Exploring the diluted ferromagnetic $p$-spin model with a Cavity Master Equation

Erik Aurell,1,∗ Eduardo Domínguez,2 David Machado,2 and Roberto Mulet2,†

1KTH Royal Institute of Technology, AlbaNova University Center, SE-106 91 Stockholm, Sweden
2Department of Theoretical Physics, University of Havana, Cuba

Dated: February 2, 2022

We introduce a new solution to Glauber multi-spin dynamics on random graphs. The solution is based on the recently introduced Cavity Master Equation (CME), a time-closure turning the in principle exact Dynamic Cavity Method into a practical method of analysis and of fast simulation. Running CME once is of comparable computational complexity as one Monte Carlo run on the same problem. We show that CME correctly models the ferromagnetic $p$-spin Glauber dynamics from high temperatures down to and below the spinoidal transition. We also show that CME allows a novel exploration of the low-temperature spin-glass phase of the model.

Nature abounds in systems of interacting units with non-trivial dynamical properties. This leads to similar questions in condensed-matter physics [1], systems biology [2], neuroscience [3] and neural networks [4], but also in many practical applications in computer science [5] and engineering [6]. Not surprisingly the study of such systems has led to the development and the use of similar techniques across different scientific communities.

A common starting point is a Markovian dynamics in continuous time of $N$ discrete interacting variables $\sigma = \{\sigma_1, \ldots, \sigma_N\}$. Such a system is described by a Master Equation defining the evolution of the probability of the states of the system $P(\sigma)$, [3, 7, 8]:

$$\frac{dP(\sigma)}{dt} = \sum_{\sigma'} \left[ r(\sigma' \rightarrow \sigma)P(\sigma') - r(\sigma \rightarrow \sigma')P(\sigma) \right] \quad (1)$$

where $r$ are transition probabilities from states.

The solution of the Master Equation (1) is in general a cumbersome task and exact results on both stationary states and transients are limited to some special cases [3, 8–10]. For fully connected [11] and dilute graphs [12] it is possible to resort to hierarchical schemes to derive dynamical equations for the probability of some macroscopic observables. A very general one of this type is the Dynamical Replica Analysis[13–15]. With this reduction of the dimensionality the problem becomes tractable, but one looses the detailed information about the microscopic state of the system and the results are not exact for fully connected [16], nor for the transients in 1D ferromagnetic systems [17].

A frequently made approximation reduces (1) to a simpler Master Equation for the probabilities of single spin variables $P(\sigma_i)$. However, this is only valid in mean-field-like models[3], or at very high temperatures. An alternative solution that successfully describes the local dynamics of these systems in a wide range of the parameter space was recently suggested in[18]. The problem is then reduced to the solution of a Master Equation for a conditional probability that is subsequently used to recover the full $P(\sigma_i)$ of equation.

In this work we generalize this last approach to systems with multi-particle interactions and explore for the first time its potential to describe processes both near to equilibrium and deep into a phase with multiple minima. We show also that it can be exploited to obtain the ground-state of models with a complex energy landscape. To keep the notation simple, and to conveniently explore the features of our solution we study the dilute ferromagnetic $p$-spin model with Glauber dynamics. However, the ideas behind our solution should be clear enough to allow the straightforward application of our method to other dynamical rules and other Hamiltonians.

The ferromagnetic $p$-spin model is defined by the Hamiltonian :

$$H = -\sum_{i_1,i_2,\ldots,i_p} J_{i_1i_2\ldots i_p} \sigma_{i_1}\sigma_{i_2}\sigma_{i_3}\ldots\sigma_{i_p} \quad (2)$$

where $J_{i_1,i_2,\ldots,i_p} = 1$ for all p-tuplets in the set of hyper-edges of a Bethe lattice and zero otherwise, and $s_i \in \{-1,1\}$ are binary variables. This model is a natural intermediary between spin glasses and structural glasses. Like the former, it is defined by binary variables and fixed in an infinite lattice. With the latter it shares a crystalline state, the absence of quenched disorder, and a finite range of the interactions.

