Cavity master equation for the continuous time dynamics of discrete-spin models

E. Aurell,[1,2,3] G. Del Ferraro,[4,*] E. Domínguez,[5] and R. Mulet[5,†]

[1]*Department of Computational Biology, AlbaNova University Center, SE-106 91 Stockholm, Sweden*

[2]*Departments of Applied Physics and Computer Science, Aalto University, FIN-00076 Aalto, Finland*

[3]*Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

[4]*Department of Computational Biology, AlbaNova University Center, SE-106, 91 Stockholm, Sweden*

[5]*Group of Complex Systems and Statistical Physics. Department of Theoretical Physics, Physics Faculty, University of Havana, Cuba*

(Received 25 July 2016; published 12 May 2017)

We present an alternate method to close the master equation representing the continuous time dynamics of interacting Ising spins. The method makes use of the theory of random point processes to derive a master equation for local conditional probabilities. We analytically test our solution studying two known cases, the dynamics of the mean-field ferromagnet and the dynamics of the one-dimensional Ising system. We present numerical results comparing our predictions with Monte Carlo simulations in three different models on random graphs with finite connectivity: the Ising ferromagnet, the random field Ising model, and the Viana-Bray spin-glass model.

DOI: 10.1103/PhysRevE.95.052119

I. INTRODUCTION

Understanding complex systems with many interacting particles is at the forefront of the research activity in many scientific communities including statistical physics [1], chemical kinetics [2], population biology [3], neuroscience [4], and more. A general approach to gain proper insight in these systems is to develop simple models that, although composed of many interacting particles, can be treated analytically and/or computationally on reasonable time scales.

A prototypical example of this class of simple systems is the Ising ferromagnetic model whose equilibrium properties, despite still lacking a proper analytical solution in three dimensions, are globally well understood [5]. The addition of disorder to this model yields a more complicated scenario, but first the replica trick, together with the possibility of a replica symmetry breaking scheme [6–8], and then the later-introduced cavity approach [9,10] opened the doors to the analytical treatment of this and other models [11–13]. Provided the model is defined on fully connected or on random graphs, and leaving aside technical difficulties associated to each particular problem, there are proper tools to correctly understand the equilibrium properties of these families of disordered models [14]. Along this direction, only finite-dimensional systems continue to be elusive albeit some progress has been obtained in the last few years [15–18].

The situation is completely different once the interest turns to the dynamical properties of complex systems. First, because it is clear that approximations that work for short time scales are not necessarily valid at long time scales and vice versa. Second, because, regardless of the approximation, the selection of the local dynamical rules that define the processes turns out to be fundamental to characterize the evolution of the macroscopic quantities of the model [19,20]. Progress in this direction has been slower. The introduction of disorder further complicates the issue, generating new families of models and behaviors that still lack a complete understanding [21].

As for the equilibrium, a first classification of complex systems in the dynamical scenario can be made between continuous and discrete state variables. The dynamical modeling in the former case is usually done by using a Langevin equation [20] as it is possible to write the differential of a spin variable in a rigorous form. For the same reason, in these models time is in most of the cases considered a continuous variable. The dynamics of the fully connected (FC) spherical p-spin model, for instance, has been studied in Ref. [22] and solved by writing equations for the correlation and response function. The FC case $p = 2$ has been considered in Ref. [23] whereas random network architectures have been studied in Ref. [24] introducing a series of approximate equations.

For discrete variables, the discontinuous nature of the spin values makes it cumbersome, if not improper, to formulate the problem starting with Langevin-like equations. Thus, in this case, instead of writing differential equations for the spin variables, one describes the stochasticity of the dynamics by writing equations for the probability of the spin state. The literature has focused on two different possible choices for the dynamics: either time evolves in discrete steps, or in continuous time.

Furthermore, due to the possible relevance of these models to understand the dynamics in biological and neural networks, studies have focused also on cases with different network connectivity symmetries. The exact solution of a dilute fully asymmetric neural network model, for both parallel and asynchronous dynamics, dates back to Ref. [25]. Directed random graphs have been investigated by using different approaches. Path integral techniques for this case have been introduced in Ref. [26] for parallel dynamics on graphs with finite connectivity. The treelike structure of such graphs led then to the extension of equilibrium techniques such as the cavity method or belief-propagation message passing [14] to the dynamical scenario. The pioneering contribution [27], followed by Ref. [28], generalized the cavity technique to the parallel dynamics of the Ising model on a Bethe lattice but the approach was limited to stationary solutions. Random sequential updates rules were considered in Ref. [29] but also

*gino@kth.se

†mulet@fisica.uh.cu

limited to stationary states. In Ref. [30] a similar technique has been applied to study models with majority dynamics update rules (i.e., linear dynamics with thresholding). A generalization to nonstationary states of these techniques for models with reversible dynamics was made in Ref. [31], followed by Ref. [32] where further improvements of the dynamic reconstruction have been introduced at the price of a more involved formalism. More recently, variational approaches have shown to give accurate results for the transient dynamic of disordered spin systems by using a simple and systematic theory [33,34].

Another line of research that considers discrete state variables focuses on irreversible dynamics, as for instance the study of cascade processes or SIR models [35,36], and poses different types of challenges compared to those of the reversible dynamics studied in this paper and mentioned above.

The above historical overview shows that the field is active but also that progress has been limited to discrete-time update rules. The continuous-time counterpart has indeed been less investigated. In this case, a proper dynamic description of the spin-state configuration, follows a master equation (ME) for the probability density of the states of N-spin interacting variables [37,38]. Once the ME is defined one must select the rules of the dynamical evolution. Although this approach can be stated easily, the full solution of the master equation in the general case is a cumbersome task and exact solutions have been limited to simple models [37,39]. A way out of this issue has been proposed through the dynamical replica analysis for fully connected [40] and diluted graphs [41] where the authors, instead of looking for an estimation of the out-of-equilibrium probability of the spin state, derive dynamical equations for the probability of some macroscopic observables. Therefore, if, on one side, this approach obviously reduces the dimensionality of the problem, on the other, it loses detailed information about the microscopic state of the system.

In this paper we focus on continuous-time dynamics for discrete-spin variables and derive a set of closed equations for marginal probabilities of the microscopic spin configuration. In Sec. II we define the model in its more general form and present the main result of our paper, a cavity master equation (CME) stated as (5), and derived in Secs. III–V. The organization of the paper and the respective contributions of the following Secs. III–VII are outlined at the end of Sec. II.

II. MODEL DYNAMICS

We consider a system of N interacting discrete spin variables $\underline{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$, with $\sigma_i = \pm 1$, where the spins may flip spontaneously. The transition rate $r_i(\underline{\sigma})$ of having a spin flip for spin i at time t may in general depend on the instantaneous values of all the spin variables in the system. The master equation describes the evolution of the probability of the system to be in state $\underline{\sigma}(t)$ at time t as [37,38]

$$\frac{dP(\underline{\sigma})}{dt} = -\sum_{i=1}^{N}[r_i(\underline{\sigma})P(\underline{\sigma}) - r_i(F_i(\underline{\sigma}))P(F_i(\underline{\sigma}))], \qquad (1)$$

where we omitted the time dependence in $P(\underline{\sigma},t)$ to shorten notation and F_i represents the inversion operator on spin i, i.e., $F_i(\underline{\sigma}) = \{\sigma_1, \ldots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \ldots, \sigma_N\}$. Equation (1) is a first-order differential equation and encodes the Markov property that the future development of the system only depends on the present state of the system and not on the past history before the present. Although (1) is a simple equation to state formally, in practice it implies the daunting task of tracking the evolution in time of 2^N discrete states.

If the transition rate $r_i(\underline{\sigma})$ does not depend on the whole system state but only on the configuration of spin i and some neighbors ∂i, the master equation can be reduced to a local form. The evolution in time of the probability of the spin configuration σ_i is then obtained by tracing (1) over all the spin states except σ_i. The resulting equation reads

$$\frac{dP(\sigma_i)}{dt} = -\sum_{\sigma_{\partial i}}[r_i(\sigma_i,\sigma_{\partial i})P(\sigma_i,\sigma_{\partial i}) - r_i(-\sigma_i,\sigma_{\partial i})P(-\sigma_i,\sigma_{\partial i})], \qquad (2)$$

where $\sigma_{\partial i}$ stands for the configuration of all the spins in the neighborhood of i. As for $P(\underline{\sigma},t)$ in (1), in the equation above and hereafter we omit for brevity the time dependence of every probability distribution, all them should be taken at time t.

Contrary to (1), Eq. (2) is not closed. On the left-hand side we have the probability $P(\sigma_i)$ that spin i is in a particular state whereas, on the right-hand side, $P(\sigma_i,\sigma_{\partial i})$ stands for the probability of a certain configuration for spin i and its neighbors. To consistently describe the evolution of the single site probability (2) in time, we then have to search for a closure of this equation. Different approximations could be made on the joint probability $P(\sigma_i,\sigma_{\partial i})$ appearing on the right-hand side of (2). A reasonable one, that we will use through this work, is to assume that

$$P(\sigma_i,\sigma_{\partial i}) = \prod_{k\in\partial i} P(\sigma_k|\sigma_i)P(\sigma_i), \qquad (3)$$

which has the desirable property of being exact at equilibrium for trees and random graphs where loops are large compared to the system size. Assuming a treelike topology and the factorization in (3), the master equation (2) can then be written as

$$\frac{dP(\sigma_i)}{dt} = -\sum_{\sigma_{\partial i}}\left[r_i(\sigma_i,\sigma_{\partial i})\left[\prod_{k\in\partial i} P(\sigma_k|\sigma_i)\right]P(\sigma_i) - r_i(-\sigma_i,\sigma_{\partial i})\left[\prod_{k\in\partial i} P(\sigma_k|-\sigma_i)\right]P(-\sigma_i)\right]. \qquad (4)$$

The above equation is also not closed, as we do not know how $P(\sigma_k|\sigma_i)$ changes with time. The knowledge of a closed equation for this latter probability would allow us to describe the evolution of the probability distribution of the variable σ_i and of the conditional probability $P(\sigma_k|\sigma_i)$ at each time in the dynamics.

