightarrow{\lambda_{ji}} I(i) + I(j), \quad (33)$$

$$I(i) \xrightarrow{\mu_i} R(i). \quad (34)$$

Note that the SIR model, defined by equations (33) and (34), represents in some sense a particular case of the generalized SI model, with  $\epsilon_{ij} = 0$  and  $\nu_i = 0$  for all  $i$  and  $j$ . At the same time, a trivial (independent on the state of neighbors) transition to the  $R$  state with probability  $\mu_i$  is added. Now the time trajectory for a node  $i$  can be fully parametrized by two flipping

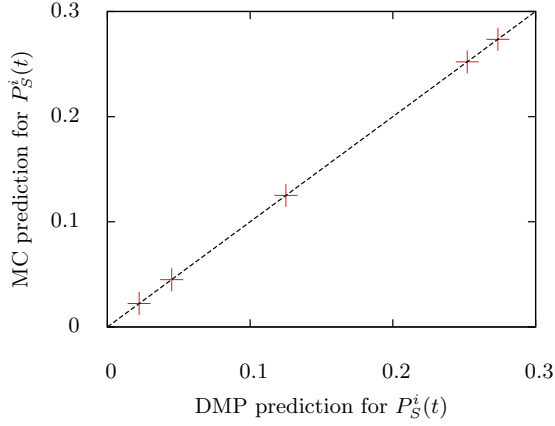


FIG. 3: (Color online) Comparison of prediction of the DMP equations for the generalized SI model with the Monte Carlo simulations. Marginal probabilities  $P_S^i(t)$  are presented for five nodes from a tree graph with  $N = 20$  nodes and  $t = 5$ , the parameters of the model are  $\lambda = 0.5$ ,  $\epsilon = 0.1$  and  $\nu = 0.1$ , there is one infected node at initial time. The MC average is performed over  $10^7$  instances. The error bars are negligible and are not shown.

times:  $\vec{\sigma}_i = |S_0 SSSSSSI_{\tau_i} IIIIIIR_{\omega_i} RRRRRR_T\rangle \leftrightarrow (\tau_i, \omega_i)$ . This leads to the following equations, already derived in a different way (with a correct choice of dynamic variables that have to be used) in [31]:

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0) \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \quad (35)$$

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) = -\lambda_{ki} \phi^{k \rightarrow i}(t), \quad (36)$$

$$\begin{aligned} \phi^{k \rightarrow i}(t) &= (1 - \lambda_{ki})(1 - \mu_k) \phi^{k \rightarrow i}(t-1) \\ &\quad - [P_S^{k \rightarrow i}(t) - P_S^{k \rightarrow i}(t-1)]. \end{aligned} \quad (37)$$

The initial conditions are given by  $\theta^{k \rightarrow i}(0) = 1$ , and  $\phi^{k \rightarrow i}(0) = \delta_{\sigma_k^0, I}$ . Similarly to the SI model, a concrete physical sense may be given to the dynamic variables  $P_S^{k \rightarrow i}(t)$ ,  $\theta^{k \rightarrow i}(t)$  and  $\phi^{k \rightarrow i}(t)$  (see [31] for details). The marginal probabilities that node  $i$  is in a given state at time  $t$  are then given as

$$P_S^i(t+1) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \rightarrow i}(t+1), \quad (38)$$

$$P_R^i(t+1) = P_R^i(t) + \mu_i P_I^i(t), \quad (39)$$

$$P_I^i(t+1) = 1 - P_S^i(t+1) - P_R^i(t+1). \quad (40)$$

The computational complexity of DMP equations for SIR model is  $\mathcal{O}(Nct)$ . The details of the derivation of DMP equations for the SIR model using dynamic belief propagation are given in the Appendix E. A numerical study of these equations has been provided in [31] for different type of networks.

Equations reminiscent of (38-40) were first derived in [27] for a more general SIR model with non-exponential transmission and recovery distributions. For this more

general case, no easily tractable Markovian form of the DMP is known: the equations in [27] are presented in a convolutional form that is complicated for numerical resolution. For a simpler case of constant recovery and transmission rates, the equations of [27] simplify. For an ensemble of random graphs with a given probability distribution it is possible to write the mean-field equations on the fraction of nodes in the states  $S$ ,  $I$  and  $R$  [28–30]. These equations are exact in the ensemble of diluted random graphs in the thermodynamic limit,  $N \rightarrow \infty$  and differ markedly from the naive mean field equations [4] that completely ignore the topology of the network and therefore provide only a crude approximation to the dynamics. The recent work [33] presented a generalization of the SIR model to the threshold models where a transition happens only if the information is received from a certain number  $A$  of neighbors. This model can also be readily solved within the DMP approach. Indeed, the expression (38) for the marginal probability that node  $i$  is in the state  $S$  would take a form similar to the second term in the right-hand side of the equation (16) in the RFIM, with  $q^{k \rightarrow i}(t)$  replaced by  $\theta^{k \rightarrow i}(t+1)$ : one would need to sum over all the subsets of  $\partial i$  that correspond to the transmission of information by at least  $A$  neighbors. Hence, this model represents a three-state model with a RFIM-like non-trivial transition to the infected state.

## B. Rumor spreading model

The definition of the rumor spreading model can be summarized as follows [4]. For the sake of simplicity, we keep the same notations for the states as those used in the SIR model. Each node  $i \in V$  at discrete time  $t$  can be in one of three states  $\sigma_i^t$ : ignorant,  $\sigma_i^t = S$ , spreader,  $\sigma_i^t = I$ , or stifter,  $\sigma_i^t = R$ . At each time step, an “infected” node  $i$  will recover with probability  $1 - \prod_{k \in \partial i} (1 - \alpha_{ki} \delta_{\sigma_k^t, I})$ , and a “susceptible” node  $i$  will become infected with probability  $1 - \prod_{k \in \partial i} (1 - \lambda_{ki} \delta_{\sigma_k^t, I})$ , where  $\partial i$  is the set of neighbors of node  $i$ . The recovered nodes never change their state. These rules can be summarized by the following scheme:

$$S(i) + I(j) \xrightarrow{\lambda_{ji}} I(i) + I(j), \quad (41)$$

$$I(i) + I(j) \xrightarrow{\alpha_{ji}} R(i) + I(j). \quad (42)$$

The interpretation of this model is as follows: a spreader node can either inform one of its ignorant neighbors on the rumor, in which case they start to communicate the rumor to their neighbors, or become uninterested in the rumor and turn to the  $R$  state if the rumor loses its “news value”. This happens in a directed way when the spreader gets in contact with another spreader. Note that some rumor spreading models include an additional modeling of such a spreading decay, described by a contact of a spreader with a stifter with the same probability  $\alpha$ . This additional transition can be easily included in our

approach, but for the purpose of this paper we stick to this “minimal” version of the ISS model that captures the main features of the rumor spreading process and its difference with respect to epidemiological spreading models, such as the SIR model considered before.

As in the previous cases, the irreversibility of dynamics of the rumor spreading model makes it possible to parametrize the time trajectory of a node  $i$  by only two flipping times,  $\vec{\sigma}_i = (\tau_i, \omega_i)$ :  $\tau_i$ , indicating a transition from  $S$  to  $I$  (the first time to be in the state  $I$ ) and  $\omega_i$ , corresponding to a  $I$  to  $R$  transition (the first time in the state  $R$ ). If the node  $i$  is initially in the state  $I$ , we set  $\tau_i = 0$ , and we put by definition  $\tau_i = T$  if the flipping happens after the termination time  $T$ , or never happens.

The rumor spreading model, defined via the transition rules (41-42), is notably more complicated than the SIR model because it has two non-trivial transitions, dependent on the state of neighbors. As we will see, it is not easy to obtain the corresponding DMP equations for this model by guessing the correct dynamic variables since the computation of the very DBP messages is required. On the other hand, they appear automatically in the dynamic cavity approach. Let us first state the DMP computational scheme for this model.

The marginal probabilities  $P_S^i(t+1)$ ,  $P_I^i(t+1)$  and  $P_R^i(t+1)$  that node  $i$  is in a state  $S$ ,  $I$  and  $R$  respectively at time  $t$  are given by the following equations that can be iterated in time starting from initial conditions at time  $t = 0$ :

$$P_S^i(t+1) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \rightarrow i}(t+1), \quad (43)$$

$$P_R^i(t+1) = P_R^i(t) + \sum_{\tau_i \leq t} m^i(\tau_i, t+1), \quad (44)$$

$$P_I^i(t+1) = 1 - P_S^i(t+1) - P_R^i(t+1), \quad (45)$$

where  $m^i(\tau_i, t+1)$  has a physical meaning of the marginal probability that the node  $i$  has switched to the state  $I$  at time  $\tau_i$  and to the state  $R$  at time  $t+1$ . The remaining computational scheme serves to compute this probabilities explicitly. To this purpose, we introduce a number of auxiliary dynamic messages that can be computed iteratively. Again, these messages may be given a physical interpretation, and are defined in the corresponding cavity dynamics. As an illustration, consider the message  $P_S^{i \rightarrow j}(t)$ , defined as the probability for node  $i$  to be in the state  $S$  at time  $t$  in the cavity graph, in which all the connections of the node  $j$ , except to  $i$ , has been removed. It is updated as follows:

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0) \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \quad (46)$$

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) = -\lambda_{ki} \phi^{k \rightarrow i}(t), \quad (47)$$

$$\begin{aligned} \phi^{k \rightarrow i}(t) &= (1 - \lambda_{ki}) \phi^{k \rightarrow i}(t-1) + P_S^{k \rightarrow i}(t-1) \\ &- P_S^{k \rightarrow i}(t) - \sum_{\tau_k \leq t-1} (1 - \lambda_{ki})^{t-\tau_k} m^{k \rightarrow i}(\tau_k, t | T, T). \end{aligned} \quad (48)$$

In these equations, the dynamic messages  $\theta^{k \rightarrow i}(t+1)$ ,  $\phi^{k \rightarrow i}(t)$  and  $m^{k \rightarrow i}(\tau_k, t | T, T)$  have the following physical sense (for precise mathematical expressions, see Appendix F):

- $\theta^{k \rightarrow i}(t+1)$  is the probability that the infection signal  $\lambda$  has not been passed from node  $k$  to node  $i$  up to time  $t+1$  in the cavity dynamics  $D_i$ ;
- $\phi^{k \rightarrow i}(t)$  is the probability that the infection signal  $\lambda$  has not been passed from the node  $k$  to the node  $i$  up to time  $t$  in the cavity dynamics  $D_i$  and that  $k$  is in the state  $I$  at time  $t$ ;
- $m^{k \rightarrow i}(\tau_k, t | T, T)$  is the marginal probability that node  $k$  has the trajectory  $(\tau_k, t)$  in the cavity dynamics  $D_i$ .

Hence, the last term in the equation (48) represents a contribution to the change of  $\phi^{k \rightarrow i}(t)$  due to the recovery of the node  $k$  exactly at time  $t$  in the cavity dynamics  $D_i$ . The initial conditions are given by  $\theta^{k \rightarrow i}(0) = 1$ , and  $\phi^{k \rightarrow i}(0) = \delta_{\sigma_k^0, I}$ .