A standard dynamics for this model uses a single spin transition rule between states $r(\sigma \rightarrow \sigma') = r_i(\sigma_1,\ldots,\sigma_i,\ldots \rightarrow \sigma_1,\ldots - \sigma_i,\ldots)$. In addition, we will consider that the transition rate of spin $i$ depends only on the state of the spin and its neighborhood; in this case, all p-tuplets it belongs to. Let us define $\partial_i$ as the set of p-tuplets that include spin $i$ and use $\sigma_\partial$ as a notation for the set of spins in $\partial_i$ excluding $\sigma_i$ and denote $\sigma_\alpha$ as the group of spins forming the p-tuple $\alpha$. The transition rate for spin $i$ will be then, for a Glauber dynamics:$r_i(\sigma,\sigma_\partial) = \frac{\alpha}{2} [1 - \sigma_i \tanh(\beta \sum_{\sigma_\alpha} J_{\alpha} \sigma_\alpha(i))]$.

Under these general settings the local dynamics of spin $i$ is described by a local Master Equation:
\[
\frac{dP(\sigma_i)}{dt} = - \sum_{\sigma_{\partial_i}} p_i(\sigma_i, \sigma_{\partial_i}) P(\sigma_i, \sigma_{\partial_i}) - r_i(-\sigma_i, \sigma_{\partial_i}) P(-\sigma_i, \sigma_{\partial_i}) \tag{3}
\]

Equation (3) looks simpler than (1), but it is not a closed set of equations. To close it we need to resort to proper approximations. In this case we first write \( P(\sigma_i, \sigma_{\partial_i}) = \prod_{a \in \partial_i} P(\sigma_{a\downarrow} | \sigma_i) P(\sigma_i) \), which is exact in equilibrium for trees and random graphs with large loops, and substitute it in (3). The goal in what follows is to find a proper approximation for the dynamics of the conditional probability \( P(\sigma_{a\downarrow} | \sigma_i) \) that substituted back in (3) makes the calculation of \( P(\sigma_i, t) \) straightforward.

We build our equations with the help of the theory of Random Point Processes [7, 19] where the dynamics of the spin variable \( \sigma_i \) is encoded in a trajectory \( \{ X_t \} \) that is parametrized by the number of jumps \( s_i \) of the corresponding spin in the studied interval \([t_0, t]\) and the times \( \{ t_1, t_2, \ldots, t_s \} \). Very generally for spins interacting through Hamiltonians defined on factor graphs the joint probability distribution \( Q \) of these trajectories can be written as: 
\( Q(X_1 \ldots X_N) = \prod_{i=1}^{N} \Phi_i(X_i | X_{\partial i}) \) where
\[
\Phi_i(X_i | X_{\partial i}) = \prod_{t_i=1}^{s_i} \phi_i(\tau_i(t_i), \sigma_{\partial i}(t_i)) 
\times \exp\left[ - \int_{t_0}^{t} r_i(\tau, \sigma_{\partial i}(\tau)) d\tau \right]. \tag{4}
\]

For locally tree-like graphs this parametrization leads to a message-passing equation (see Supplementary Material), structurally identical to the Belief Propagation equations used to approximate an equilibrium Gibbs-Boltzmann distribution[5]:
\[
\mu_{a \rightarrow i}(X_{a\downarrow} | X_i) = \prod_{j \in a \downarrow \setminus i} \sum_{X_j | X_{\partial j}} \Phi_j(X_j | X_{\partial j}) 
\times \prod_{b \in \partial_j \setminus a} \mu_{b \rightarrow j}(X_{b\downarrow} | X_j). \tag{5}
\]

Formally, equation (5) defines a set of fixed point equations for the probabilities \( \mu_{a \rightarrow i}(X_{a\downarrow} | X_i) \) of the histories in the set \( X_{a\downarrow} \) in terms of equivalent objects in the neighborhood of the spins in \( a \setminus i \), considering that the history \( X_i \) is given.