In the following sections we derive a master equation for a cavity probability distribution $p(\sigma_i|\sigma_k)$ that approximates $P(\sigma_i|\sigma_k)$ and under certain conditions is equal to it (as discussed in Appendix A). Our derivation is guided by general probabilistic principles and as a result we obtain a closure

scheme for $p(\sigma_i|\sigma_k)$. We here present the final result

$$\frac{dp(\sigma_i|\sigma_j)}{dt} = -\sum_{\sigma_{\partial i\setminus j}}\left[r_i[\sigma_i,\sigma_{\partial i}]\left[\prod_{k\in\partial i\setminus j} p(\sigma_k|\sigma_i)\right]p(\sigma_i|\sigma_j) - r_i[-\sigma_i,\sigma_{\partial i}]\left[\prod_{k\in\partial i\setminus j} p(\sigma_k|-\sigma_i)\right]p(-\sigma_i|\sigma_j)\right], \tag{5}$$

which we call the cavity master equation for reasons that will be clear below.

Equation (5) is the main result of our work and together with (4) provides a closed set of equations for the dynamics of the single site. We note that (5) can be obtained formally by conditioning both sides of (4) on the value of spin σ_j. A derivation starting from (2) and (4) of such a formal operation is not valid, as we discuss for completeness in Appendix B. The outcome of the analysis in Secs. III–V is that if the interactions are organized in a locally treelike geometry so that the dynamic cavity method can be applied, the result of the formal operation nevertheless holds. In Sec. VI we show that Eq. (5) reproduces analytical exact results for both the mean-field and the one-dimensional Ising ferromagnet. In Sec. VII we test the performances of the closure scheme (4) and (5) for different models defined on random graphs. The reader not interested in the analytical derivation of (5) can skip the following sections and continue from Sec. VII.

For completeness, we also note that Eq. (5) shares similarity with Eq. (10) of Ref. [27], for a discrete-time dynamics case, however, cannot be obtained as a continuous time limit of it. The main reason is that the basic cavity construction in that work is fundamentally different from ours. It is, nonetheless, an interesting point to check if starting from a similar discrete-time formalism may lead to the same result (5).

III. RANDOM POINT PROCESSES

In this section we introduce the random point process formalism, which will be used to parametrize probability distributions of spin histories in continuous time. To get familiar with the notation we first concentrate on just one independent spin.

The probability of having a single spin in state σ at time t given the initial condition σ_0 at time t_0, $p(\sigma,t|\sigma_0,t_0)$, can be specified by the sum of the probability weight of all trajectories that transform this initial state σ_0 into σ after a time $t-t_0$. A specific spin history or trajectory X is parametrized by the number of spin flips, the time in which they occur and the initial state of the system. The spin trajectory is then nothing but a random point process (RPP) [20,42,43] and the probability measure in this sample space may be denoted as

$$Q(X) = Q_s(t_0,t_1,\ldots,t_s,t|\sigma(t_0)=\sigma_0), \tag{6}$$

which represents the probability density of having a trajectory with s jumps at $(t_1,t_1+dt_1)\ldots(t_s,t_s+dt_s)$, etc. given the initial state σ_0. We here stress a detail on notation: there is a difference between $\sigma(\tau)$, which is the value of the spin orientation at the particular time τ and the complete trajectory X of such spin, which is specified by the value that $\sigma(\tau)$ takes for every time in the interval $[t_0,t]$. When needed, we may write X as $X(t)$ to emphasize that the final time of the spin history is precisely t.

To recover $p(\sigma,t|\sigma_0,t_0)$ from $Q(X)$, one integrates $Q_s(t_0,t_1,\ldots,t_s,t|\sigma(t_0)=\sigma_0)$ over all times for a fixed s and sums over all possible values of s. For example, in order to find $p(\sigma=\sigma_0,t|\sigma_0,t_0)$, the probability of having the same orientation at time t as in the initial state, we have to sum over all $s=2k$ possible values:

$$p(\sigma_0,t|\sigma_0,t_0) = Q_0 + \sum_{k=1}^{\infty}\int_{t_0}^{t} dt_1\int_{t_1}^{t} dt_2 \ldots \int_{t_{s-1}}^{t} dt_s\, Q_{2k}(t_0,\ldots,t). \tag{7}$$

In a nutshell, we sum the probability density of all trajectories with an even number of jumps because those, and only those, transform the initial state into itself.

If the spin flips occur independently at a rate that depends only on the instantaneous orientation σ and not on previous values, the interjump waiting time is exponentially distributed and

$$Q_{s=2k} = r(\sigma_0)e^{-r(\sigma_0)(t_1-t_0)}r(-\sigma_0)e^{-r(-\sigma_0)(t_2-t_1)} \ldots r(-\sigma_0)e^{-r(-\sigma_0)(t_s-t_{s-1})}e^{-r(\sigma_0)(t_f-t_s)}. \tag{8}$$

The term $r(\sigma)$ above is the jumping rate that, in this case of one single spin, only depends on the spin configuration σ at time t, whereas each exponential term accounts for the probability rate of remaining in a spin configuration for a time $t_s - t_{s-1}$.

If instead of a single spin we deal with the more general case of N interacting spins, the set of the individual histories $X_1,\ldots X_N$ can still be parametrized by the number of jumps of the corresponding spin s_i and the time coordinate of each trajectory, so $X_i \sim \{s_i,(t_1^i,t_2^i,\ldots,t_{s_i}^i)\} \sim \{s_i,\vec{t}^{\,i}\}$. Loosely speaking we may say that the system is at some time t in some global state $\underline{\sigma}(t)$ and can jump out of this state by flipping any of the spins σ_i. Each of these events happen at instantaneous rate $r_i(\underline{\sigma}(t))$ and the conditional probability that the state does not flip at all in a short time window $[t,t+\Delta]$ is hence $\exp(-\Delta t\sum_i r_i)$. Putting these pieces together the probability density on the collection of spin histories is

$$Q(X_1,\ldots,X_N) = \prod_{a=1}^{N}\left\{\prod_{l_a=1}^{s_a} r_a(\sigma_a(t_{l_a}),\sigma_{\partial a}(t_{l_a}))e^{-\int_{t_0}^{t} r_a[\sigma_a(\tau),\sigma_{\partial a}(\tau)])d\tau}\right\}. \tag{9}$$

Rigorous versions of this argument can be found in the mathematical literature, e.g., in Appendix A of the monograph of Kipnis and Landim [44].

We can write the quantity inside the curly brackets in (9) as $\Phi_a(X_a|X_{\partial a})$, since it represents the probability density of the history X_a with the histories of the neighborhood $X_{\partial a}$ fixed. Equation (9) can therefore be rewritten in a compact form as

$$Q(X_1,\ldots,X_N) = \prod_{a=1}^{N}\Phi_a(X_a|X_{\partial a}), \tag{10}$$

which will be used below.

IV. DYNAMIC MESSAGE PASSING EQUATIONS

In this section, starting from $Q(X_1,\ldots,X_N)$, we rederive the dynamic message passing (or dynamic cavity equation) as presented in Ref. [31] for the case of discrete-time update dynamics. This equation is an iterative relation for conditional probabilities of spin histories, exact on treelike graph in the thermodynamic limit. Then by using the RPP formalism we parametrize these probabilities and we give a proper mathematical description of how to marginalize them over time for the case of continuous-time processes. The combined use of the dynamic message-passing equation and the RPP formalism is our starting point to obtain Eq. (5).

Assuming a treelike architecture network, we start by selecting a spin, say i, and rewrite Eq. (10) expanding the tree around it and making use of its structure:

$$Q(X_1,\ldots,X_N) = \Phi_i(X_i|X_{\partial i}) \prod_{k\in\partial i}\left[\Phi_k(X_k|X_{\partial k}) \times \prod_{m\in\partial k\setminus i}\left(\Phi_m(X_m|X_{\partial m}) \prod_{l\in\partial m\setminus k}\ldots\right)\right]. \tag{11}$$

Let $G_k^{(i)}$ be the subgraph extended from the site k after removing the link (ik). We define as $\{X\}_{ik}$ the set of histories of the spins included in $G_k^{(i)}$ except X_k itself. With these definitions we express (11) as:

$$Q(X_1,\ldots,X_N) = \Phi_i(X_i|X_{\partial i}) \prod_{k\in\partial i} M_{ki}(X_k,\{X\}_{ik}|X_i). \tag{12}$$

Here M_{ki} is just a shorthand for the expression inside brackets, i.e.,

$$M_{ki}(X_k,\{X\}_{ik}|X_i) \equiv \Phi_i(X_i|X_{\partial i}) \prod_{k\in\partial i\setminus j}\left[\Phi_k(X_k|X_{\partial k}) \prod_{m\in\partial k\setminus i}\ldots\right]. \tag{13}$$

Marginalizing Q on all histories except $X_i, X_{\partial i}$ we get the joint (density) probability distribution of the history of spin i and its neighbors

$$Q(X_i,X_{\partial i}) = \Phi_i(X_i|X_{\partial i}) \prod_{k\in\partial i}\mu_{k\to(ki)}(X_k|X_i), \tag{14}$$

where the new functions $\mu_{k\to(ki)}(X_k|X_i)$ are the marginals of $M_{ki}(X_k,\{X\}_{ik}|X_i)$:

$$\mu_{k\to(ki)}(X_k|X_i) \equiv \sum_{\{X\}_{ik}} M_{ki}(X_k,\{X\}_{ik}|X_i) \tag{15}$$

and have the interpretation of the probability of history X_k holding X_i fixed. Let us observe that Eq. (11) can be rewritten by expanding the tree around two nodes, say i and j, as follows:

$$Q(X_1,\ldots,X_N) = M_{ji}(X_j,\{X\}_{ij}|X_i)\, M_{ij}(X_i,\{X\}_{ji}|X_j). \tag{16}$$

Marginalizing the above equation over all trajectories except X_i, X_j and using the definition of μ given in (15), we obtain

$$Q(X_i,X_j) = \mu_{i\to(ij)}(X_i|X_j)\mu_{j\to(ji)}(X_j|X_i). \tag{17}$$

The final step in order to derive the dynamic message-passing equation is to marginalize (14) on $X_{\partial i\setminus\{i,j\}}$ and combine with (17). Simplifying terms, we get

$$\mu_{i\to(ij)}(X_i|X_j) = \sum_{\{X_k\},k\in\partial i\setminus j} \Phi_i(X_i|X_{\partial i}) \prod_{k\in\partial i\setminus j}\mu_{k\to(ki)}(X_k|X_i), \tag{18}$$

where $X_i = X_i(t)$ is the history of spin i up to time t and the trace is over all the spin trajectories for $k \in \partial i\setminus j$, each term in the sum to be interpreted as

$$\sum_{X_k}(\cdot) = \sum_{s_k}\left[\int_{t_0}^{t} dt_1^k \int_{t_1^k}^{t} dt_2^k \ldots \int_{t_{s_k-1}^k}^{t} dt_{s_k}^k\,(\cdot)\right]. \tag{19}$$

Structurally (18) is what in the computer science and information theory literature is referred to as a belief propagation update equation. To simplify notation we will alternatively write $\mu_{i\to(ij)}(t)$ for the cavity conditional probability and, following the belief propagation literature [14,45], we may refer to it as a "message". In contrast to the use of belief propagation to compute marginals of Gibbs-Boltzmann distributions (18) is not very useful as it stands since the sum on the right-hand side is over a very high-dimensional object. Further approximations are therefore needed, which for the case of discrete dynamics were discussed in Ref. [31].