So far, the equations (46-48) are not in a closed form, we still need to know how to compute  $m^{i \rightarrow j}(\tau_i, t | T, T)$  for  $\tau_i < t$ . We have for each  $t$

$$\begin{aligned} m^{i \rightarrow j}(0, t | \tau_j, T) &= P_I^i(0) \\ &\times \left[ f_{\rho_1, \chi_1}^{i \rightarrow j}(-2, t-2 | \tau_j) - f_{\rho_1, \chi_1}^{i \rightarrow j}(-2, t-1 | \tau_j) \right], \end{aligned} \quad (49)$$

$$\begin{aligned} m^{i \rightarrow j}(\tau_i, t | \tau_j, T) &= P_S^i(0) \\ &\times \left[ f_{\rho_1, \chi_1}^{i \rightarrow j}(\tau_i - 2, t-2 | \tau_j) - f_{\rho_1, \chi_1}^{i \rightarrow j}(\tau_i - 2, t-1 | \tau_j) \right. \\ &\left. - f_{\rho_2, \chi_2}^{i \rightarrow j}(\tau_i - 1, t-2 | \tau_j) + f_{\rho_2, \chi_2}^{i \rightarrow j}(\tau_i - 1, t-1 | \tau_j) \right], \end{aligned} \quad (50)$$

for  $1 \leq \tau_i \leq t-1$  and  $0 \leq \tau_j \leq t$ . The functional  $f_{\rho, \chi}^{i \rightarrow j}(t_1, t_2 | \tau_j)$  is defined as follows:

$$f_{\rho, \chi}^{i \rightarrow j}(t_1, t_2 | \tau_j) = \rho^{j \rightarrow i}(t_1, t_2 | \tau_j) \prod_{k \in \partial i \setminus j} \chi^{k \rightarrow i}(t_1, t_2), \quad (51)$$

where the  $\tau_j$ -dependent coefficients, characterizing the influence of node  $j$  on the dynamics of  $i$ , read for  $t_2 = t-2$



or  $t_2 = t - 1$ :

$$\begin{aligned} & \rho_1^{j \rightarrow i}(\tau_i - 2, t_2 | \tau_j) \\ &= (1 - \lambda_{ji})^{\tau_i - \tau_j - 1} \left( \prod_{t'=\tau_i}^{t_2} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t']) \right), \quad (52) \end{aligned}$$

$$\begin{aligned} & \rho_2^{j \rightarrow i}(\tau_i - 1, t_2 | \tau_j) \\ &= (1 - \lambda_{ji})^{\tau_i - \tau_j} \left( \prod_{t'=\tau_i}^{t_2} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t']) \right). \quad (53) \end{aligned}$$

Let us remind at this point, that the convention (9) is used here and in the following. Note that in the update equation (48) we are only interested in the messages of the form  $m^{i \rightarrow j}(\tau_i, t | T, T)$ , that correspond to  $\tau_j = t$ , and for which the  $j$ -influence is not present: in this case  $\rho_1^{j \rightarrow i}(\tau_i - 2, t_2 | \tau_j)$  and  $\rho_2^{j \rightarrow i}(\tau_i - 1, t_2 | \tau_j)$  are simply equal to one. Still, in the computation scheme for  $\chi_1^{k \rightarrow i}(\tau_i - 2, t_1)$  and  $\chi_2^{k \rightarrow i}(\tau_i - 1, t_1)$  all the values  $0 \leq \tau_j \leq t - 1$  are also required, since the remaining update equations read

$$\begin{aligned} \chi_1^{k \rightarrow i}(\tau_i - 2, t - 1) &= \chi_1^{k \rightarrow i}(\tau_i - 2, t - 2) \\ &\quad - \alpha_{ki} \psi_1^{k \rightarrow i}(\tau_i - 2, t - 1), \quad (54) \end{aligned}$$

$$\begin{aligned} \chi_2^{k \rightarrow i}(\tau_i - 1, t - 1) &= \chi_2^{k \rightarrow i}(\tau_i - 1, t - 2) \\ &\quad - \alpha_{ki} \psi_2^{k \rightarrow i}(\tau_i - 1, t - 1), \quad (55) \end{aligned}$$

and

$$\begin{aligned} & \psi_1^{k \rightarrow i}(\tau_i - 2, t - 1) = P_S^{k \rightarrow i}(t - 2 | \tau_i) \\ & - P_S^{k \rightarrow i}(t - 1 | \tau_i) + (1 - \alpha_{ki} \mathbb{1}_{\tau_i \neq t-1}) \psi_1^{k \rightarrow i}(\tau_i - 2, t - 2) \\ & - \sum_{\tau_k \leq t-2} (1 - \lambda_{ki})^{\tau_i - \tau_k - 1} \left( \prod_{t'=\tau_i}^{t-2} (1 - \alpha_{ki} \mathbb{1}[\tau_k \leq t']) \right) \\ & \quad \times m^{k \rightarrow i}(\tau_k, t - 1 | \tau_i, T), \quad (56) \end{aligned}$$

$$\begin{aligned} & \psi_2^{k \rightarrow i}(\tau_i - 1, t - 1) = P_S^{k \rightarrow i}(t - 2 | \tau_i) \\ & - P_S^{k \rightarrow i}(t - 1 | \tau_i) + (1 - \alpha_{ki} \mathbb{1}_{\tau_i \neq t-1}) \psi_2^{k \rightarrow i}(\tau_i - 1, t - 2) \\ & - \sum_{\tau_k \leq t-2} (1 - \lambda_{ki})^{\tau_i - \tau_k} \left( \prod_{t'=\tau_i}^{t-2} (1 - \alpha_{ki} \mathbb{1}[\tau_k \leq t']) \right) \\ & \quad \times m^{k \rightarrow i}(\tau_k, t - 1 | \tau_i, T). \quad (57) \end{aligned}$$

The conditional quantity  $P_S^{k \rightarrow i}(t_1 | \tau_i)$  is defined as

$$P_S^{k \rightarrow i}(t_1 | \tau_i) = P_S^k(0) (1 - \lambda_{ik})^{t_1 - \tau_i} \prod_{l \in \partial k \setminus i} \theta^{l \rightarrow k}(t_1). \quad (58)$$

The necessary initial conditions are given by  $\chi_1^{k \rightarrow i}(-2, -1) = 1$  and  $\psi_1^{k \rightarrow i}(-2, 0) = \phi^{k \rightarrow i}(0)$ . The following border conditions are used for  $\tau_i = t - 1$ :

$$\chi_1^{k \rightarrow i}(t - 3, t - 2) = \theta^{k \rightarrow i}(t - 2), \quad (59)$$

$$\chi_2^{k \rightarrow i}(t - 2, t - 2) = \theta^{k \rightarrow i}(t - 1), \quad (60)$$

$$\psi_1^{k \rightarrow i}(t - 3, t - 2) = \phi^{k \rightarrow i}(t - 2), \quad (61)$$

$$\psi_2^{k \rightarrow i}(t - 2, t - 2) = (1 - \lambda_{ki}) \phi^{k \rightarrow i}(t - 2). \quad (62)$$

and  $\chi_1^{k \rightarrow i}(t - 3, t - 1)$ ,  $\chi_2^{k \rightarrow i}(t - 2, t - 1)$ ,  $\psi_1^{k \rightarrow i}(t - 3, t - 1)$ ,  $\psi_2^{k \rightarrow i}(t - 2, t - 1)$  follow the equations (54-57).

Therefore, the computation of  $m^{i \rightarrow j}(\tau_i, t | T, T)$  for  $\tau_i < t$  involves messages  $m^{i \rightarrow j}(\tau_i, t - 1 | \tau_j, T)$  for  $\tau_i < t - 1$  and  $\tau_j \leq t - 1$ , computed at a previous step. Finally, the marginal probabilities  $m^i(\tau_i, t + 1)$  are computed via equations (49) and (50), with replacement of the indices  $i \rightarrow j$  simply by  $i$ , and the corresponding change of product over  $k \in \partial i \setminus j$  in the definition (51) by the product over all the neighboring nodes  $k \in \partial i$ . The computational complexity of DMP equations for the rumor spreading model is  $\mathcal{O}(Nct^3)$ , where  $c$  is the average degree of the graph. The details of the derivation are presented in the Appendix F.

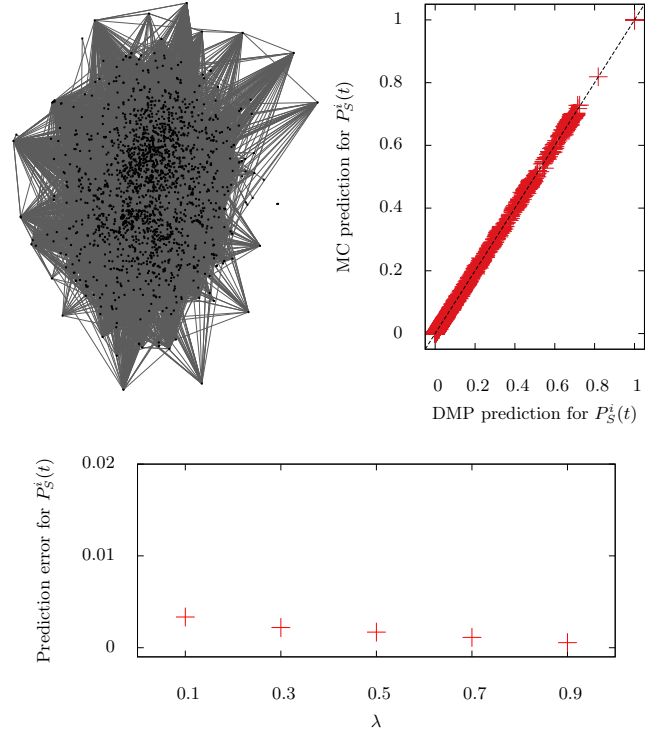


FIG. 4: (Color online) Top right: Comparison of prediction of the DMP equations for the rumor spreading model with the Monte Carlo (MC) simulations in a typical case. Marginal probabilities  $P_S^i(t)$  are presented for the Facebook-like social network with  $N = 1899$  nodes and  $t = 10$ , the parameters of the model are  $\lambda = 0.3$ ,  $\alpha = 0.2$ , there is one infected node at initial time. Top left: A representation of the topology of the network, generated with Gephi [43]. The high-degree nodes (hubs) are placed on the periphery. Bottom: Study of the average prediction error per node for marginal probabilities  $P_S^i(t)$  as a function of  $\lambda$ . In this plot,  $t = 10$  and  $\alpha = 0.2$ , so that the point  $\lambda = 0.3$  corresponds to the comparison above. For both plots, the MC average is performed over  $10^4$  instances. The error bars are smaller than the symbol size on the plots and are not shown.

The validity of these equations has been checked numerically via comparison with the Monte Carlo (MC) simulation: the marginals given by the DMP equations

appear to be exact on any tree graph. Although *a priori* the DMP equations are not guaranteed to be exact on networks that do not have a locally tree-like structure, they provide remarkably accurate predictions even for small and loopy networks. For example, we have tested the performance of the DMP equations for the rumor spreading model on two real-world networks. The first example is a Facebook-like social network with 1899 nodes and 20 296 edges that represents an online community for students at University of California, Irvine [44]; the predictions for the marginals given by DMP are compared with the values obtained from  $10^4$  MC simulations, see Fig. 4. Another test has been performed for the small Zachary’s karate club network of friendships between 34 members of a karate club at a US university [45], the results are presented in the Fig. 5. In both cases, the predictions of the DMP equations appear to be very accurate with respect to the true values of the marginals.

The DMP equations for the rumor spreading model have never been reported so far. The existing approaches include the naive mean-field equations [4] that are derived under the complete-mixing assumption and completely neglect the topology of the network, or the so-called heterogeneous mean-field equations [46] that assume equivalent behavior for different nodes of the same degree; they are averaged over the ensemble of random graphs and are not applicable on a single instance of the network.

## VI. UNIDIRECTIONAL MODELS WITH $K > 3$ STATES

The DMP equations for the rumor spreading model, described in the previous section, can be easily generalized to a more complicated pairwise model with three states, similar to the generalized SI model. In this section we will illustrate the procedure for deriving the DMP equations for pairwise models with a larger number of states, using as an example a “minimal” model with  $K = 4$  states, which is an extension of the rumor spreading model with an additional non-trivial transition to the final state. The procedure for deriving these equations from the general dynamic cavity equation (3) is very similar to the derivation of the DMP equations for the rumor spreading model. A generalization for any larger number of states seems to be straightforward. Note that the models with  $K \geq 3$  states that include direct transitions that skip some number of intermediate states can also be taken into account in this approach.