Unfortunately, the histories, defined by the variables \( X \) are cumbersome arguments to treat beyond formal statements, and actual solution of (5) is hopeless. One can however marginalize this quantity via \( p(\sigma_{a\downarrow} | X_i, t) = \sum_{X_{a\downarrow}} \mu_{a \rightarrow i}(X_{a\downarrow} | X_i), \) and then take the proper time derivative (see Supplementary Material) to obtain:
\[
\dot{p}(\sigma_{a\downarrow} | X_i) = - \sum_{j \in a \downarrow \setminus i} \sum_{\sigma_{b\downarrow}} r_j^{+} p(\sigma_{b\downarrow} | \sigma_{a\downarrow}) - \sum_{\sigma_{b\downarrow}} r_j^{-} p(\sigma_{b\downarrow} | \sigma_{a\downarrow}, F_j | \sigma_{a\downarrow} | X_i) \tag{6}
\]
where \( r_j^{+/(-)} = r_j^{+}(-\sigma_i, \sigma_{b\downarrow}) \) and \( \{ \sigma_{b\downarrow} \} \) with \( b \in \partial_j \setminus a \) is the set of instantaneous variables that characterize the nodes neighboring \( j \), except those in \( a \).

Note that although eq. (6) is exact in tree-like graphs it still contains spin histories as conditional arguments. In principle, these histories could be taken as parameters to be tracked during the solution of (6). However, it is convenient to go further assuming first that variables factorize around factor nodes: \( p(\sigma_{b\downarrow} | \sigma_{a\downarrow}, X_i) = \prod_{b \in \partial_j \setminus a} p(\sigma_{b\downarrow} | \sigma_{a\downarrow}, X_i) \approx p(\sigma_{a\downarrow} X_i) \) which is exact in tree-like graphs in equilibrium. Then, to close the system of equations it is enough to consider that locally the system has short time memory, \( p(\sigma | X_i) \sim p(\sigma | \sigma_i) \) which is equivalent to a Markov hypothesis.

With these approximations equation (6) transforms into a Master Equation but conditioned to neighboring (cavity) spins. We call this the Cavity Master Equation (CME):
\[
\dot{p}(\sigma_{a\downarrow} | \sigma_i) = - \sum_{j \in a \downarrow \setminus i} \sum_{\sigma_{b\downarrow}} r_j^{+} \prod_{b \in \partial_j \setminus a} p(\sigma_{b\downarrow} | \sigma_j) p(\sigma_{a\downarrow} | \sigma_i) 
+ \sum_{j \in a \downarrow \setminus i} \sum_{\sigma_{b\downarrow}} r_j^{-} \prod_{b \in \partial_j \setminus a} p(\sigma_{b\downarrow} | -\sigma_j) p(F_j | \sigma_{a\downarrow} | \sigma_i). \tag{7}
\]

Then, to obtain the observables of the system one just solves the set of equations (7), plugs the result in the local Master Equation (3) and integrate to obtain the local probabilities at the nodes of the network. This gives the joint probability distribution of all the variables for all times in one go, the same outcome which would require averaging over many realizations using Kinetic Monte Carlo (KMC). Although we will not pursue this issue here, let us note that a similar approach could also, as for standard Belief Propagation, be used to describe the evolution of the marginal probability of a simply connected subset of nodes in the graph.