A. Differential update equations

In Eq. (9) we showed how to properly parametrize probability densities within the RPP formalism. The messages introduced through Eq. (15), according to the RPP framework [20,42,43], can then be written as

$$\mu_{i\to(ij)}(X_i(t)X_j(t)) = \prod_{l_i=1}^{s_i}\lambda_{i\to(ij)}(t_{l_i}) \exp\left\{-\int_{t_0}^{t}\lambda_{i\to(ij)}(\tau)d\tau\right\}, \tag{20}$$

where above t_{l_i} are the jumping times for the X_i history, which has s_i jumps. In this equation, $\lambda_{i\to(ij)}$ is interpreted as the jumping rate of i at each time, in a cavity where X_j is the history of spin j. In the more general case $\lambda_{i\to(ij)}$ is a function of the spin history of the variable i and j and thus, for a spin dynamics up to time τ we may more explicitly write $\lambda_{i\to(ij)}(\tau) = \lambda_{i\to(ij)}(X_i,X_j,\tau)$. On the other hand, the interaction term $\Phi_i(X_i(t)|X_{\partial i}(t))$ in (18), already introduced through (9) and (10), can be interpreted as the probability density of X_i conditioned on the histories of spins in ∂i. where r_i is the jumping rate of spin i. For a Markov dynamics this is an instantaneous quantity, meaning that at time τ it depends only on the values of $\sigma_i(\tau)$ and $\sigma_{\partial i}(\tau)$.

In principle, through the parametrization (20), if we write (18) for every pair (i,j) in the network we get a system of coupled equations for the λ's. Solving them we may describe the dynamics of the system. Unfortunately (18) is a very involved expression that needs to be transformed into a more tractable one. We here propose to differentiate it with respect

to the parameter t, the final time, in order to have a differential equation for the messages.

Differentiation in this context should be handled carefully since increasing t means we are changing the sample space itself. Writing $\mu_{i\to(ij)}(X_i(t+\Delta t)X_j(t+\Delta t))$ in terms of $\mu_{i\to(ij)}(X_i(t)X_j(t))$ is mapping the probability of one sample space (that of all possible histories up to time t) onto another (histories up to time $t+\Delta t$). Instead of using standard differentiation rules it is safer to go by the definition. Therefore, for the left-hand side of Eq. (18) we will compute the limit:

$$\lim_{\Delta t\to 0}\frac{\mu_{i\to(ij)}(X_i(t+\Delta t)X_j(t+\Delta t))-\mu_{i\to(ij)}(X_i(t)X_j(t))}{\Delta t}. \tag{21}$$

A very important question arises at this point. What is the relation of $[X_i(t+\Delta t),X_j(t+\Delta t)]$ and $[X_i(t),X_j(t)]$? Or in other words, what happens in the interval $(t,t+\Delta t)$? The answer is important because expressions for $\mu_{i\to(ij)}(X_i(t+\Delta t)X_j(t+\Delta t))$ are different whether we consider that there can be jumps in the small Δt interval or not. The first thing that makes sense to impose is that histories must agree up to time t. In $(t,t+\Delta t)$ we can have several combinations.

An implicit assumption throughout all this theory is that on an infinitesimal interval only two things can happen to a spin; it can stay on its current state or make one and only one jump to the opposite orientation. Two or more jumps are not allowed. Considering this we have four cases to analyze:

(i) There are s_i,s_j jumps in (t_0,t) and neither i nor j jumps in $(t,t+\Delta t)$. This occurs with a probability $(1-\lambda_i\Delta t)(1-\lambda_j\Delta t)$.

(ii) There are s_i,s_j jumps in (t_0,t) and i XOR j jumps in $(t,t+\Delta t)$. This occurs with a probability $(1-\lambda_i\Delta t)(\lambda_j\Delta t)$ or $(1-\lambda_j\Delta t)(\lambda_i\Delta t)$. These are two cases in one.

(iii) There are s_i,s_j jumps in (t_0,t) and both i and j jumps in $(t,t+\Delta t)$. This has a probability of $\lambda_j\lambda_i\Delta t^2$.

When Δt goes to zero, from the previous analysis we conclude that the derivative should be computed, with probability 1, using the first option, where histories for i and j have no jumps in the interval of length Δt.

To differentiate the left-hand side of (18) we use the parametrization (20). Writing shortly $\mu_{i\to(ij)}(t)$ for the message $\mu_{i\to(ij)}(X_i(t)X_j(t))$ at time t, we get

$$\begin{aligned}
&\mu_{i\to(ij)}(t+\Delta t)\\
&\quad=\prod_{l_i=1}^{s_i}\lambda_{i\to(ij)}(t_{l_i})\exp\left\{-\int_{t_0}^{t+\Delta t}\lambda_{i\to(ij)}(\tau)d\tau\right\}\\
&\quad\approx[1-\lambda_{i\to(ij)}(t)\Delta t]\prod_{l_i=1}^{s_i}\lambda_{i\to(ij)}(t_{l_i})\\
&\qquad\times\exp\left\{-\int_{t_0}^{t}\lambda_{i\to(ij)}(\tau)d\tau\right\}+O(\Delta t^2)\\
&\quad=[1-\lambda_{i\to(ij)}(t)\Delta t]\,\mu_{i\to(ij)}(t)+O(\Delta t^2).
\end{aligned}\tag{22}$$

Then, with probability 1, the derivative of the left-hand side of (18) is equal to $-\lambda_{i\to(ij)}(t)\,\mu_{i\to(ij)}(t)$.

To calculate the derivative of the right-hand side of Eq. (18), we should compute the limit of the difference quotient for this side of the equation. To shorten notation, we define F as the argument of the sum in (18), i.e., $F(X_i,X_{\partial i},t)\equiv\Phi_i(X_i|X_{\partial i})\prod_{k\in\partial i\setminus j}\mu_{k\to(ki)}(X_k|X_i)$. So, the trace on the right-hand side of this equation can be written as

$$\begin{aligned}
&\sum_{\{X_k\},k\in\partial i\setminus j}^{[t_0,t]}F(X_i,X_{\partial i},t)\\
&\quad=\sum_{\{s_k\},k\in\partial i\setminus j}\left[\prod_{k=1}^{d}\int_{t_0}^{t}dt_1^k\int_{t_1^k}^{t}dt_2^k\ldots\int_{t_{s_k-1}^k}^{t}dt_{s_k}^k\right]F(X_i,X_{\partial i},t)
\end{aligned}\tag{23}$$

and to obtain its derivative we must compute:

$$\lim_{\Delta t\to 0}\frac{\sum_{\{X_k\},k\in\partial i\setminus j}^{[t_0,t+\Delta t]}F(X_i,X_{\partial i},t+\Delta t)-\mu_{i\to(ij)}(X_i(t)X_j(t))}{\Delta t}. \tag{24}$$

Let us focus on the first term in the numerator of the above expression. It can be expanded to first order in Δt. It is important to remember that Δt appears in the integration limits as well as in the integrand of F. In addition, we should keep in mind that all jumps for X_i and X_j must occur before t. This restriction, however, does not apply to the histories X_k for k in $\partial i\setminus j$.

The expansion of (24) can be explained as follows. First, let us remember that F is the joint probability of X_i and $\{X_k\}$ with $k\in\partial i\setminus j$, conditioned on X_j. All the histories of interest are in the interval $[t_0,t+\Delta t]$. The expression:

$$\sum_{\{X_k\},k\in\partial i\setminus j}^{[t_0,t+\Delta t]}F(X_i,X_{\partial i},t+\Delta t) \tag{25}$$

is the marginalization of the mentioned joint probability distribution. The previous sum, to order Δt, has two main contributions. One comes from summing over $\{X_k\}$ with all X_k having no jumps in $[t,t+\Delta t]$:

$$\begin{aligned}
A=&\sum_{\{X_k\},\,k\in\partial i\setminus j}^{[t_0,t]}F(X_i,X_{\partial i},t)\\
&\times\left\{1-\left[\sum_k\lambda_{k\to(ki)}(t)+r_i(t)\right]\Delta t\right\}.
\end{aligned}\tag{26}$$

The other considers all the possibilities of having one of the X_k with a jump in the interval of length Δt:

$$B=\sum_k\sum_{\{X_k\},\,k\in\partial i\setminus j}^{[t_0,t]}F(X_i,X_{\partial i},t)\lambda_{k\to(ki)}(t)\Delta t. \tag{27}$$

The expansion to first order in Δt of Eq. (25) then reads

$$\sum_{\{X_k\},k\in\partial i\setminus j}^{[t_0,t+\Delta t]}F(X_i,X_{\partial i},t+\Delta t)=A+B+O(\Delta t^2). \tag{28}$$

Writing the right-hand side of the above equation explicitly we observe that B cancels out with the λ part of A in (26) and the remaining term of order 1 is $\mu_{i\to(ij)}(X_i(t)X_j(t))$, which

cancels when inserted in the limit expression. Then, the final outcome of the differentiation of Eq. (18) reads

$$\lambda_{i\to(ij)}(X_i,X_j,t)\,\mu_{i\to(ij)}(X_i|X_j) = \sum_{\{X_k\},k\in\partial i\setminus j}^{[t_0,t]} r_i(\sigma_i(t),\sigma_j(t),\sigma_{\partial i\setminus j}(t))F(X_i,X_{\partial i},t). \quad (29)$$

We can now marginalize the right-hand side of the above equation of all the spin histories of the spins $k\in\partial i\setminus j$ by keeping the configuration of these spins at the last time t fixed. The results reads

$$\lambda_{i\to(ij)}(X_i,X_j,t)\,\mu_{i\to(ij)}(X_i|X_j) = \sum_{\sigma_{\partial i\setminus j}(t)} r_i(\sigma_i(t),\sigma_j(t),\sigma_{\partial i\setminus j}(t))p(\sigma_{\partial i\setminus j}(t),X_i|X_j), \quad (30)$$

where we introduced the function p as the marginalization of the function F over all the spin histories of the i neighbors except j, with the configuration at the final time fixed. Note that the notation $\sigma_{\partial i\setminus j}(t)$ is equivalent to $\{\sigma_k(t)\}_{k\in\partial i\setminus j}$ and that in p above appears again explicitly the conditional nature of the probability distribution F.