Let us consider four states  $S$ ,  $I_1$ ,  $I_2$  and  $R$ , and the following dynamic rules:

$$S(i) + I_1(j) \xrightarrow{\lambda_{ji}} I_1(i) + I_1(j), \quad (63)$$

$$I_1(i) + I_1(j) \xrightarrow{\alpha_{ji}} I_2(i) + I_1(j), \quad (64)$$

$$I_2(i) + I_2(j) \xrightarrow{\beta_{ji}} R(i) + I_2(j). \quad (65)$$

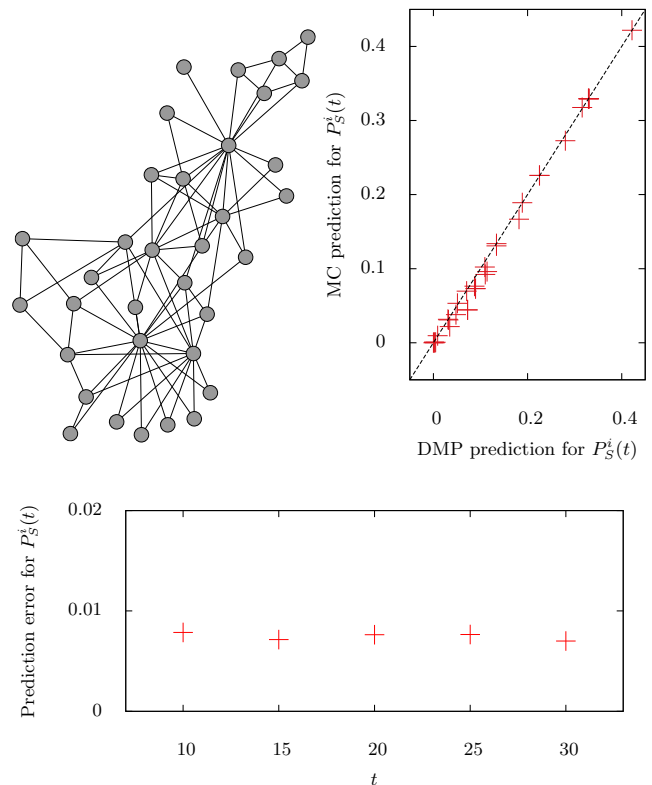


FIG. 5: (Color online) Top right: Comparison of prediction of the DMP equations for the rumor spreading model with the Monte Carlo (MC) simulations in a typical case. Marginal probabilities  $P_S^i(t)$  are presented for the Zachary’s karate club network with  $N = 34$  nodes and  $t = 10$ , the parameters of the model are  $\lambda = 0.3$ ,  $\alpha = 0.2$ , there is one infected node at initial time. Top left: A representation of the topology of the network, generated with Gephi [43]. This network has a block structure and contains many loops of small length. Bottom: Study of the average prediction error for marginal probabilities  $P_S^i(t)$  as a function of time  $t$ . In this plot,  $\lambda = 0.3$  and  $\alpha = 0.2$ , so that the point  $t = 10$  corresponds to the comparison above. For both plots, the MC average is performed over  $10^4$  instances. The error bars are smaller than the symbol size on the plots and are not shown.

Now the time trajectory of the node  $i$  can be parametrized by three flipping times:  $\tau_i$  (first time in  $I_1$ ),  $\omega_i$  (first time in  $I_2$ ) and  $\varepsilon_i$  (first time in  $R$ ). The trajectory of the spin  $i$  is hence described by  $\vec{\sigma}_i(t) = (\tau_i, \omega_i, \varepsilon_i)$ , and the corresponding marginal of the dynamic cavity equation (4) could be written as  $m^i(\tau_i, \omega_i, \varepsilon_i)$ . Similarly to the SI, the SIR and the rumor spreading model, we might expect that the expressions for the marginal prob-

abilities at time  $t$  could be written in the following form:

$$P_S^i(t+1) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \rightarrow i}(t+1), \quad (66)$$

$$P_R^i(t+1) = P_R^i(t) + \sum_{\substack{\omega_i \leq t \\ \tau_i + 1 \leq \omega_i}} m^i(\tau_i, \omega_i, t+1), \quad (67)$$

$$P_{I_2}^i(t+1) = P_{I_2}^i(t) + \sum_{\substack{\tau_i \leq t \\ \varepsilon_i > t+1}} m^i(\tau_i, t+1, \varepsilon_i), \quad (68)$$

$$P_{I_1}^i(t+1) = 1 - P_S^i(t+1) - P_{I_2}^i(t+1) - P_R^i(t+1). \quad (69)$$

The apparent difficulty in the equation (68) is that the sum runs over all the flipping times  $\varepsilon_i > t+1$ , and the number of terms can potentially be very big, of the order of the stopping time  $T$ . In the Appendix G it is shown that this difficulty can be overcome if one defines a new sort of messages:

$$\mu^{k \rightarrow i}(\tau_k, t | T, T) = \sum_{\varepsilon_k \geq t+1} m^{k \rightarrow i}(\tau_k, t, \varepsilon_k | T, T, T). \quad (70)$$

The evolution of  $\mu^{k \rightarrow i}(\tau_k, t | T, T)$  follows the same equations as for the rumor spreading model (49-62), except that now we will require the computation of  $\mu^{k \rightarrow i}(\tau_k, t | \tau_i, \omega_i)$ . The details of derivation for this case is presented in the Appendix G.

The generalization of this model may describe different models with 4 states, for example the generalization of the SIR model that include immunized or exposed states [2]. As it has been expected, the computational complexity for this  $K = 4$  model is higher:  $\mathcal{O}(Nct^5)$ , where  $c$  is the average degree of the graph. From the structure of the solution for this minimal model, which makes use of the equations for the model with a lower number of steps, we conjecture that a general model with unidirectional dynamics and  $M$  non-trivial transitions will have the computational complexity growing as  $t^{2M-1}$ . Note that  $M$  is not always equal to  $K - 1$ , for instance, compare the SIR ( $K = 3$ ,  $M = 1$ ) and the rumor spreading ( $K = 3$ ,  $M = 2$ ) models.

## VII. CONCLUSION

In this paper we have developed a general approach for deriving the dynamic message-passing equations that describe models with unidirectional dynamics and an arbitrary number of states. These equations can be iterated in physical time starting from arbitrary initial conditions, and allow one to estimate the exact values of marginal probabilities on locally tree-like graphs, providing good approximation for real-world networks. These

closed-form equations can be derived starting from the dynamic belief propagation equation on time trajectories, and using the causality and the normalization properties of messages. The dynamic variables that appear in the resulting DMP equations are typically represented by the weighted sums of dynamic BP messages, and emerge automatically in this approach.

Importantly, although the general formulation of dynamic BP on trajectories is of exponential complexity, it takes only a polynomial number of steps  $t^{2M-1}$  (where  $M$  is the number of non-trivial transitions in the considered models) to solve the corresponding DMP equations. The growth of the number of operations with the number of states of the model in the DMP equations is essentially due to the local effects of retro-action that have to be taken into account. It would be interesting to understand whether there is a way to reduce further this computational complexity, or to prove that this is the minimal number of operations required in order to provide exact equations for these models on tree networks.

The DMP approach opens a way to a number of applications aimed at a better control of the cascading processes on networks. Since the transmission probabilities can be time-dependent, the DMP equations can be used for the dynamically changing graphs. The fact that the DMP equations can be applied to a single instance of a graph has been recently used for the algorithmic application to the inverse problem in the context of the epidemic spreading: the inference of the epidemic origin of an epidemic outbreak [31]. The DMP equations are also promising for optimization and control in models that incorporate the changes of individuals' behavior during the dynamical process [4, 46]. Polynomial-complexity DMP equations for the forward dynamics could be used on their own in these potential applications. Another possible strategy would consist in adding extra check nodes (corresponding to the optimization constraints) directly in the factor graph, exemplified in the Fig. 2. This may, however, lead to convergence issues of the corresponding dynamic BP equations for sufficiently complex optimization problems; the discussion of this point is beyond the scope of the present work. These and other applications of the DMP approach to optimization problems are left for future work.

## Acknowledgments

The authors are grateful to Hiroki Ohta and Silvio Franz for fruitful discussions and valuable comments. This work has been supported by the Grant DySpaN of Triangle de la Physique.

## Appendix A: Belief propagation equations for static problems

Let us assume that a static problem is formulated in terms of a graphical model, defined on a tree graph by a joint probability distribution

$$P(\underline{\sigma}) = \frac{1}{Z} \prod_{a=1}^M \psi_a(\underline{\sigma}_{\partial a}), \quad (\text{A1})$$

where  $Z$  is the normalization constant. Note that this expression is given in a factorized form, each factor  $\psi_a$  representing a local interaction weight. Very often it is convenient to represent the graphical model in a form of a corresponding factor graph that reflects this structure of the model (A1). The factor graph can be thought of as a bipartite graph  $G = (V, F, E)$ :  $V$  is a list of variables,  $\underline{\sigma} = \{\sigma_i\}_{i \in V}$ , and  $F$  represents a list of interactions, or function nodes, so that an edge  $(i, a) \in E$  is present if the interaction  $a \in F$  involves a variable  $\sigma_i$  in the node  $i \in V$ . We assume here that the set of possible values of  $\sigma_i$  is a finite set of size  $K$  ( $K = 2$  for binary spins or boolean variables,  $K > 3$  for Potts spins or colors, *etc.*). The neighboring nodes in the factor graphs are denoted by  $\partial i$  and  $\partial a$  for the variables and function nodes, correspondingly. The marginal probability distribution (also called belief) that the variable on the site  $i$  takes value  $\sigma_i$ , is defined as

$$m^i(\sigma_i) = \sum_{\underline{\sigma}_{\setminus i}} P(\underline{\sigma}). \quad (\text{A2})$$

The basic idea of the belief propagation method (also known as the sum-product algorithm) consists in the following observation: since the model is defined on a tree, when one removes a site  $i$  from the graph  $G$  and cuts the corresponding connection to the neighboring interactions, the resulting cavity graph  $G^{(i)}$  is given by a collection of independent and statistically uncorrelated branches of a tree. Therefore, the marginal  $m^i(\sigma_i)$  can be expressed simply as a product over the conditional probabilities that represent the contributions of these branches:

$$m^i(\sigma_i) = \frac{1}{Z^i} \prod_{a \in \partial i} \hat{m}^{a \rightarrow i}(\sigma_i). \quad (\text{A3})$$

In this expression,  $Z^i$  is the normalization factor that ensures that  $\sum_{\sigma_i} m^i(\sigma_i) = 1$ , and  $\partial i$  stands for the neighbors of  $i$  in the factor graph. The probability  $\hat{m}^{a \rightarrow i}(\sigma_i)$ , called the message, is defined as the marginal probability that node  $i$  takes value  $\sigma_i$  in the modified graph, in which all the interactions around  $i$  except  $a$  have been cut out. Now, in order to compute  $m^i(\sigma_i)$ , one needs to know the values of  $\hat{m}^{a \rightarrow i}(\sigma_i)$ . These quantities obey the coupled self-consistency equations [26, 38]

$$m^{i \rightarrow a}(\sigma_i) = \frac{1}{Z^{i \rightarrow a}} \prod_{b \in \partial i \setminus a} \hat{m}^{b \rightarrow i}(\sigma_i), \quad (\text{A4})$$

$$\hat{m}^{a \rightarrow i}(\sigma_i) = \frac{1}{Z^{a \rightarrow i}} \sum_{\underline{\sigma}_{\partial a \setminus i}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{k \in \partial a \setminus i} m^{k \rightarrow a}(\sigma_k), \quad (\text{A5})$$

where we have also introduced another sort of messages,  $m^{i \rightarrow a}(\sigma_i)$ , which is defined as the marginal probability that node  $i$  takes value  $\sigma_i$  in the modified graph, in which the interaction  $a$  has been deleted. In these equations,  $Z^{i \rightarrow a}$  and  $Z^{a \rightarrow i}$  are the normalization constants. The coupled equations (A4) and (A5) are usually solved by iteration: first, one initialises all the messages to some value, and iterate the equations (A4) and (A5) until convergence. Then, the final values for the messages  $\hat{m}^{a \rightarrow i}(\sigma_i)$  are used in (A3) for computing the exact values of the marginal probability distributions  $m^i(\sigma_i)$ . This procedure explains the fine terminology of the BP algorithm: the iteration of equations (A4) and (A5) can be thought of as a message-passing protocol, each message holding a conditional information on the probability of the corresponding variable; the marginal is then given by a belief shaped by the information contained in all the messages arriving to the node.

Note that although the equations (A3)-(A5) have been derived for a tree graph, they can be viewed as an algorithm that can be run on an arbitrary interaction graph. They will provide accurate estimations of the marginals as long as the replica symmetric assumption holds for the interaction graph, i.e. that the correlations induced by loops decay fast enough, so that the approximation (A3) as a product over neighboring interactions is correct (see [26] for more details). In particular, the BP equations (A3)-(A5) give asymptotically exact (in the thermodynamic limit  $N \rightarrow \infty$ )

expressions for the beliefs on the only locally tree-like networks; sparse random graphs, as well as many real-world networks of interest, fall into this category.