Let us now discuss in more detail some of the properties of the p-spin ferromagnetic model with \( p = 3 \) in a random regular graph with \( k = 3 \). As discussed in [20] it shows three different phases. For \( T > T_{ms} \approx 1.63 \) it is paramagnetic. At \( T_{ms} \) emerges a ferromagnetic metastable consistent with a spinodal transition. This ferromagnetic state becomes stable at \( T_{fm} \approx 2.12 \). For lower
temperatures $T < T_K = 0.655[21]$ the system has a thermodynamic phase transition to a spin glass state. The origin of this spin glass states is the existence of a free-energy minimum with the same energy of the ferromagnetic state, but with higher entropy. This minimum can be described as follows: when a spin is down, it puts an effective negative interaction on the other $k-1 = 2$ neighbors that will act as an anti-ferromagnetic pair. Thus, given the tree-like structure of the lattice, is is possible to minimize the energy with a mix of ferromagnetic and anti-ferromagnetic plaquettes. The previous configurations are energy minimum and their mean magnetization is zero. There are many different realizations of such configurations, and this leads to fluctuations in the system dynamics, whose interactions are not fully satisfied because of existing loops. Eventually the spins freeze giving origin to spin glass states that define the dynamics of the model for $T_d < 0.757[21]$.

We know from previous results that the CME fails to predict KMC results deep into the SG phase and for intermediate time scales near second order phase transitions[18]. Therefore we will here focus first our attention in the behavior of the method around the spinodal transition. We compare CME to KMC starting at $t = 0$ from a totally ordered configuration with magnetization one. Then we observe the evolution of the average magnetization after a quench to a temperature $T$ and quantitatively compare the CME and KMC using the mean square difference of local magnetization’s 
\[ \delta m(t) = \sqrt{\frac{1}{N} \sum_i (m_{i}^{CME}(t) - m_{i}^{KMC}(t))^2} \]

Results for both quantities are reported in panels of Fig. 1. In this numerical experiment, below the spinodal transition ($T_{ms} < 1.63$), and deep in the high temperature region ($T > 1.90$) CME and KMC produce the same output. In the first case the system stays trapped in the magnetized state while in the later the system relax exponentially to a paramagnetic phase. Only above and near the spinodal transition and at intermediate times do the algorithms differ; the long-term results, however, are the same. Note that even around the spinodal transition the errors ultimately become zero supporting the idea that the CME properly reproduces the long time behavior of the system.

We also tested our CME in a planted model [22] near the dynamical phase transition of the model $T_d$. The planted model is built fixing the ground state, all $\sigma_i = 1$, and then choosing the value of the links in the graph following the rule: 
\[ P(J_a = 1) = \frac{1 + \tanh(\beta)}{2} \]
\[ P(J_a = -1) = \frac{1 - \tanh(\beta)}{2} \]

The results of the evolution of the magnetization in this graph are shown in Fig. 2. In this case our dynamics evolves toward the correct paramagnetic state for $T > .78$ a temperature that is close to but slightly above the expected $T_d = 0.757$. However, one most notice, that these results are strongly affected by sample to sample fluctuations, as the inset in Figure

![Fig. 1: Top: Dynamics of the average magnetization using CME (lines and symbols) and KMC (symbols). In the inset: Equilibration time near the spinodal transition temperature. Bottom: Dynamical behavior of the local error between CME and KMC. System size: $N = 99999$ spins. In the inset: Temperature dependence of the maximum local error. At time $t = 0$ the system is fully magnetized, and then begins to interact with a heat bath at a given temperature.](image1)

![Fig. 2: Dynamical behavior of the planted model for $N = 10002$ near $T_d$. In the plot are represented curves for a single planted graph. In the inset we show the fraction of the samples in which the magnetization decays to zero as a function of $T$.](image2)
Motivated by the long time behavior of the CME we explored the behavior of the observables of the system in an adiabatic protocol. Although the set of differential equations (7) was written with the aim to study dynamical processes one can equally well let the integration evolve in time and define the equilibrium states when the cavity probabilities reach a fixed point of the dynamics. This idea is used here to reproduce adiabatic heating and cooling experiments in the model. In Fig. 3 we show the result of these experiments and compare them with KMC. The heating process reflects quite well the existence of $T_{ms}$. Starting from an homogeneous ferromagnetic configuration at low $T$, the temperature is increased in small steps. Both CME and KMC remain in the ferromagnetic state until it disappears beyond the spinodal temperature and then jumps to the paramagnetic state.

