Rapid Communications

Matrix product algorithm for stochastic dynamics on networks applied to nonequilibrium Glauber dynamics

Thomas Barthel,[1,2] Caterina De Bacco,[2,3] and Silvio Franz[2]
[1]*Department of Physics, Duke University, Durham, North Carolina 27708, USA*
[2]*Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, CNRS UMR 8626, 91405 Orsay Cedex, France*
[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

(Received 13 August 2015; revised manuscript received 16 October 2017; published 29 January 2018)

We introduce and apply an efficient method for the precise simulation of stochastic dynamical processes on locally treelike graphs. Networks with cycles are treated in the framework of the cavity method. Such models correspond, for example, to spin-glass systems, Boolean networks, neural networks, or other technological, biological, and social networks. Building upon ideas from quantum many-body theory, our approach is based on a matrix product approximation of the so-called edge messages—conditional probabilities of vertex variable trajectories. Computation costs and accuracy can be tuned by controlling the matrix dimensions of the matrix product edge messages (MPEM) in truncations. In contrast to Monte Carlo simulations, the algorithm has a better error scaling and works for both single instances as well as the thermodynamic limit. We employ it to examine prototypical nonequilibrium Glauber dynamics in the kinetic Ising model. Because of the absence of cancellation effects, observables with small expectation values can be evaluated accurately, allowing for the study of decay processes and temporal correlations.

DOI: 10.1103/PhysRevE.97.010104

I. INTRODUCTION

In recent years, we have seen increased efforts by statistical physicists to tackle stochastic dynamical processes in networks in order to study various phenomena [1,2] such as ordering processes, the spreading of epidemics and opinions, synchronization, collective behavior in social networks, stability under perturbations, and avalanche dynamics.

A drastic simplification can be achieved when short cycles in the network, defined by interaction terms, are very rare. This is the case for locally treelike graphs such as random regular graphs, Erdős-Rény graphs, and Gilbert graphs. For such random graphs with N vertices, almost all cycles have length $\gtrsim \log N$ such that their effect is negligible in the thermodynamic limit [3]. For static problems, this has been exploited in the so-called cavity method [4], where conditional nearest-neighbor probabilities are computed iteratively within the Bethe-Peierls approximation. The method was very successfully applied to study, for example, equilibrium properties of spin glasses [4], computationally hard satisfiability problems [5,6], and random matrix ensembles [7].

This big success has motivated the generalization of the cavity method to dynamical problems [8,9], which is known as the dynamic cavity method or dynamic belief propagation. Unfortunately, the number of possible trajectories and, hence, the computational complexity increase exponentially in time. Applications have hence been restricted to either very short times [8,10], oriented graphs [8], or unidirectional dynamics with local absorbing states [9,11–13]. In the latter case, one can exploit that vertex trajectories can be parametrized by a few switching times. The problem is to find good approximations to the exact solution of the dynamic cavity equations with polynomial computations costs. Simple approaches are to neglect temporal correlations completely as in the one-step method [8,14–16] or to retain only some $\Delta t = 1$ correlations as in the one-step Markov ansatz [17]. While this can be expected to work well for stationary states at high temperatures, such approximations are usually quite severe for short to intermediate times or low temperatures. Also, for dense networks, where the cavity method is not applicable, approximation schemes like the cluster variational method [18–20] or perturbative schemes [15,21,22] have been developed.

In this paper, we present an efficient algorithm for precise solutions of the parallel dynamic cavity equations for generic (locally treelike) graphs and generic bidirectional dynamics. The main feature is the reduction of the computational complexity from exponential to polynomial in the duration of the dynamical process. The central objects in the dynamic cavity method are conditional probabilities for vertex trajectories of nearest neighbors—the so-called edge messages. As temporal correlations are decaying in time and/or time difference $|t - t'|$, we exploit that the edge messages can be approximated by matrix products; i.e., there is one matrix for every edge, edge state, and time step, encoding the temporal correlations in the corresponding part of the evolution. It turns out that the dimensions of these matrices do not have to be increased exponentially in time. One can obtain quasiexact results with relatively small matrix dimensions. Computation costs and accuracy can be tuned by controlling the dimensions through controlled truncations. The idea of exploiting the decay of temporal correlations to approximate edge messages in matrix product form is in analogy with the use of matrix product states [23–27] for the simulation of strongly correlated, mostly one-dimensional, quantum many-body systems. These have been used very successfully in algorithms like the density-matrix renormalization group [28,29] to study, for example, quantum ground-state properties, often with machine precision [30]. Besides lifting the restrictions of the aforementioned approaches, the matrix product edge-message (MPEM) algorithm can also outperform Monte Carlo simulations (MC) of the dynamics in important respects. In particular, besides allowing for the simulation of single instances, alternatively, one can work directly in the thermodynamic limit. Perhaps more importantly, it has a favorable error scaling. While statistical errors in MC decay very slowly with the number of samples N_s as $1/\sqrt{N_s}$, MPEM yields also observables with absolutely small expectation values with very good accuracy which is essential for the study of decay processes and temporal correlations. As a first application, we solve the prototypical example for nonequilibrium dynamics on networks—Glauber dynamics of the kinetic Ising model [31]—and study the equilibration of the magnetization as well as temporal correlations.