Sometimes, it is easier to eliminate one sort of messages in (A4) and (A5) and to use a single iteration equation for messages instead of two:

$$\widehat{m}^{a \rightarrow i}(\sigma_i) = \frac{1}{Z^{\rightarrow i}} \sum_{\underline{\sigma}_{\partial a \setminus i}} \psi_a(\underline{\sigma}_{\partial a}) \prod_{k \in \partial a \setminus i} \prod_{b \in \partial k \setminus a} \widehat{m}^{b \rightarrow k}(\sigma_k), \quad (\text{A6})$$

with  $Z^{\rightarrow i} = Z^{a \rightarrow i} \prod_{k \in \partial a \setminus i} Z^{k \rightarrow a}$ . This expression further simplifies in the important case of pairwise models, when the variables interact pairs by pairs, and the joint probability distribution factorizes over the links in the graph:

$$P(\underline{\sigma}) = \frac{1}{Z} \prod_{(ij)} \psi_{ij}(\sigma_i, \sigma_j). \quad (\text{A7})$$

In this case, the update equation (A6) can be rewritten exclusively in terms of messages  $m^{i \rightarrow j}(\sigma_i)$ , a short-cut for  $m^{i \rightarrow (ij)}(\sigma_i)$ :

$$m^{i \rightarrow j}(\sigma_i) = \frac{1}{Z^{i \rightarrow j}} \prod_{k \in \partial i \setminus j} \sum_{\sigma_k} \psi_{ik}(\sigma_i, \sigma_k) m^{k \rightarrow i}(\sigma_k). \quad (\text{A8})$$

## Appendix B: Properties of the dynamic belief propagation equations

The exact BP equation for the joint probability distribution of pairs of time trajectories (2) in terms of messages  $m^{i \rightarrow j}(\vec{\sigma}_i, \vec{\sigma}_j)$  reads:

$$m^{i \rightarrow j}(\vec{\sigma}_i, \vec{\sigma}_j) = \frac{1}{Z^{i \rightarrow j}} \sum_{\{\vec{\sigma}_k\}_{k \in \partial i \setminus j}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} \mid \{\sigma_k^t\}_{k \in \partial i \setminus j}, \sigma_j^t) P(\{\sigma_i^0\}_{i \in V}) \right] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\vec{\sigma}_k, \vec{\sigma}_i). \quad (\text{B1})$$

The normalization constant  $Z^{i \rightarrow j}$  can be calculated explicitly for the Markovian dynamics from the normalization condition

$$\sum_{\vec{\sigma}_i, \vec{\sigma}_j} m^{i \rightarrow j}(\vec{\sigma}_i, \vec{\sigma}_j) = 1 \quad (\text{B2})$$

For example, in the case of general Markov dynamics we use the fact that  $m^{k \rightarrow i}(\vec{\sigma}_k, \vec{\sigma}_i)$  does not depend on  $\sigma_i^T$  and perform the summation first over  $\sigma_j^T$ , then over  $\sigma_i^T$ , and so on for the times  $T-1, \dots, 0$ . Finally, we get the normalization factor

$$Z^{i \rightarrow j} = \frac{1}{2^{(T+1)(d_i-2)}} \quad (\text{B3})$$

for this case, where  $d_i$  is the number of neighbors of the node  $i$  in the initial graph.

The message  $m^{i \rightarrow j}(\vec{\sigma}_i, \vec{\sigma}_j)$  has a meaning of probability for the trajectories  $\vec{\sigma}_i, \vec{\sigma}_j$  in the transformed cavity graph, where the factor node  $j$  has been removed. We can also rewrite the equation (B1) in terms of conditional probabilities  $m^{i \rightarrow j}(\vec{\sigma}_i \mid \vec{\sigma}_j)$  on the cavity graph. Thus, for the dynamics obeying the Markov property we get

$$\sum_{\vec{\sigma}_i} m^{i \rightarrow j}(\vec{\sigma}_i, \vec{\sigma}_j) = \frac{1}{2^{T+1}}, \quad (\text{B4})$$

and hence recover the equations (3) and (4):

$$m^{i \rightarrow j}(\vec{\sigma}_i \mid \vec{\sigma}_j) = \sum_{\{\vec{\sigma}_k\}_{k \in \partial i \setminus j}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} \mid \{\sigma_k^t\}_{k \in \partial i \setminus j}, \sigma_j^t) P(\{\sigma_i^0\}_{i \in V}) \right] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\vec{\sigma}_k \mid \vec{\sigma}_i), \quad (\text{B5})$$

$$m^i(\vec{\sigma}_i) = \sum_{\{\vec{\sigma}_k\}_{k \in \partial i}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial i}) P(\{\sigma_i^0\}_{i \in V}) \right] \prod_{k \in \partial i} m^{k \rightarrow i}(\vec{\sigma}_k | \vec{\sigma}_i). \quad (\text{B6})$$

The message  $m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$  has a meaning of probability for the trajectory  $\vec{\sigma}_i$  given the trajectory  $\vec{\sigma}_j$  in the transformed cavity graph, where the factor node  $j$  has been removed. The normalization factor in this equation is exactly equal to 1, due to the Markov property of the dynamics. Note that, again, by construction,  $m^{k \rightarrow i}(\vec{\sigma}_k | \vec{\sigma}_i)$  does not depend on  $\sigma_i^T$ , so this variable can be erased. Then, as far as  $\prod_{t=1}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial i \setminus j}, \sigma_j^t)$  does not depend on  $\sigma_k^T$ , we can perform the sum over  $\sigma_k^T$  in the right-hand side of the equation (B5), which gives exactly the Kanoria-Montanari expression [21]:

$$m_{T+1}^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j) = \sum_{\{\sigma_k^0, \dots, \sigma_k^{T-1}\}_{k \in \partial i \setminus j}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial i \setminus j}, \sigma_j^t) P(\{\sigma_i^0\}_{i \in V}) \right] \prod_{k \in \partial i \setminus j} m_T^{k \rightarrow i}(\vec{\sigma}_k | \vec{\sigma}_i), \quad (\text{B7})$$

where we denote  $m_{T+1}^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j) = m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$  and  $m_T^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j) = \sum_{\sigma_i^{T+1}} m_{T+1}^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$ .

### Purely directed case

In the purely directed case, the direct influence between neighboring nodes  $i$  and  $j$  runs only in one direction. It means that  $m^{i \rightarrow j}(\vec{\sigma}_i | \vec{\sigma}_j)$  does not depend on the variable  $\vec{\sigma}_j$ , and the equation (B5) is reduced to

$$m^{i \rightarrow j}(\vec{\sigma}_i) = \sum_{\{\sigma_k\}_{k \in \partial_{in} i}} \left[ \prod_{t=0}^{T-1} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial_{in} i}) P(\{\sigma_i^0\}_{i \in V}) \right] \prod_{k \in \partial_{in} i} m^{k \rightarrow i}(\vec{\sigma}_k), \quad (\text{B8})$$

where  $\partial_{in} i$  denotes the set of neighbors of  $i$  that have a direct influence on  $i$ . Therefore, writing the marginal in a factorized form

$$m^i(\vec{\sigma}_i) = \prod_t m_t^i(\sigma_i^t), \quad (\text{B9})$$

we immediately get from (B5) for  $t > 0$

$$m_{t+1}^i(\sigma_i^{t+1}) = \sum_{\{\sigma_k^t\}_{k \in \partial_{in} i}} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial_{in} i}) \prod_{k \in \partial_{in} i} m_t^k(\sigma_k^t). \quad (\text{B10})$$

Note, however, that in the case of the dynamics in which the state of a node depends on the state of the same node at previous time, the factorization (B9) does not lead to a decoupled expression, and we have

$$\prod_t m_{t+1}^i(\sigma_i^{t+1}) = \prod_t \left[ \sum_{\{\sigma_k^t\}_{k \in \partial_{in} i}} w_i(\sigma_i^{t+1} | \{\sigma_k^t\}_{k \in \partial_{in} i}, \sigma_i^t) P(\{\sigma_i^0\}_{i \in V}) \prod_{k \in \partial_{in} i} m_t^k(\sigma_k^t) \right]. \quad (\text{B11})$$

### Appendix C: Derivation of the DMP equations for the zero-temperature RFIM

Using the dynamic properties of the model, the equation (3) for  $\tau_i < T$  can be expressed as follows:

$$m^{i \rightarrow j}(\tau_i | \tau_j) = \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} W_{RFIM} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i), \quad (\text{C1})$$

where

$$W_{RFIM} = \prod_{t'=0}^{\tau_i-2} \left(1 - \frac{1}{\tau} \mathbb{1}[\Delta_i^{t'} > 0]\right) \frac{1}{\tau} \mathbb{1}[\Delta_i^{\tau_i-1} > 0]. \quad (\text{C2})$$

Recall that here and in what follows we use a convention

$$\prod_{t=a}^{a-\epsilon} (\dots) \equiv 1 \quad (\text{C3})$$

for any fixed  $a$  and  $\epsilon > 0$ . We define the marginal probabilities in the cavity dynamics  $D_j$  as

$$p^{i \rightarrow j}(t) = \sum_{\tau_i > t} m^{i \rightarrow j}(\tau_i | T), \quad (\text{C4})$$

$$q^{i \rightarrow j}(t) = \sum_{\tau_i \leq t} m^{i \rightarrow j}(\tau_i | T), \quad (\text{C5})$$

and the probabilities of interest in the original dynamics as

$$p^i(t) = \sum_{\tau_i > t} m^i(\tau_i), \quad (\text{C6})$$

$$q^i(t) = 1 - p^i(t). \quad (\text{C7})$$

The corresponding DMP equation for this model can be derived, starting from the equation (C1) and using elementary properties of the messages, such as normalization and causality constraints. First of all, let us rewrite (C1) in an equivalent form, explicitly using the monotonicity of the local field (6)

$$m^{i \rightarrow j}(\tau_i | \tau_j) = \frac{1}{\tau} \sum_{\Theta_i^* = 0}^{\tau_i-1} \left(1 - \frac{1}{\tau}\right)^{(\tau_i-1-\tau_i^*)} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i^*-1} < 0] \mathbb{1}[\Delta_i^{\tau_i^*} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i), \quad (\text{C8})$$

where  $\tau_i^*$  appears explicitly as the moment when the local field becomes positive for the first time and the convention  $\mathbb{1}[\Delta_i^{-1} < 0] \equiv 1$  is used. In what follows, we use two natural properties of messages.

**Property 1.** The norm conservation. For every fixed  $\tau_j$

$$\sum_{\tau_i=0}^T m^{i \rightarrow j}(\tau_i | \tau_j) = 1. \quad (\text{C9})$$

**Property 2.** If  $\tau_j > \tau_i$ , then for every  $t' > \tau_i$

$$m^{i \rightarrow j}(\tau_i | \tau_j) = m^{i \rightarrow j}(\tau_i | t'). \quad (\text{C10})$$

Using the definition (C4), we get

$$p^{i \rightarrow j}(t+1) = \sum_{\tau_i \geq t+1} \sum_{\tau_i^* = 0}^{\tau_i-1} \frac{1}{\tau} \left(1 - \frac{1}{\tau}\right)^{(\tau_i-1-\tau_i^*)} \sum_{\{\Theta_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i^*-1} < 0] \mathbb{1}[\Delta_i^{\tau_i^*} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i). \quad (\text{C11})$$

An important observation that can be made on this expression is that nothing changes if we replace in the right-hand side of the equation above the messages  $m^{k \rightarrow i}(\tau_k | \tau_i)$ , conditioned on the flipping time  $\tau_i$ , by the messages conditioned on the stopping time  $T$ ,  $m^{k \rightarrow i}(\tau_k | T)$ . The easiest way to see it consists in observing that the value of the probability  $p^{i \rightarrow j}(t+1)$  should not depend on the value of the stopping time provided that  $T > t+1$ , and can be assigned to an arbitrary value. Since in (C11) we are only interested in  $\tau_i > t+1$ , we can in particular choose  $T = t+2$ , and since the stopping time by definition comprises all the events that happen after the time  $T$ , we get

$$p^{i \rightarrow j}(t+1) = \sum_{\tau_i \geq t+1} \sum_{\tau_i^* = 0}^{\tau_i-1} \frac{1}{\tau} \left(1 - \frac{1}{\tau}\right)^{(\tau_i-1-\tau_i^*)} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i^*-1} < 0] \mathbb{1}[\Delta_i^{\tau_i^*} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T). \quad (\text{C12})$$

The sums over  $\tau_i$  and  $\tau_i^*$  in this expression can be split into two terms:

$$\sum_{\tau_i > t+1} \sum_{\tau_i^* = 0}^{\tau_i - 1} = \sum_{\tau_i > t+1} \delta_{\tau_i^*, \tau_i - 1} + \sum_{\tau_i > t+1} \sum_{\tau_i^* = 0}^{\tau_i - 2}, \quad (\text{C13})$$

and we get

$$\begin{aligned} p^{i \rightarrow j}(t+1) &= \sum_{\tau_i > t+1} \frac{1}{\tau} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i - 2} < 0] \mathbb{1}[\Delta_i^{\tau_i - 1} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &+ \sum_{\tau_i > t+1} \sum_{\tau_i^* = 0}^{\tau_i - 2} \frac{1}{\tau} \left(1 - \frac{1}{\tau}\right)^{(\tau_i - 1 - \tau_i^*)} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i^* - 1} < 0] \mathbb{1}[\Delta_i^{\tau_i^*} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &= \frac{1}{\tau} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^t < 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &+ \left(1 - \frac{1}{\tau}\right) \sum_{\tau_i' > t+1} \sum_{\tau_i^* = 0}^{\tau_i' - 1} \frac{1}{\tau} \left(1 - \frac{1}{\tau}\right)^{(\tau_i' - 1 - \tau_i^*)} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^{\tau_i^* - 1} < 0] \mathbb{1}[\Delta_i^{\tau_i^*} > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &= \frac{1}{\tau} \left(1 - \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^t > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T)\right) + \left(1 - \frac{1}{\tau}\right) p^{i \rightarrow j}(t). \end{aligned} \quad (\text{C14})$$