![Graph showing the comparison between the quasi-equilibrium evolution of the CME (continuous lines) and KMC (dots). The CME was applied to a $N=1002$ system, and KMC to a $N=10002$ system. The temperature step is in both cases $\Delta T = 0.02$. Both approaches reproduce the existence of $T_{ms} \approx 1.62$. Below $T_d$ both dynamics differentiate.](image)

**FIG. 3:** Comparison between the quasi-equilibrium evolution of the CME (continuous lines) and KMC (dots). The CME was applied to a $N=1002$ system, and KMC to a $N=10002$ system. The temperature step is in both cases $\Delta T = 0.02$. Both approaches reproduce the existence of $T_{ms} \approx 1.62$. Below $T_d$ both dynamics differentiate.

On the other hand, if we start cooling the system from a paramagnetic state at high temperatures (Figure (3)), the CME and KMC will coincide only down to the temperature $T_d \approx 0.757$ for which the spin glass phase dynamically traps the stochastic simulation [20]. Below this temperature the behavior of the CME depends on the criterion adopted to define the convergence of the algorithm. Our criterion is to stop whenever for each plaquette $| p(\sigma_k, \sigma_j \mid \sigma_i, t+dt) - p(\sigma_k, \sigma_j \mid \sigma_i, t+dt) | \leq \alpha$, where $\alpha$ is a parameter controlling the time that the system spends at a given temperature. The smaller the $\alpha$ the larger the time we gave the system to equilibrate. Remarkably, if $\alpha$ is small enough the probability measure of the CME splits evenly among ordered states with antiferromagnetic and ferromagnetic plaquettes of the same energy. Moreover, not only the total but also the local magnetizations obtained in this temperature region are zero. This suggests that the quasi-dynamics given by the CME deep into the glassy phase may be useful to explore the low energy structure of the system at low temperatures.

This low temperature structure is indeed one of the most interesting features of the model, since despite the clear presence of a crystalline state it can not be reached by any known local dynamical rule. In that scenario, we decided to fix a fraction of the spins using strong local fields pointing in the same direction. We then inquired how large should be the fraction of spins aligned to obtain the correct crystalline state after a quench to a very low temperature. The results are shown in Figure 4 where we compare the results of the CME and KMC. It is evident that while the KMC dynamics converges to the proper equilibrium state when almost 30% of the spin are oriented in the ferromagnetic state, the pseudo-dynamics of the CME at low temperature recognize the ground state of the model with only a 15% of the spins correctly oriented. The smaller number of spins needed to drive the dynamics of the model into the ground state of the system for the CME suggests that it can be used as proper proxy for similar problems in Combinatorial Optimization.

![Graph showing the final magnetization of a system as a function of the number of fixed spins at $T = 0.12$ following the dynamics of the CME $N = 2001$ and KMC $N = 3000$.](image)

**FIG. 4:** Final magnetization of a system as a function of the number of fixed spins at $T = 0.12$ following the dynamics of the CME $N = 2001$ and KMC $N = 3000$.

In summary, we have presented a formal approach to study the continuous dynamics of a discrete system with multiple-particle interactions. We illustrate the power of this approach studying the dynamical behavior of the $p$-spin dilute ferromagnet. We have shown that the Cavity Master Equation reproduces the long time behavior of Kinetic Monte Carlo simulations in a wide temperature range. CVM also reproduces exactly the spinodal Temperature $T_{ms}$ and is a very good proxy for the exact value of the dynamical temperature $T_d$. Below $T_d$, where the spin glass states dominate the dynamics the CME fails to predict the KMC results. However, this allows a deeper exploration of the structure of the glassy phase and in particular to find the ground state of the system by fixing a small number of spins.
Acknowledgments: This project has received funding from the European Unions Horizon 2020 research and innovation programme MSCA-RISE-2016 under grant agreement No. 734439 INFERNET. It has been partially supported by the STINT (Sweden) project "Enhancing cooperation opportunities with Havana University through the Erasmus Programme actions", by an Erasmus+ International Credit Mobility to KTH (EU), and by the COIN Centre of Excellence, of the Academy of Finland (grant no. 251170).