Equation (30) represents the differential dynamic message-passing update equation obtained by differentiating (18) in time. It connects the derivative of the dynamic message $\mu_{i\to(ij)}$, and so the jumping rate $\lambda_{i\to(ij)}$ of spin i in the cavity used to parametrize the message in (20), with the transition rate of the same spin $r_i(\sigma_i(t),\sigma_j(t),\sigma_{\partial i\setminus j}(t))$ at time t. This result will be used in next section for our final derivation. In what follows we will use the same symbol p for different probability densities and distinguish them only on the basis of their arguments.

V. CAVITY MASTER EQUATION

The cavity message $\mu_{i\to(ij)}(X_i|X_j)$ is a complicated object with high dimensionality. It is a real valued functional of X_i given the history X_j, where both μ and X depend parametrically on t. For our purposes, it is convenient to reduce the dimensionality of this message by partially marginalizing over the spin history of spin i. We therefore introduce an easier mathematical object to deal with, which is the marginal of $\mu_{i\to(ij)}(X_i|X_j)$ with the final state fixed

$$p(\sigma_i,t|X_j) = \sum_{X_i|\sigma_i(t)=\sigma_i} \mu_{i\to(ij)}(X_i|X_j). \quad (31)$$

By differentiating the above equation we can obtain an equation for the evolution of this probability distribution. As we have seen in the previous section, the derivative must be computed by using the standard definition of the limit of the increment ratio:

$$\dot{p}(\sigma_i,t|X_j) = \lim_{\Delta t\to 0}\frac{p(\sigma_i,t+\Delta t|X_j(t+\Delta t)) - p(\sigma_i,t|X_j(t+\Delta t))}{\Delta t}. \quad (32)$$

We hereafter write $p(\sigma_i,t+\Delta t|X_j(t+\Delta t))$ as the marginalization of $\mu_{i\to(ij)}(X_i(t+\Delta t)|X_j(t+\Delta t))$ and then, following the procedure developed in the last section, we expand it to first order in Δt, similarly to what we did for (24). Using the shorthand notation for μ, we get:

$$p(\sigma_i,t+\Delta t|X_j(t+\Delta t)) = \sum_{s=0}^{\infty}\int_{t_0}^{t+\Delta t}dt_1\int_{t_1}^{t+\Delta t}dt_2\ldots\int_{t_{s-1}}^{t+\Delta t}dt_s\,\mu_{i\to(ij)}(t+\Delta t). \quad (33)$$

Let us call the series of the s integrals above as $I_{t+\Delta t}$ and note that they can be written separating the $O(\Delta t^2)$ terms as follows

$$I_{t+\Delta t} \doteq \int_{t_0}^{t+\Delta t}dt_1\int_{t_1}^{t+\Delta t}dt_2\ldots\int_{t_{s-1}}^{t+\Delta t}dt_s = \int_{t_0}^{t}dt_1\int_{t_1}^{t}dt_2\ldots\int_{t_{s-1}}^{t+\Delta t}dt_s + O(\Delta t^2) \quad (34)$$

and that, furthermore, the last one can be split into two intervals from $[t_s,t]$ and $[t,t+\Delta t]$

$$I_{t+\Delta t} = \int_{t_0}^{t}dt_1\int_{t_1}^{t}dt_2\ldots\int_{t_{s-1}}^{t}dt_s + \int_{t_0}^{t}dt_1\int_{t_1}^{t}dt_2\ldots\int_{t}^{t+\Delta t}dt_s + O(\Delta t^2). \quad (35)$$

The first term in (35) is a trace over trajectories that have all jumps in $[t_0,t]$; the second term considers that the sth jump occurs in $[t,t+\Delta t]$. We then insert $\mu_{i\to(i,j)}(t+\Delta t)$ into $I_{t+\Delta t}$ expressed as above in order to compute (33). Let us observe that in the first integral $\mu_{i\to(ij)}(t+\Delta t)$ should be expanded to first order in Δt, whereas in the second integral it can be left to the order zero expansion since the integral has an order Δt itself.

In compact notation the expansion of $\mu_{i\to(i,j)}(X_i(t+\Delta t)|X_j(t+\Delta t))$ reads:

$$\mu_{i\to(i,j)}(t+\Delta t) = \left[\prod_{i=1}^{s}\lambda(t_i)\right]e^{-\int_{t_0}^{t+\Delta t}\lambda(\tau)d\tau} = \mu_{i\to(i,j)}(t)(1-\lambda(t)\Delta t + O(\Delta t^2)), \quad (36)$$

where hereafter we omit the subscript of λ to further shorten notation. Adding all together these results as described above we obtain:

$$p(\sigma_i,t+\Delta t|X_j(t+\Delta t)) = \sum_{X_i|\sigma_i(t)=\sigma_i}[1-\lambda(t)\Delta t]\mu_{i\to(i,j)}(X_i(t)|X_j(t)) + \sum_{s}\int_{t_0}^{t}dt_1\int_{t_1}^{t}dt_2\ldots\int_{t}^{t+\Delta t}dt_s\left[\prod_{i=1}^{s}\lambda(t_i)\right]e^{-\int_{t_0}^{t}\lambda(\tau)d\tau} + O(\Delta t^2). \quad (37)$$

As previously mentioned, the last integral on the second term of (37) corresponds to the probability of having the last jump s in the interval $[t,t+\Delta t]$. Since the orientation before the last jump is $-\sigma_i$, the corresponding jumping rate in this case is $\lambda(t)=\lambda(-\sigma_i(t),X_i^-,X_j)$, where to stress this difference we separated the value of the last spin of i at time t from its previous history and X_i^- denotes that the final state of this

history is $-\sigma_i$. With this notation, the last integral in (37) can be expanded as

$$\begin{aligned}
&\int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \ldots \int_{t}^{t+\Delta t} dt_s \left[\prod_{i=1}^{s} \lambda(t_i)\right] e^{-\int_{t_0}^{t} \lambda(\tau) d\tau} \\
&= \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \ldots \int_{t_{s-2}}^{t} dt_{s-1} \left\{ \left[\prod_{i=1}^{s-1} \lambda(t_i)\right] e^{-\int_{t_0}^{t} \lambda(\tau) d\tau} \lambda(-\sigma_i(t), X_i^-, X_j) \right\} \Delta t + O(\Delta t^2) \\
&= \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \ldots \int_{t_{s-2}}^{t} dt_{s-1} \{\mu_{i\to(i,j)}[X_i^-(t)|X_j(t)] \lambda(-\sigma_i(t), X_i^-, X_j)\} \Delta t + O(\Delta t^2). \qquad (38)
\end{aligned}$$

We highlight that the histories X_i^- and X_j as arguments of λ above run up to time t and, as mentioned, the superscript in X_i^- is included to explicitly state that the corresponding integrals are tracing histories of i with $s-1$ jumps and it means that the last state is the opposite to $\sigma_i(t)$.

Combining together the results in (33), (37), and (38) and taking the limit $\Delta t \to 0$ we get that the derivative expressed in (32) reads

$$\dot{p}(\sigma_i, t|X_j) = -\sum_{X_i|\sigma_i(t)=\sigma_i} \lambda_{i\to(ij)}[\sigma_i(t), X_i, X_j]\ \mu_{i\to(ij)}(X_i|X_j) + \sum_{X_i^-|\sigma_i(t)=-\sigma_i} \lambda_{i\to(ij)}[-\sigma_i(t), X_i^-, X_j]\ \mu_{i\to(ij)}(X_i^-|X_j). \qquad (39)$$

Finally, plugging (30) into (39) and after some rearrangement:

$$\begin{aligned}
\dot{p}(\sigma_i, t|X_j) &= -\sum_{\sigma_{\partial i\setminus j}} \left[\sum_{X_i|\sigma_i(t)=\sigma_i} r_i[\sigma_i, \sigma_{\partial i}] p(\sigma_{\partial i\setminus j}, X_i|X_j) - \sum_{X_i^-|\sigma_i(t)=-\sigma_i} r_i[-\sigma_i, \sigma_{\partial i}] p(\sigma_{\partial i\setminus j}, X_i^-|X_j) \right] \qquad (40) \\
&= -\sum_{\sigma_{\partial i\setminus j}} [r_i[\sigma_i, \sigma_{\partial i}] p(\sigma_{\partial i\setminus j}, \sigma_i|X_j) - r_i[-\sigma_i, \sigma_{\partial i}] p(\sigma_{\partial i\setminus j}, -\sigma_i|X_j)], \qquad (41)
\end{aligned}$$

where in the second equality we performed the traces inside the brackets.

To arrive at the result expressed in Eq. (41) we only assumed that the graph topology is treelike. In what follows, in order to finally derive Eq. (5), we will make two further approximations.

Approximation I. The conditional distributions are short range and factorize:

$$\begin{aligned}
p(\sigma_{\partial i\setminus j}, \sigma_i|X_j) &= p(\sigma_{\partial i\setminus j}|\sigma_i, X_j) p(\sigma_i|X_j) \\
&\approx p(\sigma_{\partial i\setminus j}|\sigma_i) p(\sigma_i|X_j) \approx \left[\prod_{k\in\partial i\setminus j} p(\sigma_k|\sigma_i)\right] p(\sigma_i|X_j). \qquad (42)
\end{aligned}$$

This first approximation is equivalent to say that there are short-length correlations between the variables and the stories of their neighbors. Indeed $p(\sigma_{\partial i\setminus j}|\sigma_i, X_j) = p(\sigma_{\partial i\setminus j}|\sigma_i)$ assumes that the states of the neighbors of i different from j do not depend on the story of j. The factorization that follows is an immediate consequence of the treelike structure of the system. Indeed, the approximations in the equation above become very accurate on a treelike topology, in the thermodynamic limit. Consider, for this network architecture, a node i and its neighbors ∂i. The influence of j on the other i neighbors $\{\partial i\setminus j\}$ can propagate only through either i itself or through loops of $O(\log N)$. Therefore, when σ_i is conditioned as in Eq. (42), $p(\sigma_{\partial i\setminus j}|X_i, X_j) \approx p(\sigma_{\partial i\setminus j}|X_i)$. Similarly, since the nodes $\partial i\setminus j$ are connected to each other through the node i, conditioning on X_i makes its neighbors probabilistically independent, as in the last approximation of Eq. (42). On a more general topology, the accuracy of the above approximations is a matter of local degree connectivity and presence of short loops and so, in other words, of the correlation regime among spin variables.