The remaining sum over flipping times  $\{\tau_k\}_{k \in \partial i \setminus j}$  can be further simplified by pushing the sum over the indicator function:

$$\begin{aligned} &\sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\Delta_i^t > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &= \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \left[ \sum_{\{\sigma_k^t\}_{k \in \partial i \setminus j} = \pm 1} \prod_{k \in \partial i \setminus j} \delta\left(\frac{\sigma_k^t + 1}{2}, \mathbb{1}[\tau_k < t]\right) \right] \mathbb{1}[\Delta_i^t > 0] \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T) \\ &= \sum_{\{\sigma_k^t\}_{k \in \partial i \setminus j} = \pm 1} \mathbb{1}\left[h + h_i + \sum_{k \in \partial i \setminus j} J_{ki} \sigma_k^t - J_{ji} > 0\right] \prod_{k \in \partial i \setminus j} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} m^{k \rightarrow i}(\tau_k | T) \delta\left(\frac{\sigma_k^t + 1}{2}, \mathbb{1}[\tau_k < t]\right) \\ &= \sum_{\{\sigma_k^t\}_{k \in \partial i \setminus j} = \pm 1} \mathbb{1}\left[h + h_i + \sum_{k \in \partial i \setminus j} J_{ki} \sigma_k^t - J_{ji} > 0\right] \prod_{k \in \partial i \setminus j} \prod_{k \in \partial i \setminus j: \sigma_k^t = +1} [1 - p^{k \rightarrow i}(t)] \prod_{k \in \partial i \setminus j: \sigma_k^t = -1} p^{k \rightarrow i}(t). \end{aligned} \quad (\text{C15})$$

Finally, using (C15) and rewriting (C14) in terms of  $q^{i \rightarrow j}(t) = 1 - p^{i \rightarrow j}(t)$ , we recover the DMP equations (15) and (16).

#### Appendix D: Derivation of the DMP equations for the generalized SI model

The dynamic belief propagation equation (3) for  $\tau_i < T$  takes the following form in the generalized SI model:

$$m^{i \rightarrow j}(\tau_i | \tau_j) = \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} W_{SI} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i), \quad (\text{D1})$$

where the dynamic kernel  $W_{SI}$  for the generalized SI model has the following form:

$$\begin{aligned} W_{SI} &= P_I^i(0) \mathbb{1}[\tau_i = 0] + P_S^i(0) \mathbb{1}[\tau_i > 0] \prod_{t'=0}^{\tau_i - 2} (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \\ &\times \left(1 - (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq \tau_i]) (1 - \lambda_{ki} \mathbb{1}[\tau_i \geq \tau_k + 1])\right). \end{aligned} \quad (\text{D2})$$



As it is described in the main text, the messages  $m^{i \rightarrow j}(\tau_i | \tau_j)$  allow one to define the marginal probabilities describing the dynamics of the SI model:

$$P_S^i(t) = \sum_{\tau_i > t} m^i(\tau_i), \quad (\text{D3})$$

$$P_I^i(t) = 1 - P_S^i(t). \quad (\text{D4})$$

It would also be useful to define the marginal probability that the node  $i$  is in the state  $S$  at a given time in the cavity dynamics  $D_j$ , in which the node  $j$  is fixed to the state  $S$  for all times:

$$P_S^{i \rightarrow j}(t) = \sum_{\tau_i > t} m^{i \rightarrow j}(\tau_i | T). \quad (\text{D5})$$

By analogy with (C9) and (C10), the messages in the dynamic cavity equation (D1) have the normalization and causality properties.

**Property 1.** For every fixed  $\tau_j$

$$\sum_{\tau_i=0}^T m^{i \rightarrow j}(\tau_i | \tau_j) = 1. \quad (\text{D6})$$

**Property 2.** If  $\tau_j > \tau_i$ , then for every  $t' > \tau_i$

$$m^{i \rightarrow j}(\tau_i | \tau_j) = m^{i \rightarrow j}(\tau_i | t'). \quad (\text{D7})$$

Using the definition (D5), we get for  $t > 0$

$$\begin{aligned} P_S^{i \rightarrow j}(t+1) &= P_S^i(0) \sum_{\tau_i > t+1} \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \mathbb{1}[\tau_j = T] \prod_{t'=0}^{\tau_i-2} (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \\ &\times \left( 1 - (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq \tau_i]) (1 - \lambda_{ki} \mathbb{1}[\tau_i \geq \tau_k + 1]) \right) \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i). \end{aligned} \quad (\text{D8})$$

Using the same arguments as in the derivation of the DMP equations for the zero-temperature RFIM, it can be shown that one can replace  $\prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | \tau_i)$  in the right-hand side of the last expression by  $\prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k | T)$  for arbitrary value of the stopping time  $T > t+1$ . Since we are interested in the messages  $m^{i \rightarrow j}(\tau_i | T)$  with  $\tau_i > t+1$ , we have

$$\begin{aligned} P_S^{i \rightarrow j}(t+1) &= P_S^i(0) (1 - \nu_i)^{t+1} (1 - \epsilon_{ji})^{t+1} \\ &\times \sum_{\{\tau_k\}_{k \in \partial i \setminus j}} \prod_{k \in \partial i \setminus j} \prod_{t'=0}^t (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) m^{k \rightarrow i}(\tau_k | T) \\ &\times \left[ \sum_{\tau_i > t+1} \mathbb{1}[\tau_j = T] \prod_{t'=t+1}^{\tau_i-2} (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \right. \\ &\quad \left. \times \left( 1 - (1 - \nu_i) \prod_{k \in \partial i} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq \tau_i]) (1 - \lambda_{ki} \mathbb{1}[\tau_i \geq \tau_k + 1]) \right) \right]. \end{aligned} \quad (\text{D9})$$

Developing the sum in the square brackets, one may ascertain that it gives exactly one:

$$\begin{aligned}
& 1 - (1 - \nu_i)(1 - \epsilon_{ji}) \prod_{k \in \partial i \setminus j} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 2]) (1 - \lambda_{ki} \mathbb{1}[t + 2 \geq \tau_k + 1]) \\
& + (1 - \nu_i)(1 - \epsilon_{ji}) \prod_{k \in \partial i \setminus j} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 2]) (1 - \lambda_{ki} \mathbb{1}[t + 1 \geq \tau_k]) \\
& - (1 - \nu_i)^2 (1 - \epsilon_{ji})^2 \prod_{k \in \partial i \setminus j} \left[ (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 2]) (1 - \lambda_{ki} \mathbb{1}[t + 2 \geq \tau_k + 1]) \right] \\
& \quad \times \prod_{k \in \partial i \setminus j} \left[ (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 3]) (1 - \lambda_{ki} \mathbb{1}[t + 3 \geq \tau_k + 1]) \right] \\
& + (1 - \nu_i)^2 (1 - \epsilon_{ji})^2 \prod_{k \in \partial i \setminus j} \left[ (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 2]) (1 - \lambda_{ki} \mathbb{1}[t + 1 \geq \tau_k]) \right] \\
& \quad \times \prod_{k \in \partial i \setminus j} \left[ (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t + 3]) (1 - \lambda_{ki} \mathbb{1}[t + 2 \geq \tau_k]) \right] + \dots = 1, \tag{D10}
\end{aligned}$$

and therefore we obtain the factorized expression (24)

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0)(1 - \epsilon_{ji})^{t+1}(1 - \nu_i)^{t+1} \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \tag{D11}$$

where  $\theta^{k \rightarrow i}(t+1)$  are given by

$$\theta^{k \rightarrow i}(t+1) = \sum_{\tau_k} \prod_{t'=0}^t (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) m^{k \rightarrow i}(\tau_k | T). \tag{D12}$$

In order to close the equations on  $P_S^{k \rightarrow i}(t)$ , we recover the computational scheme for  $\theta^{k \rightarrow i}(t+1)$ :

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) \tag{D13}$$

$$= \sum_{\tau_k} \left( \prod_{t'=0}^{t-1} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \right) (-\epsilon_{ki} \mathbb{1}[\tau_k \geq t + 1] - \lambda_{ki} \mathbb{1}[t \geq \tau_k]) m^{k \rightarrow i}(\tau_k | T) \tag{D14}$$

$$= -\epsilon_{ki} \sum_{\tau_k} \left( \prod_{t'=0}^{t-1} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) \right) m^{k \rightarrow i}(\tau_k | T) \mathbb{1}[\tau_k \geq t + 1] \tag{D15}$$

$$- \lambda_{ki} \sum_{\tau_k} \left( \prod_{t'=0}^{t-1} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \right) m^{k \rightarrow i}(\tau_k | T) \mathbb{1}[t \geq \tau_k] \tag{D16}$$

$$\equiv -\epsilon_{ki} \phi_1^{k \rightarrow i}(t) - \lambda_{ki} \phi_2^{k \rightarrow i}(t), \tag{D17}$$

where, using the identity  $\mathbb{1}[\tau_k \geq t + 1] = \mathbb{1}[\tau_k \geq t] - \delta(\tau_k, t)$ , we get for  $\phi_1^{k \rightarrow i}(t)$

$$\phi_1^{k \rightarrow i}(t) = \sum_{\tau_k} \left( \prod_{t'=0}^{t-2} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) \right) (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t]) m^{k \rightarrow i}(\tau_k | T) \mathbb{1}[\tau_k \geq t + 1] \tag{D18}$$

$$= (1 - \epsilon_{ki}) \phi_1^{k \rightarrow i}(t-1) - \left( \prod_{t'=0}^{t-1} (1 - \epsilon_{ki}) \right) m^{k \rightarrow i}(t | T) = (1 - \epsilon_{ki}) \phi_1^{k \rightarrow i}(t-1) - (1 - \epsilon_{ki})^t (P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t)), \tag{D19}$$

and for  $\phi_2^{k \rightarrow i}(t)$

$$\phi_2^{k \rightarrow i}(t) = \sum_{\tau_k} \left( \prod_{t'=0}^{t-2} (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t' + 1]) (1 - \lambda_{ki} \mathbb{1}[t' \geq \tau_k]) \right) (1 - \epsilon_{ki} \mathbb{1}[\tau_k \geq t]) \tag{D20}$$

$$\times (1 - \lambda_{ki} \mathbb{1}[t \geq \tau_k + 1]) m^{k \rightarrow i}(\tau_k | T) \mathbb{1}[t \geq \tau_k] = (1 - \lambda_{ki}) \phi_2^{k \rightarrow i}(t-1) + (1 - \epsilon_{ki})^t (P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t)). \tag{D21}$$

This gives exactly the computational scheme (24-27).

### Appendix E: Derivation of the DMP equations for the SIR model

The derivation of the DMP equations for the SIR model follows the very same lines as the derivation for the SI model, but includes several subtleties. First of all, since now each node can be in one of three states ( $S$  - susceptible,  $I$  - infected and  $R$  - recovered), the trajectory of the node  $i$  can be parametrized by two times  $\tau_i$  and  $\omega_i$ :  $\tau_i$  is defined as a first time to be in the state  $I$ , while  $\omega_i$  is a first time to be in the state  $R$ . We will denote the trajectory of the spin  $i$  as  $\vec{\sigma}_i(t) = (\tau_i, \omega_i)$ . Therefore, the dynamic BP equation (3) for  $\omega_i < T$  becomes in this case:

$$m^{i \rightarrow j}(\tau_i, \omega_i | \tau_j, \omega_j) = \sum_{\{\tau_k, \omega_k\}_{k \in \partial i \setminus j}} W_{SIR} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k, \omega_k | \tau_i, \omega_i), \quad (\text{E1})$$

where

$$W_{SIR} = \left[ P_S^i(0) \mathbb{1}[\tau_i > 0] \prod_{t'=0}^{\tau_i-2} \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \left( 1 - \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq \tau_i] \mathbb{1}[\tau_k \leq \tau_i - 1]) \right) + P_I^i(0) \mathbb{1}[\tau_i = 0] \right] \times \left( \prod_{t'=\tau_i}^{\omega_i-2} (1 - \mu_i) \right) \times \mu_i \times \prod_{k \in \partial i} \mathbb{1}[\omega_k \geq \tau_k + 1] \mathbb{1}[\omega_i \geq \tau_i + 1]. \quad (\text{E2})$$

The marginals of interest in the SIR model can be defined as

$$P_S^i(t) = \sum_{\tau_i > t} \sum_{\omega_i > \tau_i} m^i(\tau_i, \omega_i), \quad (\text{E3})$$

$$P_I^i(t) = \sum_{\tau_i \leq t} \sum_{\omega_i > t} m^i(\tau_i, \omega_i), \quad (\text{E4})$$

$$P_R^i(t) = \sum_{\omega_i \leq t} \sum_{\tau_i < \omega_i} m^i(\tau_i, \omega_i). \quad (\text{E5})$$

We also define the marginal probability for the susceptible state in the corresponding cavity graph:

$$P_S^{i \rightarrow j}(t) = \sum_{\tau_i > t} \sum_{\omega_i > \tau_i} m^{i \rightarrow j}(\tau_i, \omega_i | T, T). \quad (\text{E6})$$

Let us point out the properties of the messages.