* eaurell@kth.se
† mulet@fisica.uh.cu

[1] Akira Onuki. Phase Transition Dynamics. Cambridge University Press, 2002.
[2] Martin Nowak. Evolutionary Dynamics: Exploring the Equations of Life. The Belknap Press, 2006.
[3] Daniel Amit. Modelling brain function: the world of attractor neural networks. Cambridge University Press, 1989.
[4] Kevin Gurney. An introduction to Neural Networks. University College of London, 1999.
[5] Marc Mezard and Andrea Montanari. Information, physics, and computation. Oxford University Press, 2009.
[6] Steven J. Schiff. Neural Control Engineering. The MIT Press, 2011.
[7] NG van Kampen. Stochastic Processes in Physics and Chemistry, volume 1. Elsevier, Amsterdam, 1992.
[8] Anthony CC Coolen, Reimer Kühn, and Peter Sollich. Theory of neural information processing systems. Oxford University Press, 2005.
[9] Roy J Glauber. Time-dependent statistics of the ising model. Journal of mathematical physics, 4:294, 1963.
[10] Peter Mayer and Peter Sollich. General solutions for multispin two-time correlation and response functions in the glauber–ising chain. Journal of Physics A: Mathematical and General, 37(1):9, 2004.
[11] SN Laughton, ACC Coolen, and D Sherrington. Order-parameter flow in the sk spin-glass: Ii. inclusion of microscopic memory effects. Journal of Physics A: Mathematical and General, 29(4):763, 1996.
[12] Guilhem Semerjian and Martin Weigt. Approximation schemes for the dynamics of diluted spin models: The ising ferromagnet on a bethe lattice. Journal of Physics A: Mathematical and General, 37(21):5525, 2004.
[13] J.P.L: Hatchett, I. Pérez-Castillo, A.C.C. Coolen, and N. S. Slantzos. Dynamical replica analysis of disordered ising spin systems on finitely connected random graph. Physical Review Letters, 95:117204, 2005.
[14] Alexander Mozeika and ACC Coolen. Dynamical replica analysis of processes on finitely connected random graphs: I. vertex covering. Journal of Physics A: Mathematical and Theoretical, 41(11):115003, 2008.
[15] Alexander Mozeika and ACC Coolen. Dynamical replica analysis of processes on finitely connected random graphs: II. dynamics in the griffiths phase of the diluted ising ferromagnet. Journal of Physics A: Mathematical and Theoretical, 42(19):195006, 2009.
[16] Hidetoshi Nishimori and Michiko Yamana. Dynamical probability distribution function of the sk model at high temperatures. Journal of the Physical Society of Japan, 65(1):3–6, 1996.
[17] Gino Del Ferraro and Erik Aurell. Perturbative large deviation analysis of non-equilibrium dynamics. Journal of the Physical Society of Japan, 83(8):084001, 2014.
[18] Erik Aurell, Gino Del Ferraro, Eduardo Domínguez, and Roberto Mulet. A cavity master equation for the continuous time dynamics of discrete spins model. Physical Review E, 95:052119, 2017.
[19] C. Kipnis and C. Landim. Scaling Limits of Interacting Particle Systems, volume 320. Springer Verlag, Berlin, 1999.
[20] Silvio Franz, Marc Mézard, and Federico Ricci-Tersenghi. A ferromagnet with a glass transition. Europhysics Letters, page 465, 2001.
[21] Andrea Montanari and Federico Ricci-Tersenghi. Cooling-schedule dependence of the dynamics of mean-field glasses. Physical Review B, 70:134406, 2004.
[22] Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: Thresholds and algorithms. Advances in Physics, page 453, 2016.