Approximation II. The probability distribution of an instantaneous variable, conditioned on the history of a neighbor, depends only on the last state of that neighbor:

$$p(\sigma_i|X_j) \approx p(\sigma_i|\sigma_j). \qquad (43)$$

The nature of this assumption is that of a short memory, it is equivalent to *de facto* reintroducing the Markov hypothesis for the local equation of the cavity probability distribution, and in that sense similar to the Markov assumption introduced by two of us for the discrete time case in Ref. [31] and more recently improved to consider longer temporal correlations by Ref. [32].

Combining these two approximation into Eq. (42) leads to the cavity master equation, already anticipated in (5):

$$\frac{dp(\sigma_i|\sigma_j)}{dt} = -\sum_{\sigma_{\partial i\setminus j}} \left[r_i[\sigma_i, \sigma_{\partial i}] \left[\prod_{k\in\partial i\setminus j} p(\sigma_k|\sigma_i)\right] p(\sigma_i|\sigma_j) - r_i[-\sigma_i, \sigma_{\partial i}] \left[\prod_{k\in\partial i\setminus j} p(\sigma_k|-\sigma_i)\right] p(-\sigma_i|\sigma_j) \right]. \qquad (44)$$

The relations between the above conditional probability density and the probability density appearing in Eq. (4) are discussed in Appendix A.

VI. COMPARISON WITH KNOWN EXACT RESULTS

In this section we show that the cavity master equation (44) compares to exact results derived for two specific models, well studied in the literature. We show that the mean-field (fully connected) ferromagnet is well described with the help of the CME. We also use the exactly solvable one-dimensional lattice to illustrate that (44) gives the right result in a low correlation limit.

A. Mean-field ferromagnet with Glauber dynamics

In this section we assume a fully connected graph topology in the thermodynamic limit. We start considering the cavity master equation (44) and assuming that the Glauber transition rate [37]

$$r_i(\sigma_i,\sigma_{\partial i}) = \frac{\alpha}{2}\left[1-\sigma_i \tanh\left(\beta\frac{J}{N}\sum_{k\neq i}\sigma_k\right)\right], \tag{45}$$

where all the spins are considered to be at time t. One can then introduce a δ function for the variable $m=\frac{1}{N}\sum_{k\neq i}\sigma_k$, and use its integral representation to rewrite the right-hand side of (44) as:

$$\begin{aligned}\dot{p}(\sigma_i|\sigma_j) = &-\frac{1}{2\pi}\int d\hat{m}\, dm\, e^{-i\hat{m}m}\, r_i(\sigma_i,m)\, e^{i\hat{m}\frac{\sigma_j}{N}} \\ &\times\left[\sum_{\sigma_k} e^{i\hat{m}\frac{\sigma_k}{N}} p(\sigma_k|\sigma_i)\right]^{N-2} p(\sigma_i|\sigma_j) \\ &+\frac{1}{2\pi}\int d\hat{m}\, dm\, e^{-i\hat{m}m} r_i(-\sigma_i,m) e^{i\hat{m}\frac{\sigma_j}{N}} \\ &\times\left[\sum_{\sigma_k} e^{i\hat{m}\frac{\sigma_k}{N}} p(\sigma_k|-\sigma_i)\right]^{N-2} p(-\sigma_i|\sigma_j).\end{aligned} \tag{46}$$

Where we assumed that the initial state is homogeneous and the probability density $p(\sigma_k|\sigma_i)$ does not depend explicitly on k and i. Note now that the term within brackets can be expanded as:

$$\begin{aligned}\left[\sum_{\sigma_k} e^{i\hat{m}\frac{\sigma_k}{N}} p(\sigma_k|\sigma_i)\right]^{N-2} = &\sum_{k=0}^{N-2} C_{N-2,k} e^{i\hat{m}\left(\frac{2(k+1)}{N}-1\right)} \\ &\times p(1|\sigma_i)^k\, p(-1|\sigma_i)^{N-k-2},\end{aligned} \tag{47}$$

which substituted in (46) allows a simple integration over $\hat{m}$ such that now:

$$\begin{aligned}\dot{p}(\sigma_i|\sigma_j) = &-\int dm P(m,t|\sigma_i,\sigma_j) r_i(\sigma_i,m) p(\sigma_i|\sigma_j) \\ &+\int dm P(m,t|-\sigma_i,\sigma_j) r_i(-\sigma_i,m) p(-\sigma_i|\sigma_j),\end{aligned} \tag{48}$$

where $P(m,t|\sigma_i,\sigma_j) = \sum_k C_{N-2,k} p^k(1|\sigma_i) p^{N-k-1}(-1|\sigma_i)\delta\left(m-\frac{\sigma_j}{N}-\left(\frac{2(k+1)}{N}-1\right)\right)$.

One can then define $m_i(\sigma_j)=\sum_{\sigma_i} p(\sigma_i|\sigma_j)\sigma_i$ as the magnetization of spin i at time t, provided that the spin j is in state σ_j. By defining $P(m,t|\sigma_j)=\sum_{\sigma_i} P(m,t|\sigma_i,\sigma_j)p(\sigma_i|\sigma_j)$, it is then straightforward to show that:

$$\dot{m}_i(\sigma_j) = -\alpha\, m_i(\sigma_j) + \alpha\int dm\, P(m,t|\sigma_j)\, \tanh(\beta J m). \tag{49}$$

A further average of (49) over the single spins (not including j) then gives:

$$\dot{\bar{m}}(\sigma_j) = -\alpha\,\bar{m}(\sigma_j) + \alpha\int dm\, P(m,t|\sigma_j)\, \tanh(\beta J m), \tag{50}$$

which looks similar to the standard results of the literature, although it may have a more subtle interpretation due to the conditional distribution on the spin σ_j. One can proceed assuming that in the thermodynamic limit the importance of the state of one individual spin at time t is negligible and that the average magnetization is defined by only one possible trajectory, recovering the more standard result:

$$\dot{\bar{m}} = -\alpha\,\bar{m} + \alpha\, \tanh(\beta J \bar{m}). \tag{51}$$

B. One-dimensional ferromagnet with Glauber dynamics

One-dimensional (1D) lattices are often a good starting point to search for analytical results that could be used to understand the nature and limits of approximations made in more involved contexts. In this section we consider the 1D ferromagnet using the CME (44) and compare this solution to Glauber's exact result [37].

Let us start by quoting the exact results. Observe that the transition rate (45) can be written for this 1D case as

$$\begin{aligned}r_i(\sigma_i,\sigma_{\partial i}) &= \frac{\alpha}{2}\{1-\sigma_i \tanh[\beta J(\sigma_{i-1}+\sigma_{i+1})]\} \\ &= \frac{\alpha}{2}\left[1-\frac{\sigma_i}{2}(\sigma_{i-1}+\sigma_{i+1})\tanh(2\beta J)\right].\end{aligned} \tag{52}$$

Note that we only have two neighbors of spin i, therefore $\partial i=\{i-1,i+1\}$. This rate, put into the exact ME (2), gives the following closed set of equations for the single site magnetizations:

$$\dot{m}_i = -\alpha\, m_i + \frac{\alpha}{2}\tanh(2\beta J)(m_{i-1}+m_{i+1}) \tag{53}$$

and equivalently for the probabilities:

$$\begin{aligned}\dot{p}(\sigma_i) &= A(m_i,m_{i-1},m_{i+1})\sigma_i \\ &= -\frac{\alpha}{2}\left[m_i - \frac{\tanh(2\beta J)}{2}(m_{i-1}+m_{i+1})\right]\sigma_i.\end{aligned} \tag{54}$$

If we now average (53) over all the sites i in the graph, in order to get the equation for the global magnetization, we obtain

$$\dot{m} = -\alpha\,(1-\tanh(2\beta J))m \tag{55}$$

whose solution is $m(t)=m(0)e^{-\gamma t}$, with $\gamma = 1-\tanh(2\beta J)$. Equations (53), (54), and (55) correspond to the exact results found by Glauber for the evolution of the magnetization of an Ising chain.

Now let us see what would be the dynamics if, instead of the exact ME, we considered the cavity master equation equation (44) as a starting point in our analysis. Substituting in (44) the transition rate expressed in (52) reads

$$\frac{dp(\sigma_i|\sigma_{i+1})}{dt} = -\sum_{\sigma_{i-1}}\left[\frac{\alpha}{2}\left[1-\frac{\sigma_i}{2}(\sigma_{i-1}+\sigma_{i+1})\tanh(2\beta J)\right]p(\sigma_{i-1}|\sigma_i)p(\sigma_i|\sigma_{i+1})\right.$$
$$\left.-\frac{\alpha}{2}\left[1+\frac{\sigma_i}{2}(\sigma_{i-1}+\sigma_{i+1})\tanh(2\beta J)\right]p(\sigma_{i-1}|-\sigma_i)p(-\sigma_i|\sigma_{i+1})\right]. \quad (56)$$

Now we can multiply both sides of the above equation by σ_i and marginalize over it. Similarly to what is done for the mean-field ferromagnet, we define $m_i(\sigma_{i+1}) = \sum_{\sigma_i} p(\sigma_i|\sigma_{i+1})\sigma_i$ to be the magnetization of spin i at time t, given that the spin $i+1$ is in state σ_{i+1}. Separating terms the above equation then reads

$$\dot{m}_i(\sigma_{i+1}) = -\sum_{\sigma_i,\,\sigma_{i-1}}\left[\frac{\alpha}{2}\sigma_i(p(\sigma_{i-1}|\sigma_i)p(\sigma_i|\sigma_{i+1}) - p(\sigma_{i-1}|-\sigma_i)p(-\sigma_i|\sigma_{i+1}))\right.$$
$$\left.-\frac{\alpha}{4}\tanh(2\beta J)(\sigma_{i-1}+\sigma_{i+1})(p(\sigma_{i-1}|\sigma_i)p(\sigma_i|\sigma_{i+1}) + p(\sigma_{i-1}|-\sigma_i)p(-\sigma_i|\sigma_{i+1}))\right]. \quad (57)$$

Marginalizing we get

$$\dot{m}_i(\sigma_{i+1}) = -\alpha\, m_i(\sigma_{i+1}) + \frac{\alpha}{2}\tanh(2\beta J)(m_{i-1}(\sigma_{i+1}) + \sigma_{i+1}). \quad (58)$$

This approximate equation for the local magnetization reflects the structure of the equation (53) deduced starting from the master equation, but it has a different meaning. As already said the quantity $m_i(\sigma_{i+1})$ should be interpreted as the magnetization of spin i, conditioned to spin $i+1$ being in state σ_{i+1}.