**Property 1.**  $m^{i \rightarrow j}(\tau_i, \omega_i | T, T) = 0$  if  $\tau_i \geq \omega_i$ ;

**Property 2.** If  $\tau_j \geq \tau_i$ , then  $m^{i \rightarrow j}(\tau_i, \omega_i | \tau_j, \omega_j) = m^{i \rightarrow j}(\tau_i, \omega_i | t', \omega_j)$  for every  $\tau_i \leq t' < \omega_j$ ;

**Property 3.**  $\sum_{\tau_i, \omega_i} m^{i \rightarrow j}(\tau_i, \omega_i | T, T) = 1$ ;

**Property 4.**  $m^{i \rightarrow j}(\tau_i, \omega_i + 1 | T, T) = (1 - \mu_i) m^{i \rightarrow j}(\tau_i, \omega_i | T, T)$ .

The properties, equivalent to properties 1, 2 and 4, are also valid for marginals  $m^i(\tau_i, \omega_i)$ . It is straightforward to establish first two evolution equations on the quantities  $P_S^i(t)$ ,  $P_I^i(t)$  and  $P_R^i(t)$ . According to the definitions,

$$P_R^i(t+1) = \sum_{\omega_i \leq t+1} \sum_{\tau_i < \omega_i} m^i(\tau_i, \omega_i) = \sum_{\omega_i \leq t} \sum_{\tau_i < \omega_i} m^i(\tau_i, \omega_i) + \delta_{\omega_i, t+1} \sum_{\tau_i \leq t} m^i(\tau_i, \omega_i) = P_R^i(t) + \mu_i P_I^i(t), \quad (\text{E7})$$

where we used the property 4 of marginals, because

$$\sum_{\omega_i \geq t+1} m^i(\tau_i, \omega_i) = \frac{1}{1 - (1 - \mu_i)} m^i(\tau_i, t+1) = \frac{1}{\mu_i} m^i(\tau_i, t+1). \quad (\text{E8})$$

Since the expressions defined in (E3-E5) sum to one, it is obvious that

$$P_I^i(t+1) = 1 - P_S^i(t+1) - P_R^i(t+1). \quad (\text{E9})$$

In what follows we show that we can put  $P_S^{i \rightarrow j}(t+1)$  in the form

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0) \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \quad (\text{E10})$$

where  $\theta^{k \rightarrow i}(t+1)$  can be calculated via  $P_S^{k \rightarrow i}(t)$  at each time step. The equations (E7), (E9), (E10) and its marginalized version, together with a computational scheme for  $\theta^{i \rightarrow j}(t+1)$ , form a closed set of equations for the SIR model.

We proceed in the same way as for the SI model. The quantity  $\theta^{k \rightarrow i}(t+1)$  is now defined as

$$\theta^{k \rightarrow i}(t+1) = \sum_{\tau_k, \omega_k} \mathbb{1}[\omega_k \geq \tau_k + 1] \prod_{t'=0}^t (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) m^{k \rightarrow i}(\tau_k, \omega_k | T, T). \quad (\text{E11})$$

As for the SI model, we can write

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) \equiv -\lambda_{ki} \phi^{k \rightarrow i}(t), \quad (\text{E12})$$

where we get for  $\phi^{k \rightarrow i}(t)$

$$\begin{aligned} \phi^{k \rightarrow i}(t) &= \\ &= \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{t-2} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) \right) (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t-1]) \mathbb{1}[\omega_k \geq t+1] \mathbb{1}[\tau_k \leq t] m^{k \rightarrow i}(\tau_k, \omega_k | T, T) \\ &= (1 - \lambda_{ki})(1 - \mu_k) \phi^{k \rightarrow i}(t-1) - \sum_{\omega_k} \mathbb{1}[\omega_k \geq t+1] m^{k \rightarrow i}(t, \omega_k | T, T) \\ &= (1 - \lambda_{ki})(1 - \mu_k) \phi^{k \rightarrow i}(t-1) - (P_S^{k \rightarrow i}(t) - P_S^{k \rightarrow i}(t-1)). \end{aligned} \quad (\text{E13})$$

In the last expression the factor  $(1 - \mu_k)$  is due to the property 4 of messages. Equations (E12) and (E13) complete the computational scheme for  $\theta^{k \rightarrow i}(t+1)$  for each time step, and we recover the DMP equations (35-40).

## Appendix F: Derivation of the DMP equations for the rumor spreading model

Let us follow the main lines of derivation of these DMP equations, concentrating mainly on the subtleties with the respect to the SIR model. Again, the trajectory of the node  $i$  can be parametrized by two flipping times:  $\tau_i$  (first time in  $I$ ) and  $\omega_i$  (first time in  $R$ ). The trajectory of the spin  $i$  is hence described by  $\vec{\sigma}_i(t) = (\tau_i, \omega_i)$ . The dynamic BP equation (3) in the rumor spreading model reads for  $\omega_i < T$ :

$$m^{i \rightarrow j}(\tau_i, \omega_i | \tau_j, \omega_j) = \sum_{\{\tau_k, \omega_k\}_{k \in \partial i \setminus j}} W_{RS} \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k, \omega_k | \tau_i, \omega_i), \quad (\text{F1})$$

where

$$\begin{aligned} W_{RS} &= \left[ P_S^i(0) \mathbb{1}[\tau_i > 0] \prod_{t'=0}^{\tau_i-2} \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \left( 1 - \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq \tau_i] \mathbb{1}[\tau_k \leq \tau_i - 1]) \right) \right] + \\ &+ P_I^i(0) \mathbb{1}[\tau_i = 0] \times \prod_{t''=\tau_i}^{\omega_i-2} \prod_{k \in \partial i} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \left( 1 - \prod_{k \in \partial i} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq \omega_i] \mathbb{1}[\tau_k \leq \omega_i - 1]) \right) \times \\ &\times \prod_{k \in \partial i} \mathbb{1}[\omega_k \geq \tau_k + 1] \mathbb{1}[\omega_i \geq \tau_i + 1]. \end{aligned} \quad (\text{F2})$$

Quantities  $P_S^i(t)$ ,  $P_I^i(t)$ ,  $P_R^i(t)$  and  $P_S^{i \rightarrow j}(t)$  are defined in the same way as in (E3-E6).

The following properties of the messages hold.

**Property 1.**  $m^{i \rightarrow j}(\tau_i, \omega_i | T, T) = 0$  if  $\tau_i \geq \omega_i$ ;

**Property 2.** If  $\tau_j \geq \tau_i$ , then  $m^{i \rightarrow j}(\tau_i, \omega_i | \tau_j, \omega_j) = m^{i \rightarrow j}(\tau_i, \omega_i | t', \omega_j)$  for every  $\tau_i \leq t' < \omega_j$ ;

**Property 3.**  $\sum_{\tau_i, \omega_i} m^{i \rightarrow j}(\tau_i, \omega_i | T, T) = 1$ .

According to the definitions, we have

$$P_R^i(t+1) = \sum_{\omega_i \leq t+1} \sum_{\tau_i < \omega_i} m^{i \rightarrow j}(\tau_i, \omega_i) = \sum_{\omega_i \leq t} \sum_{\tau_i < \omega_i} m^i(\tau_i, \omega_i) + \delta_{\omega_i, t+1} \sum_{\tau_i \leq t} m^i(\tau_i, \omega_i) = P_R^i(t) + \sum_{\tau_i \leq t} m^i(\tau_i, \omega_i). \quad (\text{F3})$$

The expressions defined in (E3-E5) sum to one, hence

$$P_I^i(t+1) = 1 - P_S^i(t+1) - P_R^i(t+1). \quad (\text{F4})$$

As for the SIR model, we show that we can put  $P_S^{i \rightarrow j}(t+1)$  in the form

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0) \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \quad (\text{F5})$$

where  $\theta^{k \rightarrow i}(t+1)$  is defined as

$$\theta^{k \rightarrow i}(t+1) = \sum_{\tau_k, \omega_k} \mathbb{1}[\omega_k \geq \tau_k + 1] \prod_{t'=0}^t (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) m^{k \rightarrow i}(\tau_k, \omega_k | T, T), \quad (\text{F6})$$

and we have in the same way

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) \equiv -\lambda_{ki} \phi^{k \rightarrow i}(t). \quad (\text{F7})$$

The quantity  $\phi^{k \rightarrow i}(t)$  is defined as

$$\phi^{k \rightarrow i}(t) = \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{t-1} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) \right) \mathbb{1}[\omega_k \geq t+1] \mathbb{1}[\tau_k \leq t] m^{k \rightarrow i}(\tau_k, \omega_k | T, T). \quad (\text{F8})$$

Since  $\mathbb{1}[\omega_k \geq t+1] = \mathbb{1}[\omega_k \geq t] - \delta_{\omega_k, t}$  and  $\mathbb{1}[\tau_k \leq t] = \mathbb{1}[\tau_k \leq t-1] + \delta_{\tau_k, t}$ , we obtain

$$\phi^{k \rightarrow i}(t) = (1 - \lambda_{ki}) \phi^{k \rightarrow i}(t-1) + P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t) - \sum_{\tau_k \leq t-1} \prod_{t'=0}^{t-1} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) m^{k \rightarrow i}(\tau_k, t | T, T). \quad (\text{F9})$$

The last term corresponds to the probability of recovering at time step  $t$ . As for the SIR model, the initial conditions for  $\theta^{k \rightarrow i}(t)$  and  $\phi^{k \rightarrow i}(t)$  are given by  $\theta^{k \rightarrow i}(0) = 1$ , and  $\phi^{k \rightarrow i}(0) = \delta_{\sigma_k^0, I}$ . Since these equations are in not in a closed form, we proceed to the computation of  $m^{i \rightarrow j}(\tau_i, t | T, T)$  for  $\tau_i < t$ . The equations (49) and (50) follow directly from (F1) with dynamics (F2), if we denote (for  $t_1 = t-2$  or  $t_1 = t-1$ )

$$\begin{aligned} \chi_1^{k \rightarrow i}(\tau_i - 2, t_1) = \\ \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{\tau_i - 2} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \right) \left( \prod_{t''=\tau_i}^{t_1} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \right) m^{k \rightarrow i}(\tau_k, \omega_k | \tau_i, t), \quad (\text{F10}) \end{aligned}$$

and

$$\begin{aligned} \chi_2^{k \rightarrow i}(\tau_i - 1, t_1) = \\ \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{\tau_i - 1} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \right) \left( \prod_{t''=\tau_i}^{t_1} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \right) m^{k \rightarrow i}(\tau_k, \omega_k | \tau_i, t). \quad (\text{F11}) \end{aligned}$$

Similarly to the evolution of  $\theta^{k \rightarrow i}(t)$ , the evolution equations for  $\chi_1^{k \rightarrow i}(\tau_i - 2, t_1)$  and  $\chi_2^{k \rightarrow i}(\tau_i - 1, t_1)$ , given their definitions, read

$$\chi_1^{k \rightarrow i}(\tau_i - 2, t - 1) = \chi_1^{k \rightarrow i}(\tau_i - 2, t - 2) - \alpha_{ki} \psi_1^{k \rightarrow i}(\tau_i - 2, t - 1), \quad (\text{F12})$$

$$\chi_2^{k \rightarrow i}(\tau_i - 1, t - 1) = \chi_2^{k \rightarrow i}(\tau_i - 1, t - 2) - \alpha_{ki} \psi_2^{k \rightarrow i}(\tau_i - 1, t - 1), \quad (\text{F13})$$

where  $\psi_1^{k \rightarrow i}(\tau_i - 2, t_1)$  and  $\psi_2^{k \rightarrow i}(\tau_i - 1, t_1)$  are defined as

$$\begin{aligned} \psi_1^{k \rightarrow i}(\tau_i - 2, t_1) = \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{\tau_i-2} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \right) \left( \prod_{t''=\tau_i}^{t_1-1} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \right) \\ \times \mathbb{1}[\omega_k \geq t_1 + 1] \mathbb{1}[\tau_k \leq t_1] m^{k \rightarrow i}(\tau_k, \omega_k \mid \tau_i, t), \end{aligned} \quad (\text{F14})$$

and

$$\begin{aligned} \psi_2^{k \rightarrow i}(\tau_i - 1, t_1) = \sum_{\tau_k, \omega_k > \tau_k} \left( \prod_{t'=0}^{\tau_i-1} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \right) \left( \prod_{t''=\tau_i}^{t_1-1} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \right) \\ \times \mathbb{1}[\omega_k \geq t_1 + 1] \mathbb{1}[\tau_k \leq t_1] m^{k \rightarrow i}(\tau_k, \omega_k \mid \tau_i, t). \end{aligned} \quad (\text{F15})$$

In order to close the computational scheme, we have to write the update equations for  $\psi_1^{k \rightarrow i}(\tau_i - 2, t_1)$  and  $\psi_2^{k \rightarrow i}(\tau_i - 1, t_1)$ . One has to be careful, because the messages involved in these quantities are conditioned on the state of the node  $i$ , with flipping times  $(\tau_i, t)$ . Using again the identities  $\mathbb{1}[\omega_k \geq t_2 + 1] = \mathbb{1}[\omega_k \geq t_2] - \delta_{\omega_k, t_2}$  and  $\mathbb{1}[\tau_k \leq t_2] = \mathbb{1}[\tau_k \leq t_2 - 1] + \delta_{\tau_k, t_2}$ , we recover the equations (56) and (57), with the definition of conditional probability of staying in the state  $S$  (58).