SUPPLEMENTARY MATERIAL

Message Passing Equation

A possible parametric expression for the joint probability of the spin trajectories is:

\[ Q(X_1 \ldots X_N) = \prod_{i=1}^{N} \Phi_i(X_i | X_{\partial i}) \] (8)

where

\[ \Phi_i(X_i | X_{\partial i}) = \prod_{l=1}^{s_i} r_i(\sigma_l(t_l), \sigma_{\partial l}(t_l)) \times \exp[- \int_{t_0}^{t} r_i(\sigma(\tau), \sigma_{\partial}(\tau))d\tau] \] (9)

For a tree-like structure the expression (8) can be written as an expansion that takes a factor node as a starting point. In a mathematical language:
\[ Q(X_1, \ldots, X_N) = \Phi_i(X_i \mid X_{\partial i}) \prod_{a \in \partial i} M_{a \mid i}(X_{a \setminus i}, \{X\}_{a,i} \mid X_i) \]  

(10)

where \( M_{a \mid i} \) is:

\[ M_{a \mid i}(X_{a \setminus i}, \{X\}_{a,i} \mid X_i) = \prod_{j \in a \setminus i} \Phi_j(X_j \mid X_{\partial j}) \prod_{b \in \partial j \setminus a} \left[ \prod_{m \in b \setminus j} \Phi_m(X_m \mid X_{\partial m}) \ldots \right] \]  

(11)

Now, marginalizing conveniently:

\[ Q(X_i, X_{\partial i}) = \Phi_i(X_i \mid X_{\partial i}) \prod_{a \in \partial i} \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) \]  

(12)

where:

\[ \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) = \sum_{\{X\}_{a,i}} M_{a \mid i}(X_{a \setminus i}, \{X\}_{a,i} \mid X_i) \]  

(13)

can be interpreted as probabilities of the histories in \( X_{a \setminus i} \) given a fixed \( X_i \), or in other words, in a cavity where \( X_i \) is the history of spin \( i \).

On the other hand, the nodes connected to a factor node can also be taken as starting point of a new expansion of (10). Again, after marginalization:

\[ Q(X_i, X_{\partial i}) = \sum_{\{X\}_{a,i}} \prod_{j' \in a} \Phi_{j'}(X_{j'} \mid X_{\partial j'}) \prod_{b \in \partial j' \setminus a} \mu_{b \rightarrow (j')}(X_{b \setminus j'} \mid X_{j'}) \]  

(14)

and putting (12) and (14) together gives an equation for \( \mu \):

\[ \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) = \sum_{\{X\}_{a,i}} \prod_{j' \in a \setminus i} \Phi_{j'}(X_{j'} \mid X_{\partial j'}) \prod_{b \in \partial j' \setminus a} \mu_{b \rightarrow (j')}(X_{b \setminus j'} \mid X_{j'}) \]  

(15)

In order to seek some simplicity in (15), one can interchange the product and the sum:

\[ \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) = \prod_{j \in a \setminus i} \sum_{\{X\}_{a,i}} \Phi_j(X_j \mid X_{\partial j}) \prod_{b \in \partial j \setminus a} \mu_{b \rightarrow (j)}(X_{b \setminus j} \mid X_j) \]  

(16)

Note that equation (16) gives the probabilities of the histories in any set \( X_{a \setminus i} \) when \( X_i \) is fixed in terms of similar objects that depends of all the neighbors of the nodes in \( a \setminus i \). This derivation is exact only in tree-like graphs (with long loops).