To compare to the exact results we need the real magnetization m_i. It is simply related to $m_i(\sigma_{i+1})$ through $m_i = \sum_{\sigma_{i+1}} m_i(\sigma_{i+1})p(\sigma_{i+1})$. The time derivative of m_i is:

$$\dot{m}_i = \sum_{\sigma_{i+1}}[\dot{m}_i(\sigma_{i+1})p(\sigma_{i+1}) + m_i(\sigma_{i+1})\dot{p}(\sigma_{i+1})]. \quad (59)$$

We can use (58) and (54) in the equation above and obtain:

$$\dot{m}_i = -\alpha m_i + \frac{\alpha}{2}\tanh(2\beta J)(m_{i-1}+m_{i+1}) + A(m_i,m_{i+1},m_{i+2})[m_i(\sigma_{i+1}=1) - m_i(\sigma_{i+1}=-1)]. \quad (60)$$

The first line in (60) is the exact result. The second line is an extra term that goes to zero for weak correlation or high temperature because $m_i(\sigma_{i+1})$ becomes independent of σ_{i+1}. Therefore we can conclude that our CME is consistent with exact results when appropriate limits are taken. For models with higher connectivities we may expect some agreement with numerical simulations even for not so high temperatures since a higher number of neighbors translates directly into a weaker correlation between spins.

VII. NUMERICAL RESULTS

To test numerically the performance of the approximate dynamics described by (4) and (5) we compare the numerical solutions of this set of equations with the results obtained after running a large number of kinetic Monte Carlo simulations. The numerical protocol is described below.

Three typical models are considered: an Ising ferromagnet with zero external magnetic field, the same ferromagnet with disordered local magnetic field [also known as random field Ising model (RFIM)] and the Viana-Bray spin-glass model, where interaction constants $J_{ij} = \pm 1$ are drawn positive or negative with equal probability from a bimodal distribution. All three systems share the same underlying topology of an instance of an Erdös-Rényi random graph, generated with $N = 1000$ nodes and mean connectivity $c = 3$. The rate of change for individual spins is taken according to Glauber's rule:

$$r_i(\sigma_i,\sigma_{\partial i}) = \frac{\alpha}{2}\left(1-\sigma_i\tanh\left[\beta\left(\sum_{j\in\partial i}J_{ij}\sigma_j + h_i\right)\right]\right), \quad (61)$$

where α is a constant defining the time unit, $t_0 = 1/\alpha$. For the actual simulations, the interaction constants are rescaled by a factor $1/c$. Monte Carlo simulations are run on a (typical) specific realization of the graph (single instance) and a sweep is composed by N sequential MC single-spin updates. The new state probability $W(\sigma_i|\sigma_{\partial i})$ used in the update rule is determined from the transition rates: $W(\sigma_i|\sigma_{\partial i}) = r_i(-\sigma_i,\sigma_{\partial i})/\alpha$. Final results correspond to an average over 10^5 MC trajectories, obtained starting from the same initial conditions and with a different random sequence for each run.

The solution of the CME is not substantially affected by finite-size character of available simulations. A similar experiment using 10^2 spins was performed, obtaining essentially the same results for global quantities. This behavior is natural given the local character of the approximation. We may say that even for rather small systems, the CME behaves as in the thermodynamic limit. On the other hand, kinetic MC simulations are indeed affected by the system size; the smaller the system the more likely a transition between different basins of attraction of the dynamics.

The CME for the conditional probabilities (5) is solved using Euler's method for ODEs. The integration step size h is a fraction of this time unit, $\Delta t = 0.05\, t_0$. Initial conditions, both for the differential equations and the stochastic simulations,

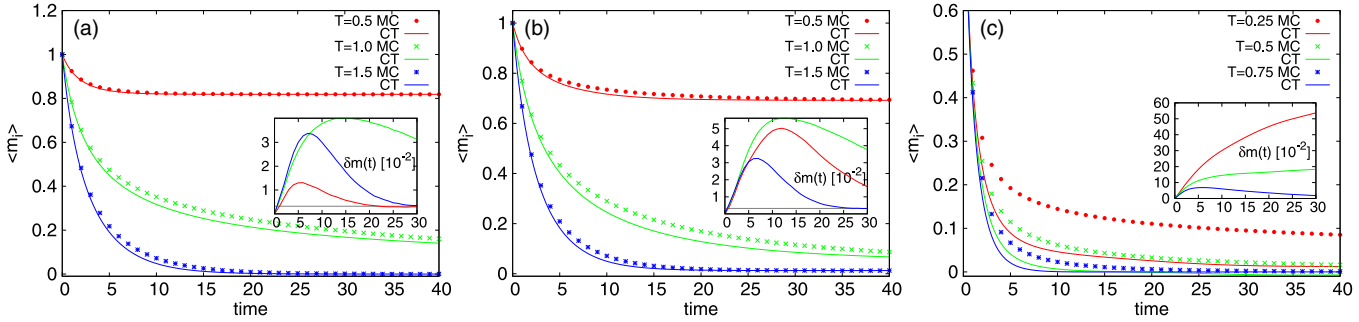

FIG. 1. Evolutionary dynamics of the global magnetization parameter for three spin models defined on a single instance of a random graph. (a) the Ising ferromagnet ($T_c \approx 0.96$). (b) the RFIM ($T_c \approx 0.78$). (c) the Viana-Bray model ($T_{SG} = 0.506$). Dots represent kinetic MC simulation on a single instance realization of an ER graph, averaged over 10^5 runs starting with the same initial conditions. Solid lines represent the cavity master equation approach presented in the main text. The insets show the mean error in local magnetization, as defined in the main text.

correspond to a frozen state with all spins pointing in the same (positive) direction. At finite temperature this is not an equilibrium state and the system will evolve and relax towards it. Once integrated, the derived equations give the time evolution of conditional pairwise probabilities, but for observables we need complete probability distributions. These can then be obtained by integrating the factorized master equation for a local variable (4) where the conditional probabilities that appear in the previous equation are given by the solution of (5). For a fixed length of the total time interval to analyze, $[t_0,t]$, the running time of the numerical integration is inversely proportional to the time update step size h. A typical computational time value for the results shown here is in the order of a few minutes on a desktop PC. On the other hand, for the equivalent numerical accuracy, the MC simulation time was much longer; in the order of several hours. Therefore, when local marginals are the relevant quantity to determine and for a model where CME performs well, it is a convenient choice to use this method rather than MC.

Starting with local probability distributions, local magnetizations are defined as usual, $m_i(t) = \sum_{\sigma_i(t)} \sigma_i(t)\, p(\sigma_i(t))$, where $p(\sigma_i(t))$ is estimated by (4) and (5). Global magnetization is computed as the average of local ones over the network $m(t) = \frac{1}{N}\sum_i m_i(t)$. For disordered systems it is also useful to investigate the evolution of the Edwards-Anderson (EA) parameter, defined as the average of the squared local magnetization $q_{EA}(t) = \frac{1}{N}\sum_i m_i^2(t)$.

Figures 1 and 2 show, respectively, the relaxation of global magnetization $m(t)$ and $q_{EA}(t)$ for our three test models, using MC simulations (dots) and the CME formalism (lines). The insets include the mean error of local magnetization with respect to the MC predictions, $\delta m = \sqrt{\frac{1}{N}\sum_i [m_i^{CME}(t) - m_i^{MC}(t)]^2}$. For the EA parameter we define the equivalent error measure.

The CME approach for the ferromagnetic case shown in Figs. 1(a) and 2(a) displays a good agreement with MC simulations for both the transient regime and the long time behavior at temperatures above and below the critical one of this model, i.e., $T_c \approx 0.96$. For a value $T = 1$, quite close to the phase transition, the qualitative behavior of the magnetization dynamics is fairly reproduced but its accuracy, as expected, diminishes. An important part of the procedure leading to (5) relies on the factorization of distributions and this is equivalent to set almost all (connected) correlations to zero. It is therefore natural to find a failure near a region where correlations become fundamental, as it is the case for a second-order phase transition.

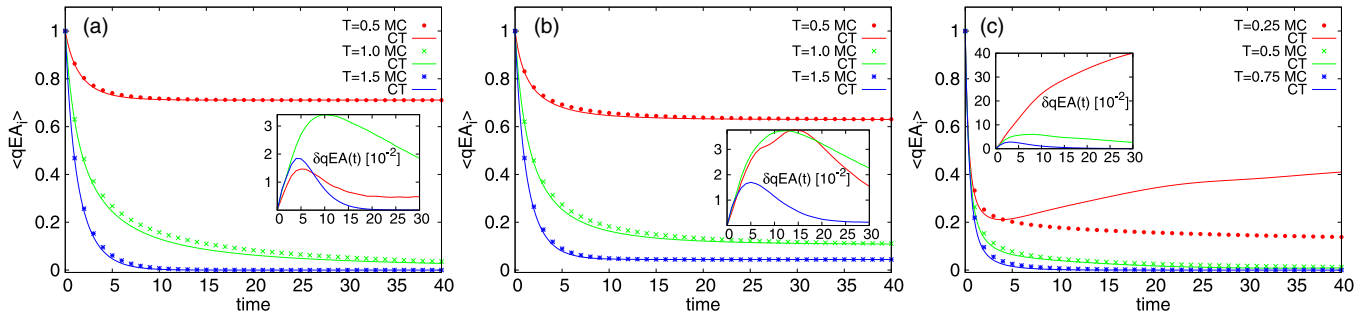

FIG. 2. Dynamics of the global EA parameter for three spin models defined on a single instance of a random graph. (a) The Ising ferromagnet ($T_c \approx 0.96$). (b) The RFIM ($T_c \approx 0.78$). (c) The Viana-Bray model ($T_{SG} = 0.506$). Dots represent kinetic MC simulations on a single instance realization of an ER graph, averaged over 10^5 runs starting with the same initial conditions. Solid lines represent the cavity master equation approach presented in the main text. The behavior for low temperatures in the Viana-Bray model shows that even though global magnetization is close to zero for long times [see Fig. 1(c)], spins are locally magnetized for long times, as it is expressed by the non zero value of $q_{EA}(t)$ for $T = 0.25$. The insets show the mean error of local quantities, as defined in the main text.