Finally one can see, directly from the definitions, that the initial conditions for this computational scheme are given by  $\chi_1^{k \rightarrow i}(-2, -1) = 1$  and  $\psi_1^{k \rightarrow i}(-2, 0) = \phi^{k \rightarrow i}(0)$ , and the border conditions for  $\tau_i = t - 1$  are given for each time step by expressions (59-62).

### Appendix G: Example of derivation of the DMP equations for the $K = 4$ states model with unidirectional dynamics

The dynamic BP equation (3) in the minimal  $K = 4$  model reads for  $\varepsilon_i < T$ :

$$m^{i \rightarrow j}(\tau_i, \omega_i, \varepsilon_i \mid \tau_j, \omega_j, \varepsilon_j) = \sum_{\{\tau_k, \omega_k, \varepsilon_k\}_{k \in \partial i \setminus j}} W_4 \prod_{k \in \partial i \setminus j} m^{k \rightarrow i}(\tau_k, \omega_k, \varepsilon_k \mid \tau_i, \omega_i, \varepsilon_i), \quad (\text{G1})$$

where

$$\begin{aligned} W_4 = \left[ P_S^i(0) \mathbb{1}[\tau_i > 0] \prod_{t'=0}^{\tau_i-2} \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) \left( 1 - \prod_{k \in \partial i} (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq \tau_i] \mathbb{1}[\tau_k \leq \tau_i - 1]) \right) \right. \\ \left. + P_I^i(0) \mathbb{1}[\tau_i = 0] \right] \times \prod_{t''=\tau_i}^{\omega_i-2} \prod_{k \in \partial i} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq t'' + 1] \mathbb{1}[\tau_k \leq t'']) \left( 1 - \prod_{k \in \partial i} (1 - \alpha_{ki} \mathbb{1}[\omega_k \geq \omega_i] \mathbb{1}[\tau_k \leq \omega_i - 1]) \right) \\ \times \prod_{t'''=\omega_i}^{\varepsilon_i-2} \prod_{k \in \partial i} (1 - \beta_{ki} \mathbb{1}[\varepsilon_k \geq t''' + 1] \mathbb{1}[\omega_k \leq t''']) \left( 1 - \prod_{k \in \partial i} (1 - \beta_{ki} \mathbb{1}[\varepsilon_k \geq \varepsilon_i] \mathbb{1}[\omega_k \leq \varepsilon_i - 1]) \right) \\ \times \prod_{k \in \partial i} \mathbb{1}[\varepsilon_k \geq \omega_k + 1] \mathbb{1}[\omega_k \geq \tau_k + 1] \mathbb{1}[\omega_i \geq \tau_i + 1]. \end{aligned} \quad (\text{G2})$$

Quantities  $P_S^i(t)$ ,  $P_{I_1}^i(t)$ ,  $P_{I_2}^i(t)$ ,  $P_R^i(t)$  and  $P_S^{i \rightarrow j}(t)$  are defined in a way similar to (E3-E6). For instance,

$$P_S^{i \rightarrow j}(t) = \sum_{\tau_i > t} \sum_{\omega_i > \tau_i} \sum_{\varepsilon_i > \omega_i} m^{i \rightarrow j}(\tau_i, \omega_i, \varepsilon_i | T, T, T). \quad (\text{G3})$$

As previously, we can put  $P_S^{i \rightarrow j}(t+1)$  in the form

$$P_S^{i \rightarrow j}(t+1) = P_S^i(0) \prod_{k \in \partial i \setminus j} \theta^{k \rightarrow i}(t+1), \quad (\text{G4})$$

where  $\theta^{k \rightarrow i}(t+1)$  is now defined as

$$\theta^{k \rightarrow i}(t+1) = \sum_{\tau_k, \omega_k, \varepsilon_k} \mathbb{1}[\varepsilon_k \geq \omega_k + 1] \mathbb{1}[\omega_k \geq \tau_k + 1] \prod_{t'=0}^t (1 - \lambda_{ki} \mathbb{1}[\omega_k \geq t' + 1] \mathbb{1}[\tau_k \leq t']) m^{k \rightarrow i}(\tau_k, \omega_k, \varepsilon_k | T, T, T), \quad (\text{G5})$$

and we have in the same way

$$\theta^{k \rightarrow i}(t+1) - \theta^{k \rightarrow i}(t) \equiv -\lambda_{ki} \phi^{k \rightarrow i}(t). \quad (\text{G6})$$

The quantity  $\phi^{k \rightarrow i}(t)$  is now defined as

$$\phi^{k \rightarrow i}(t) = \sum_{\substack{\tau_k, \omega_k > \tau_k \\ \varepsilon_k > \omega_k}} \left( \prod_{t'=0}^{t-1} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) \right) \mathbb{1}[\omega_k \geq t + 1] \mathbb{1}[\tau_k \leq t] m^{k \rightarrow i}(\tau_k, \omega_k, \varepsilon_k | T, T, T). \quad (\text{G7})$$

Since  $\mathbb{1}[\omega_k \geq t + 1] = \mathbb{1}[\omega_k \geq t] - \delta_{\omega_k, t}$  and  $\mathbb{1}[\tau_k \leq t] = \mathbb{1}[\tau_k \leq t - 1] + \delta_{\tau_k, t}$ , we obtain

$$\phi^{k \rightarrow i}(t) = (1 - \lambda_{ki}) \phi^{k \rightarrow i}(t-1) + P_S^{k \rightarrow i}(t-1) - P_S^{k \rightarrow i}(t) - \sum_{\tau_k \leq t-1} \prod_{t'=0}^{t-1} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) \sum_{\varepsilon_k \geq t+1} m^{k \rightarrow i}(\tau_k, t, \varepsilon_k | T, T, T). \quad (\text{G8})$$

The last term corresponds to the probability of switching to the  $I_2$  state at time step  $t$ . We see that for this term, exactly the same problem as the one indicated in the main text for the equation (68) remains: the sum over  $\varepsilon_k$  goes over a number of terms of order of  $T$ . At the same time, we see that both equations (68) and (G8) have a finite number of terms if we define new variables

$$\mu^{k \rightarrow i}(\tau_k, t | T, T) = \sum_{\varepsilon_k \geq t+1} m^{k \rightarrow i}(\tau_k, t, \varepsilon_k | T, T, T). \quad (\text{G9})$$

On the other hand, in order to close the computational scheme (66-69), (G4) (G6), (G8), one needs to write the evolution equations for both  $m^{i \rightarrow j}(\tau_i, \omega_i, t | T, T, T)$  and  $\mu^{i \rightarrow j}(\tau_i, t | T, T)$  variables. The evolution of  $\mu^{i \rightarrow j}(\tau_i, t | T, T)$  follows the same equations as for the rumor spreading model (49-62), except that now we will require the computation of  $\mu^{i \rightarrow j}(\tau_i, t | \tau_j, \omega_j)$ . Since all the functions  $\chi_1$  and  $\chi_2$  are independent on  $(\tau_j, \omega_j)$ , we will only need to compute the equivalents of (52) and (53). These equivalents are given by the coefficients  $\hat{\rho}_1^{j \rightarrow i}$  and  $\hat{\rho}_2^{j \rightarrow i}$ :

$$\hat{\rho}_1^{j \rightarrow i}(\tau_i - 2, \omega_i - 2, t_1 | \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 2} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 2} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t'' + 1]), \quad (\text{G10})$$

$$\hat{\rho}_2^{j \rightarrow i}(\tau_i - 2, \omega_i - 1, t_1 | \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 2} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 1} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t'' + 1]), \quad (\text{G11})$$

which are the restrictions of  $\zeta_1^{j \rightarrow i}$  and  $\zeta_2^{j \rightarrow i}$  (see definitions below) to the case  $\beta_{ij} = 0$ .

For the  $m^{i \rightarrow j}(\tau_i, \omega_i, t | T, T, T)$ , we proceed in the similar way as for the derivation of (49) and (50) in the rumor spreading model, but now we will have, instead of two and four terms, correspondingly four and eight. For each  $t$ ,

directly from (G1) and (G2), we have

$$m^{i \rightarrow j}(0, \omega_i, t \mid \tau_j, \omega_j, T) = P_I^i(0) \left[ g_{\zeta_1, \xi_1}^{i \rightarrow j}(-2, \omega_i - 2, t - 2 \mid \tau_j, \omega_j) - g_{\zeta_1, \xi_1}^{i \rightarrow j}(-2, \omega_i - 2, t - 1 \mid \tau_j, \omega_j) \right. \\ \left. - g_{\zeta_2, \xi_2}^{i \rightarrow j}(-2, \omega_i - 1, t - 2 \mid \tau_j, \omega_j) + g_{\zeta_2, \xi_2}^{i \rightarrow j}(-2, \omega_i - 1, t - 1 \mid \tau_j, \omega_j) \right] \quad (\text{G12})$$

for  $1 \leq \omega_i \leq t - 1$ ,  $0 \leq \tau_j \leq t - 1$  and  $1 \leq \omega_j \leq t$ , and

$$m^{i \rightarrow j}(\tau_i, \omega_i, t \mid \tau_j, \omega_j, T) = P_S^i(0) \left[ g_{\zeta_1, \xi_1}^{i \rightarrow j}(\tau_i - 2, \omega_i - 2, t - 2 \mid \tau_j, \omega_j) - g_{\zeta_1, \xi_1}^{i \rightarrow j}(\tau_i - 2, \omega_i - 2, t - 1 \mid \tau_j, \omega_j) \right. \\ \left. - g_{\zeta_2, \xi_2}^{i \rightarrow j}(\tau_i - 2, \omega_i - 1, t - 2 \mid \tau_j, \omega_j) + g_{\zeta_2, \xi_2}^{i \rightarrow j}(\tau_i - 2, \omega_i - 1, t - 1 \mid \tau_j, \omega_j) - g_{\zeta_3, \xi_3}^{i \rightarrow j}(\tau_i - 1, \omega_i - 2, t - 2 \mid \tau_j, \omega_j) \right. \\ \left. + g_{\zeta_3, \xi_3}^{i \rightarrow j}(\tau_i - 1, \omega_i - 2, t - 1 \mid \tau_j, \omega_j) + g_{\zeta_4, \xi_4}^{i \rightarrow j}(\tau_i - 1, \omega_i - 1, t - 2 \mid \tau_j, \omega_j) - g_{\zeta_4, \xi_4}^{i \rightarrow j}(\tau_i - 1, \omega_i - 1, t - 1 \mid \tau_j, \omega_j) \right] \quad (\text{G13})$$

for  $1 \leq \tau_i \leq t - 2$ ,  $1 \leq \omega_i \leq t - 1$ ,  $0 \leq \tau_j \leq t - 1$  and  $1 \leq \omega_j \leq t$ , where the functional  $g_{\zeta, \xi}^{i \rightarrow j}(t_1, t_2, t_3 \mid \tau_j, \omega_j)$  is defined as