### Time derivative

The marginalization of \( \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) \) is:

\[ p(a_{a \setminus i} \mid X_i, t) = \sum_{\{X\}_{a \setminus i}} \mu_{a \rightarrow (i)}(X_{a \setminus i} \mid X_i) \]  

(17)

and the derivative:
\[
\frac{dp(\sigma_{a \setminus i}|X_i, t)}{dt} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ p(\sigma_{a \setminus i}|X_i, t + \Delta t) - p(\sigma_{a \setminus i}|X_i, t) \right]
\] (18)

Then, to ease the notation from it will be written \( \Delta t \) in place of \( t + \Delta t \) and 0 in place of \( t \). So the previous formula would be:

\[
\frac{dp(\sigma_{a \setminus i}|X_i, t)}{dt} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ p(\sigma_{a \setminus i}|X_i, \Delta t) - p(\sigma_{a \setminus i}|X_i, 0) \right]
\] (19)

The first term inside the limit is the important one:

\[
p(\sigma_{a \setminus i}|X_i, \Delta t) = \sum_{X_{a\setminus i}((\Delta t) \atop \sigma_{a\setminus i}((\Delta t) = \sigma_{a\setminus i})} \mu_{a \to (i)}(X_{a\setminus i}((\Delta t)|X_i((\Delta t)))
\] (20)

With a precision of \( o(\Delta t) \), the trace in (20) can be divided in two terms \( A \) and \( B \) such that:

\[
p(\sigma_{a \setminus i}|X_i, \Delta t) = A + B + o(\Delta t)
\] (21)

where

\[
A = \sum_{X_{a\setminus i}((\Delta t) \atop \sigma_{a\setminus i}((0) = \sigma_{a\setminus i})} \mu_{a \to (i)}(X_{a\setminus i}((\Delta t)|X_i((\Delta t)))
\] (22)

and

\[
B = \sum_{j \in a \setminus i} \sum_{X_{a\setminus i}((\Delta t) \atop \sigma_{a\setminus i}((0) = F_j[\sigma_{a\setminus i}])} \mu_{a \to (i)}(X_{a\setminus i}((\Delta t)|X_i((\Delta t)))
\] (23)

Expanding \( A \) and \( B \) also to order \( \Delta t \) and putting it together into the limit:

\[
\frac{dp(\sigma_{a \setminus i}|X_i, 0)}{dt} = -\sum_{j \in a \setminus i} \sum_{X_{a\setminus i}((t) = \sigma_{a\setminus i})} \lambda_j(X_j, X_{a\setminus i}, X_i) \mu_{a \to (i)}(X_{a\setminus i}|X_i)
\]

\[
+ \sum_{j \in a \setminus i} \sum_{X_{a\setminus i}((t) = F_j[\sigma_{a\setminus i}])} \lambda_j(X_j, X_{a\setminus i}, X_i) \mu_{a \to (i)}(X_{a\setminus i}|X_i)
\] (24)

Now it’s needed an expression for \( \lambda_j(X_j, X_{a\setminus i}, X_i) \mu_{a \to (i)}(X_{a\setminus i}|X_i) \). In order to find it, one can differentiate (??) but with some care, because an assumption is required: the fact that the only nodes that can change its state in the \([t, t + \Delta t]\) time interval are those in \( a \) implies that the traces for the other nodes should be done to time \( t \). Inserting the result into (24) and explicitly doing some required marginalizations gives:

\[
\dot{p}(\sigma_{a \setminus i}|X_i, 0) = -\sum_{j \in a \setminus i} \sum_{\{\sigma_{b \setminus j}\} \atop b \in \partial_j \setminus \sigma_{a \setminus i} X_i} r_j^+ p(\{\sigma_{b \setminus j}\} | X_i)
\]

\[
+ \sum_{j \in a \setminus i} \sum_{\{\sigma_{b \setminus j}\} \atop b \in \partial_j \setminus \sigma_{a \setminus i} F_j[\sigma_{a \setminus i}] X_i)
\] (25)

where \( r_j^+ = r_j(\sigma_{j}, \{\sigma_{b \setminus j}\} b \in \partial_j \setminus \sigma_{a \setminus i}, \sigma_{a \setminus i}), r_j^- = r_j(-\sigma_{j}, \{\sigma_{b \setminus j}\} b \in \partial_j \setminus \sigma_{a \setminus i}, \sigma_{a \setminus i}). \)