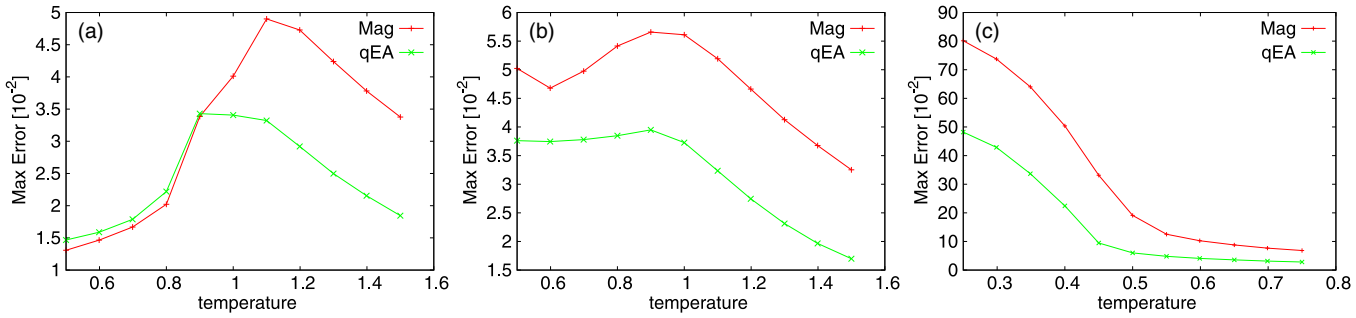

FIG. 3. Maximum mean error dependence with temperature between MC simulations and the CME approach presented in the main text for three spin models defined on a single instance of a random graph. (a) The Ising ferromagnet ($T_c \approx 0.96$). (b) The RFIM ($T_c \approx 0.78$). (c) The Viana-Bray model ($T_{SG} = 0.506$). Simulation details are the same as for Figs. 1 and 2. For the ferromagnet and the RFIM, error increases before and decreases after the ferromagnetic transition as the temperature changes. For the Viana-Bray spin glass, error grows monotonically lowering the temperature.

For the RFIM, which is one of the standard literature examples of disordered system, the dynamic cavity equation reproduces the dynamical behavior with a quality comparable to the ferromagnetic case, see Figs. 1(b) and 2(b). In this case we used a value $h_i = \pm 0.3$ chosen from a bimodal distribution for the external field acting on each node, which is kept constant in time. The corresponding ferromagnetic transition temperature for this case is $T_c(h = 0.3) \approx 0.78$. The critical transition errors found in this case are of the same order of magnitude to the situation where $h_i = 0$.

The Viana-Bray model, on the other hand, shows errors one order of magnitude larger than the previous models, which worsen as temperature decreases, see Figs. 1(c) and 2(c). It is known that this model has a spin-glass 1RSB transition at $T_{SG} = 0.506$ and this implies a fundamental difference with respect to the previously considered models. The state space collapses for low temperature into a hierarchy of low-energy configurations and it is no longer well described by only one equilibrium solution.

Note that in the insets of Figs. 1 and 2 errors are low at very short times as well as in the stationary regime for both the ferromagnet and the RFIM [Figs. 1(a), 2(a) and 1(b), 2(b)]. For these cases, the error is maximum at intermediate times, during the transient regime. The SG model shows a different behavior instead [see Figs. 1(c) and 2(c)]. In this case, indeed, the error increases monotonically with time for low temperatures in opposition to what happens with the ferromagnet and RFIM. As we said before, this indicates that there is a wrong assumption regarding the structure of phase space.

In Fig. 3 we present, for all models, the temperature dependence of the maximum value of the mean error. More precisely, for each given temperature T at which simulations are run, there is a time $\bar{t}(T)$ at which the error is maximum. Figure 3 illustrates the behavior of this error as a function of the temperature. Let us note that for the ferromagnet and RFIM the error increases by decreasing the temperature from high value towards the critical transition temperature and start decreasing when the temperature is decreased below the transition temperature, see Figs. 3(a) and 3(b). Differently, for the Viana-Bray model, the performances of our approach progressively worsen near the SG phase transition and the error monotonically increases by decreasing the temperature in the low-temperature regime [see Fig. 3(c)].

VIII. CONCLUSIONS

In this work we have derived a method to close the master equation for the continuous dynamics of interacting spins. The approach relies on the factorization of the conditional distribution of the state of spin i and its neighbors and on the formalism of the theory of random point processes. By assuming a treelike graph topology, using this formalism, we are able to rederive a known equation for conditional probabilities of spin histories, which is called dynamic message-passing or dynamic cavity equation in the literature. Combining this result with the approximated master equation and using the random point process formalism, we are able to parametrize probability distributions of the spin histories and obtain a rigorous derivation of a dynamic equation for the conditional distribution of spin variables, the cavity master equation. This equation together with the master equation for the single spin dynamics completely determine the temporal evolution of the model.

We have shown that our approach reproduces the known analytical solution of two prototypical models and we have tested numerically the performances of the method for three more complex cases defined on random graphs. Numerical results show a quantitative good agreement with Monte Carlo simulations for those models, which do not have a spin-glass phase. For the Viana-Bray model, the technique fails below the glass transition.

We believe that the general nature of our method allows us to apply it on models with different transition rates and networks with various connectivity symmetries (asymmetric and partially symmetric graphs), and therefore could bring us to several further developments to investigate the dynamics of physical and biological systems. Extension to models with a glassy phase is a topic for future research.

ACKNOWLEDGMENTS

This research has been supported by the Swedish Science Council through Grant No. 621-2012-2982, by the

Academy of Finland through its Center of Excellence COIN (E.A.) and by the Chinese Academy of Sciences CAS Presidents International Fellowship Initiative (PIFI) Grant No. 2016VMA002 (E.A.), by European Union through Marie Curie ITN "NETADIS" FP7/2007-2013/Grant Agreement No. 290038 (G.D.F.), by STINT and EUSBIOS.

APPENDIX A: RELATIONS BETWEEN PROBABILITY DENSITIES AND CAVITY PROBABILITY DENSITIES

In this Appendix we discuss the relation linking the probability density $P(\sigma_i|\sigma_k)$ shown in (4) with the cavity conditional probability $p(\sigma_i|\sigma_j)$ appearing in Eqs. (5) and (44). We start showing that the probability of spin i conditioned on the history of spin j is properly defined by a density distribution of spin histories such as those introduced in Sec. III:

$$P(\sigma_i|X_j) = \sum_{X_i|\sigma_i(t)=\sigma_i} Q(X_i|X_j) = \sum_{X_i|\sigma_i(t)=\sigma_i} \frac{Q(X_i,X_j)}{Q(X_j)}. \tag{A1}$$

On the other hand, the cavity probability density $p(\sigma_i|\sigma_j)$ has been defined, in the main text, by Eq. (31), which we here report for completeness

$$p(\sigma_i|X_j) = \sum_{X_i|\sigma_i(t)=\sigma_i} \mu_{i\to(ij)}(X_i|X_j). \tag{A2}$$

To clarify the mathematical relation between the two above probability densities, let us write more explicitly Eq. (A1), with the use of (17), it becomes:

$$\begin{aligned} P(\sigma_i|X_j) &= \sum_{X_i|\sigma_i(t)=\sigma_i} \mu_{i\to(ij)}(X_i|X_j) \\ &\quad \times \frac{\mu_{j\to(ji)}(X_j|X_i)}{\sum_{X_i}\mu_{i\to(ij)}(X_i|X_j)\mu_{j\to(ji)}(X_j|X_i)} \\ &= \sum_{X_i|\sigma_i(t)=\sigma_i} \mu_{i\to(ij)}(X_i|X_j)\Delta\mu_{j\to(ji)}(X_j,X_i), \end{aligned} \tag{A3}$$

where in the second equality we introduced the quantity $\Delta\mu_{j\to(ji)}(X_j,X_i)$ as the rate appearing on the right-hand side of the first equality. The comparison of (A2) with (A3) shows that the difference between $P(\sigma_i|\sigma_k)$ and $p(\sigma_i|\sigma_j)$ resides in this rate. From its definition, such rate can be trivially rewritten as

$$\begin{aligned} &\Delta\mu_{j\to(ji)}(X_j,X_i) \\ &\quad = 1+\frac{\mu_{j\to(ji)}(X_j|X_i)-\sum_{X_i}\mu_{i\to(ij)}(X_i|X_j)\mu_{j\to(ji)}(X_j|X_i)}{\sum_{X_i}\mu_{i\to(ij)}(X_i|X_j)\mu_{j\to(ji)}(X_j|X_i)}. \end{aligned} \tag{A4}$$

In a fully asymmetric network, where if a link connecting i to j is present then the opposite link (j to i) is absent, it is immediate to show [31] that the cavity message $\mu_{i\to(ij)}(X_i|X_j) = \mu_{i\to(ij)}(X_i)$. Therefore, in this case, from Eq. (A4), it follows that $\Delta\mu_{j\to(ji)}(X_j,X_i) = 1$, and so Eq. (A3) reduces to (31) or, in a nutshell, $P(\sigma_i|X_j) = p(\sigma_i|X_j)$. In a more general setting, from Eqs. (A3) and (A4) it follows that the conditional probability $P(\sigma_i|X_j)$ can be rewritten as

$$P(\sigma_i|X_j) = p(\sigma_i|X_j) + \text{correction terms}, \tag{A5}$$

where the correction terms are equal to the second term on the right-hand side of (A4) traced over the spin history of spin i as expressed in (A3). In more general topologies and dynamics rules the magnitude of these correction terms depends on the reciprocal influence that the trajectories of spin i and j have on each other and on their degree of statistical independence.