$$g_{\zeta, \xi}^{i \rightarrow j}(t_1, t_2, t_3 \mid \tau_j, \omega_j) = \zeta^{j \rightarrow i}(t_1, t_2, t_3 \mid \tau_j, \omega_j) \prod_{k \in \partial i \setminus j} \xi^{k \rightarrow i}(t_1, t_2, t_3), \quad (\text{G14})$$

and the  $(\tau_j, \omega_j)$ -dependent coefficients, characterizing the influence of node  $j$  on the dynamics of  $i$ , are defined as follows for  $t_1 = t - 2$  or  $t_1 = t - 1$ :

$$\zeta_1^{j \rightarrow i}(\tau_i - 2, \omega_i - 2, t_1 \mid \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 2} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 2} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\omega_j \geq t'' + 1]) \\ \times \prod_{t'''=\omega_i}^{t_1} (1 - \beta_{ji} \mathbb{1}[\omega_j \leq t''']), \quad (\text{G15})$$

$$\zeta_2^{j \rightarrow i}(\tau_i - 2, \omega_i - 1, t_1 \mid \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 2} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 1} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\omega_j \geq t'' + 1]) \\ \times \prod_{t'''=\omega_i}^{t_1} (1 - \beta_{ji} \mathbb{1}[\omega_j \leq t''']), \quad (\text{G16})$$

$$\zeta_3^{j \rightarrow i}(\tau_i - 1, \omega_i - 2, t_1 \mid \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 1} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 2} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\omega_j \geq t'' + 1]) \\ \times \prod_{t'''=\omega_i}^{t_1} (1 - \beta_{ji} \mathbb{1}[\omega_j \leq t''']), \quad (\text{G17})$$

$$\zeta_4^{j \rightarrow i}(\tau_i - 1, \omega_i - 1, t_1 \mid \tau_j, \omega_j) = \prod_{t'=0}^{\tau_i - 1} (1 - \lambda_{ji} \mathbb{1}[\tau_j \leq t'] \mathbb{1}[\omega_j \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i - 1} (1 - \alpha_{ji} \mathbb{1}[\tau_j \leq t''] \mathbb{1}[\omega_j \geq t'' + 1]) \\ \times \prod_{t'''=\omega_i}^{t_1} (1 - \beta_{ji} \mathbb{1}[\omega_j \leq t''']). \quad (\text{G18})$$

The functions  $\xi_l^{k \rightarrow i}$ ,  $l = 1 \dots 4$ , follow the evolution equations, similar to (54) and (55):

$$\xi_1^{k \rightarrow i}(\tau_i - 2, \omega_i - 2, t - 1) = \xi_1^{k \rightarrow i}(\tau_i - 2, \omega_i - 2, t - 2) - \beta_{ki} \eta_1^{k \rightarrow i}(\tau_i - 2, \omega_i - 2, t - 1), \quad (\text{G19})$$

$$\xi_2^{k \rightarrow i}(\tau_i - 2, \omega_i - 1, t - 1) = \xi_2^{k \rightarrow i}(\tau_i - 2, \omega_i - 1, t - 2) - \beta_{ki} \eta_2^{k \rightarrow i}(\tau_i - 2, \omega_i - 1, t - 1), \quad (\text{G20})$$

$$\xi_3^{k \rightarrow i}(\tau_i - 1, \omega_i - 2, t - 1) = \xi_3^{k \rightarrow i}(\tau_i - 1, \omega_i - 2, t - 2) - \beta_{ki} \eta_3^{k \rightarrow i}(\tau_i - 1, \omega_i - 2, t - 1), \quad (\text{G21})$$

$$\xi_4^{k \rightarrow i}(\tau_i - 1, \omega_i - 1, t - 1) = \xi_4^{k \rightarrow i}(\tau_i - 1, \omega_i - 1, t - 2) - \beta_{ki} \eta_4^{k \rightarrow i}(\tau_i - 1, \omega_i - 1, t - 1). \quad (\text{G22})$$



It is straightforward to convince oneself that  $\eta_l^{k \rightarrow i}$ ,  $l = 1 \dots 4$ , follow the equations of the type

$$\begin{aligned}
\eta_1^{k \rightarrow i}(\tau_i - 2, \omega_i - 2, t - 1) &= (1 - \beta_{ki} \mathbb{1}_{\omega_i \neq t-1}) \eta_1^{k \rightarrow i}(\tau_i - 2, \omega_i - 2, t - 2) \\
&+ \sum_{\tau_k \leq t-2} \prod_{t'=0}^{\tau_i-2} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t']) \prod_{t''=\tau_i}^{t-2} (1 - \alpha_{ki} \mathbb{1}[\tau_k \leq t'']) \mu^{k \rightarrow i}(\tau_k, t - 1 \mid \tau_i, \omega_i) \\
&- \sum_{\tau_k < \omega_k \leq t-2} \prod_{t'=0}^{\tau_i-2} (1 - \lambda_{ki} \mathbb{1}[\tau_k \leq t'] \mathbb{1}[\omega_k \geq t' + 1]) \prod_{t''=\tau_i}^{\omega_i-2} (1 - \alpha_{ki} \mathbb{1}[\tau_k \leq t''] \mathbb{1}[\omega_k \geq t'' + 1]) \\
&\times \prod_{t'''=\omega_i}^{t-2} (1 - \beta_{ki} \mathbb{1}[\omega_k \leq t''']) m^{k \rightarrow i}(\tau_k, \omega_k, t - 1 \mid \tau_i, \omega_i, T). \tag{G23}
\end{aligned}$$

The border conditions for  $\omega_i = t - 1$  are as follows:

$$\xi_1^{k \rightarrow i}(\tau_i - 2, t - 3, t - 2) = \chi_1^{k \rightarrow i}(\tau_i - 2, t - 3), \quad \eta_1^{k \rightarrow i}(\tau_i - 2, t - 3, t - 2) = 1 - \psi_1^{k \rightarrow i}(\tau_i - 2, t - 2), \tag{G24}$$

$$\xi_2^{k \rightarrow i}(\tau_i - 2, t - 2, t - 2) = \chi_1^{k \rightarrow i}(\tau_i - 2, t - 2), \quad \eta_2^{k \rightarrow i}(\tau_i - 2, t - 2, t - 2) = 1 - (1 - \alpha_{ki}) \psi_1^{k \rightarrow i}(\tau_i - 2, t - 2), \tag{G25}$$

$$\xi_3^{k \rightarrow i}(\tau_i - 1, t - 3, t - 2) = \chi_2^{k \rightarrow i}(\tau_i - 1, t - 3), \quad \eta_3^{k \rightarrow i}(\tau_i - 1, t - 3, t - 2) = 1 - \psi_2^{k \rightarrow i}(\tau_i - 1, t - 2), \tag{G26}$$

$$\xi_4^{k \rightarrow i}(\tau_i - 1, t - 2, t - 2) = \chi_2^{k \rightarrow i}(\tau_i - 1, t - 2), \quad \eta_4^{k \rightarrow i}(\tau_i - 1, t - 2, t - 2) = 1 - (1 - \alpha_{ki}) \psi_2^{k \rightarrow i}(\tau_i - 1, t - 2), \tag{G27}$$

and initial conditions are  $\xi_1^{k \rightarrow i}(-2, -1, 0) = \xi_2^{k \rightarrow i}(-2, 0, 0) = 1$ ,  $\eta_1^{k \rightarrow i}(-2, -1, 1) = \eta_2^{k \rightarrow i}(-2, 0, 1) = \mu^{k \rightarrow i}(0, 1 \mid 0, 1)$ .

- 
- [1] J. D. Murray, *Mathematical biology* (Springer-Verlag, 1989).
- [2] R. Anderson and R. May, *Infectious diseases of humans: dynamics and control* (Oxford University Press, 1991).
- [3] H. W. Hethcote, *SIAM Rev.* **42**, 599 (2000).
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Physics Reports* **424**, 175 (2006).
- [5] E. M. Rogers, *Diffusion of innovations* (Simon and Schuster, 2010).
- [6] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, *ACM Transactions on Information and System Security* **10**, 1 (2008).
- [7] M. Llas, P. M. Gleiser, J. M. López, and A. Díaz-Guilera, *Physical Review E* **68**, 066101 (2003).
- [8] X. Guardiola, A. Díaz-Guilera, C. J. Pérez, A. Arenas, and M. Llas, *Physical Review E* **66**, 026121 (2002).
- [9] D. Strang and S. A. Soule, *Annual Review of Sociology* **24**, 265 (1998).
- [10] D. Dhar, P. Shukla, and J. P. Sethna, *Journal of Physics A: Mathematical and General* **30**, 5259 (1997).
- [11] H. Ohta and S. Sasa, *EPL (Europhysics Letters)* **90**, 27008 (2010).
- [12] A. Demers, M. Gealy, D. Greene, C. Hauser, W. Irish, J. Larson, S. Manning, H. Sturgis, D. Swinehart, D. Woods, et al., *ACM Proceedings of the sixth annual ACM Symposium on Principles of distributed computing* pp. 1–12 (1987).
- [13] W. Vogels, R. V. Renesse, and K. Birman, *ACM SIGCOMM Computer Communication Review* **33(1)**, 131 (2003).
- [14] R. O’Dea, J. J. Crofts, and M. Kaiser, *Journal of The Royal Society Interface* **10**, 20130016 (2013).
- [15] B. Kholodenko, M. B. Yaffe, and W. Kolch, *Science signaling* **5**, re1 (2012).
- [16] A. J. Kimmel, *Journal of Behavioral Finance* **5**, 134 (2004).
- [17] M. Kosfeld, *Journal of Economic Literature* (1998).
- [18] J. P. L. Hatchett, B. Wemmenhove, I. P. Castillo, T. Nikolettopoulos, N. S. Skantzos, and A. C. C. Coolen, *Journal of Physics A: Mathematical and General* **37**, 6201 (2004).
- [19] K. Mimura and A. C. C. Coolen, *Journal of Physics A: Mathematical and Theoretical* **42**, 415001 (2009).
- [20] I. Neri and D. Bollé, *Journal of Statistical Mechanics: Theory and Experiment* **2009**, P08009 (2009).
- [21] Y. Kanoria and A. Montanari, *The Annals of Applied Probability* **21**, 1694 (2011).
- [22] C. De Dominicis, *Phys. Rev. B* **18**, 4913 (1978).
- [23] J. P. L. Hatchett, I. P. Castillo, A. C. C. Coolen, and N. S. Skantzos, *Phys. Rev. Lett.* **95**, 117204 (2005).
- [24] A. Mozeika and A. Coolen, *Journal of Physics A: Mathematical and Theoretical* **42**, 195006 (2009).
- [25] M. Mézard and G. Parisi, *Eur. Phys. J. B* **20**, 217 (2001).
- [26] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, 2009).
- [27] B. Karrer and M. E. J. Newman, *Physical Review E* **82**, 016101 (2010).
- [28] E. Volz, *J. Math. Biol.* **56**, 293 (2008).
- [29] J. C. Miller, A. C. Slim, and E. M. Volz, *J. R. Soc. Interface* **9**, 890 (2012).
- [30] A. Lucas, arXiv:1206.6294 (2012).
- [31] A. Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborová, *Phys. Rev. E* **90**, 012801 (2014).
- [32] F. Altarelli, A. Braunstein, L. Dall’Asta, and R. Zecchina, *Physical Review E* **87**, 062115 (2013).

- [33] M. Shrestha and C. Moore, *Physical Review E* **89**, 022805 (2014).
- [34] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P09011 (2013).
- [35] D. G. Daley, D. J., & Kendall, *Nature* **204** (1964).
- [36] D. G. Daley, D. J., & Kendall, *IMA Journal of Applied Mathematics* **1(1)**, 42 (1965).
- [37] M. Maki, D. P., & Thompson, *Mathematical models and applications: with emphasis on the social, life, and management sciences* (Englewood Cliffs: Prentice-Hall, 1973).
- [38] Y. Yedidia, J. S., Freeman, W. T., & Weiss, *Exploring artificial intelligence in the new millennium* **8**, 236 (2003).
- [39] R. G. Gallager, *IEEE Trans. Inform. Theory* **8**, 21 (1962).
- [40] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference* (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988).
- [41] M. Mézard, G. Parisi, and R. Zecchina, *Science* **297**, 812 (2002).
- [42] R. J. Glauber, *Journal of mathematical physics* **4**, 294 (1963).
- [43] M. Bastian, S. Heymann, M. Jacomy, et al., *ICWSM* **8**, 361 (2009).
- [44] T. Opsahl and P. Panzarasa, *Social networks* **31**, 155 (2009).
- [45] W. W. Zachary, *Journal of Anthropological Research* **33**, 452 (1977).
- [46] A. Vespignani, *Nature Physics* **8**, 32 (2012).