Therefore, more generally, if the history of i and j are weakly temporally correlated we expect that $\Delta\mu_{j\to(ji)}(X_j,X_i) \simeq 1$ and so $P(\sigma_i|X_j) \approx p(\sigma_i|X_j)$. This reasoning motivates us to use (44) to close (4) with regard to the conditional probabilities and obtain a closure scheme to approximate the dynamic of the system.

APPENDIX B: CONDITIONING THE LOCAL EXACT MASTER EQUATION

In the main text, when discussing (5) we mentioned the possibility of a different derivation of that result, based solely on the conditioning of the exact ME (2). This straightforward approach, appealing as it is, is not formally correct. The problem is that conditioning a ME on a specific variable is not a procedure that can be done without further considerations.

Suppose for example that we take a master equation such as the one in (1), but conditioned to one of the spins:

$$\begin{aligned} &\frac{dP(\underline{\sigma}|\sigma_j)}{dt} \\ &\quad = -\sum_{i=1..N, i\neq j} \left[r_i(\underline{\sigma})P(\underline{\sigma}|\sigma_j) - r_i(F_i(\underline{\sigma}))P(F_i(\underline{\sigma})|\sigma_j))\right], \end{aligned} \tag{B1}$$

from which one can marginalize both sides to obtain [after the factorizations (42)] the cavity equation presented in (44). We give two arguments why this approach is formally wrong.

The first argument is that one cannot get the conditional probability $P(\underline{\sigma}|\sigma_j)$ or the marginal $P(\sigma_i|\sigma_j)$ directly from the full probability distribution $P(\underline{\sigma})$. What one can do is to use the definition of conditional probability $P(\sigma_i|\sigma_j) = \frac{P(\sigma_i,\sigma_j)}{P(\sigma_i)}$ and then compute $P(\sigma_i,\sigma_j)$ by marginalizing from $P(\underline{\sigma})$.

Doing so one will obtain a differential equation for the joint probability of two spins, analogous to (2), and which we write out as

$$\begin{aligned} \frac{dP(\sigma_i,\sigma_j)}{dt} &= -\sum_{\sigma_{\partial i}}[r_i(\sigma_i,\sigma_{\partial i})P(\sigma_i,\sigma_{\partial i},\sigma_j) \\ &\quad - r_i(-\sigma_i,\sigma_{\partial i},\sigma_j)P(-\sigma_i,\sigma_{\partial i},\sigma_j)] \\ &\quad - \sum_{\sigma_{\partial j}}[r_j(\sigma_j,\sigma_{\partial j},\sigma_i)P(\sigma_j,\sigma_{\partial j},\sigma_i) \\ &\quad - r_j(-\sigma_j,\sigma_{\partial j},\sigma_i)P(-\sigma_j,\sigma_{\partial j},\sigma_i)]. \end{aligned} \tag{B2}$$

In the above we implicitly understand that the arguments of the various terms are only counted once, i.e., if $j \in \partial i$ the

argument σ_j is already included in $\sigma_{\partial i}$ and vice versa. The time derivative of the conditional probability of interest can then be written as

$$\frac{dP(\sigma_i|\sigma_j)}{dt} = \frac{dP(\sigma_i,\sigma_j)}{dt}\frac{1}{P(\sigma_i)} - P(\sigma_i|\sigma_j)\frac{dP(\sigma_i)}{dt}\frac{1}{P(\sigma_i)^2} \tag{B3}$$

and (2) and (B2) inserted in the right-hand side. It is clear that this will give something different than (44), e.g., (B2) contains terms from flipping both spin i and spin j while (44) only has flips of spin i.

As a second argument, let us imagine that we get (B1) starting as usual from the Chapman-Kolmogorov equation where a specific spin, say σ_j, is actually fixed, and should be treated like a local field. In this sense, the conditioning notation $P(\cdots|\sigma_j)$ is only a formal statement. However, then it must be defined from outside, because it is not implicity in this equation, the dynamical equation for σ_j or its probability, and this information is naturally hidden in the model and there is not fundamental reason to choose for it the local master equation that one derives for all the other spins. Worse, this is true for every spin you fix and therefore for each spin in the model.

Alternatively one may start over and over from different Chapman-Kolmogorov equations in which different spins are fixed and get different master equations such as (B1). Then, to trace over these master equations and get, again following (42), the set of equations represented by (44) in the manuscript. However, this procedure is clearly not consistent. All the CMEs were derived from different master equations that in turn were derived starting from different Chapman-Kolmogorov equations. Therefore, there is not a physical reason that justifies the fact that they are coupled in a rigourous sense. The failure of these approaches justify the more involved analysis followed in the main body of this paper.

[1] L. Kadanoff, *Statistical Physics: Statics, Dynamics and Renormalization* (World Scientific, Singapore, 2000).

[2] P. L. Houson, *Chemical Kinetics and Reaction Dynamics* (Dover Books on Chemistry, Mineola, 2006).

[3] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, 2007).

[4] E. M. Izhikevich, *Dynamical Systems in Neurosciences* (MIT Press, Cambridge, 2007).

[5] K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).

[6] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[7] G. Parisi, A sequence of approximated solutions to the SK model for spin glasses, J. Phys. A: Math. Gen. **13**, L115 (1980).

[8] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro, Replica symmetry breaking and the nature of the spin glass phase, J. Phys. **45**, 843 (1984).

[9] M. Mézard and G. Parisi, The cavity method at zero temperature, J. Stat. Phys. **111**, 1 (2003).

[10] M. Mézard and G. Parisi, The Bethe lattice spin glass revisited, Eur. Phys. J. B **20**, 217 (2001).

[11] M. Mézard and R. Zecchina, Random k-satisfiability problem: From an analytic solution to an efficient algorithm, Phys. Rev. E **66**, 056126 (2002).

[12] R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, Coloring Random Graphs, Phys. Rev. Lett. **89**, 268701 (2002).

[13] M. Weigt and A. K. Hartmann, *Phase Transitions and Combinatorial Optimization Problems* (Wiley, New York, 2006).

[14] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).

[15] T. Rizzo, A. Lage-Castellanos, R. Mulet, and F. Ricci-Tersenghi, Replica cluster variational method, J. Stat. Phys. **139**, 375 (2010).

[16] A. Lage-Castellanos, R. Mulet, F. Ricci-Tersenghi, and T. Rizzo, Replica cluster variational method: The replica symmetric solution for the 2D random bond ising model, J. Phys. A: Math. Theor. **46**, 135001 (2013).

[17] J.-Q. Xiao and H. Zhou, Partition function loop series for a general graphical model: Free-energy corrections and message-passing equations, J. Phys. A: Math. Theor. **44**, 425001 (2011).

[18] H. Zhou and C. Wang, Region graph partition function expansion and approximate free energy landscapes: Theory and some numerical results, J. Stat. Phys. **148**, 513 (2012).

[19] C. W. Gardiner *et al.*, *Handbook of Stochastic Methods*, Vol. 3 (Springer, Berlin, 1985).

[20] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Vol. 1 (Elsevier, Amsterdam, 1992).

[21] T. Castellani and A. Cavagna, Spin-glass theory for pedestrians, J. Stat. Mech. (2005) P05012.

[22] A. Crisanti, H. Horner, and H.-J. Sommers, The spherical p-spin interaction spin-glass model: The dynamics, Z. Phys. B **92**, 257 (1993).

[23] L. F. Cugliandolo and D. S. Dean, Full dynamical solution for a spherical spin-glass model, J. Phys. A: Math. Gen. **28**, 4213 (1995).

[24] G. Semerjian, L. F. Cugliandolo, and A. Montanari, On the stochastic dynamics of disordered spin models, J. Stat. Phys. **115**, 493 (2004).

[25] B. Derrida, E. Gardner, and A. Zippelius, An exactly solvable asymmetric neural network model, Europhys. Lett. **4**, 167 (1987).

[26] J. P. L. Hatchett, B. Wemmenhove, I. Pérez Castillo, T. Nikoletopoulos, N. S. Skantzos, and A. C. C. Coolen, Parallel dynamics of disordered ising spin systems on finitely connected random graphs, J. Phys. A: Math. Gen. **37**, 6201 (2004).

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

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

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

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

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

[32] T. Barthel, C. De Bacco, and S. Franz, A matrix product algorithm for stochastic dynamics on locally tree-like graphs, arXiv:1508.03295.

[33] A. Pelizzola, Variational approximations for stationary states of ising-like models, Eur. Phys. J. B **86**, 120 (2013).

[34] E. Domínguez Vázquez, G. Del Ferraro, and F. Ricci-Tersenghi, A simple analytical description of the non-stationary dynamics in ising spin systems, J. Stat. Mech.: Theory Exp. (2017) 033303.

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

[36] R. R. Wilkinson and K. J. Sharkey, Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks, Phys. Rev. E **89**, 022808 (2014).

[37] R. J. Glauber, Time-dependent statistics of the ising model, J. Math. Phys. **4**, 294 (1963).

[38] A. C. C. Coolen, R. Kühn, and P. Sollich, *Theory of Neural Information Processing Systems* (Oxford University Press, Oxford, 2005).

[39] P. Mayer and P. Sollich, General solutions for multispin two-time correlation and response functions in the Glauber–Ising chain, J. Phys. A: Math. Gen. **37**, 9 (2004).

[40] S. N. Laughton, A. C. C. Coolen, and D. Sherrington, Order-parameter flow in the SK spin-glass: II. Inclusion of microscopic memory effects, J. Phys. A: Math. Gen. **29**, 763 (1996).

[41] A. Mozeika and A. C. C. Coolen, Dynamical replica analysis of processes on finitely connected random graphs: I. vertex covering, J. Phys. A: Math. Theor. **41**, 115003 (2008).

[42] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes Volume I: Elementary Theory and Methods* (Springer Science & Business Media, Berlin, 2002).

[43] S. Srinivasan *et al.*, *Stochastic Theory and Cascade Processes* (Elsevier, Amsterdam, 1969).

[44] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Vol. 320 (Springer-Verlag, Berlin, 1999).

[45] J. S. Yedidia, W. T. Freeman, and Y. Weiss, in *Exploring Artificial Intelligence in the New Millennium* (Morgan Kaufmann, Burlington, 2003), pp. 239–269.