II. THE DYNAMIC CAVITY METHOD

Let σ_i^t denote the state of vertex i at time step t, and $\boldsymbol{\sigma}^t := (\sigma_1^t, \sigma_2^t, \dots)$ the state of the full system at time t. Given the state probabilities $P(\boldsymbol{\sigma}^t)$ for time t, we evolve to the next time step, $P(\boldsymbol{\sigma}^{t+1}) = \sum_{\boldsymbol{\sigma}^t} W(\boldsymbol{\sigma}^{t+1}|\boldsymbol{\sigma}^t)P(\boldsymbol{\sigma}^t)$, by applying the stochastic transition matrix W. As vertex i interacts only with its nearest neighbors $j \in \partial i$, the probability for σ_i^{t+1} only depends on the states σ_j^t of these vertices at the previous time step such that the global transition matrix W is a product of local transition matrices w_i,

$$W(\boldsymbol{\sigma}^{t+1}|\boldsymbol{\sigma}^t) = \prod_i w_i\big(\sigma_i^{t+1}\big|\boldsymbol{\sigma}_{\partial i}^t\big). \tag{1}$$

Here $\sum_{\sigma_i} w_i(\sigma_i|\boldsymbol{\sigma}'_{\partial i}) = 1$, and $\boldsymbol{\sigma}_{\partial i}^t$ is the state of the nearest neighbors of vertex i at time t. In the cavity method [4,8,9], one neglects cycles of the (locally treelike) graph according to the Bethe-Peierls approximation to reduce this computationally complex evolution to the dynamic cavity equation [8,9]

$$\mu_{i\to j}\big(\bar{\sigma}_i^{t+1}\big|\bar{\sigma}_j^t\big) = \sum_{\{\bar{\sigma}_k^t\}_{k\in\partial i\setminus\{j\}}} p_i\big(\sigma_i^0\big)\left[\prod_{s=0}^{t} w_i\big(\sigma_i^{s+1}\big|\boldsymbol{\sigma}_{\partial i}^s\big)\right] \times \left[\prod_{k\in\partial i\setminus\{j\}} \mu_{k\to i}\big(\bar{\sigma}_k^t\big|\bar{\sigma}_i^{t-1}\big)\right] \tag{2}$$

which only involves the so-called edge messages μ for the edges of a single vertex i. For simplicity, we have assumed that vertices are uncorrelated in the initial state such that $P(\boldsymbol{\sigma}^0) = \prod_i p_i(\sigma_i^0)$. The edge messages $\mu_{i\to j}(\bar{\sigma}_i^t|\bar{\sigma}_j^{t-1})$ in the dynamic cavity equation (2) are conditional probabilities for the trajectories $\bar{\sigma}_i^t := (\sigma_i^0, \sigma_i^1, \dots, \sigma_i^t)$ and $\bar{\sigma}_j^{t-1}$ on edge (i,j). Specifically, if we consider a tree graph and remove all descendants of vertex

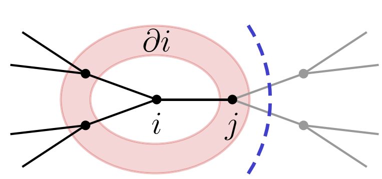

FIG. 1. Part of a locally treelike interaction graph with vertex degrees $z = 3$.

j as indicated in Fig. 1 by the dashed line, $\mu_{i\to j}(\bar{\sigma}_i^t|\bar{\sigma}_j^{t-1})$ denotes the conditional probability of a trajectory $\bar{\sigma}_i^t$ on vertex i, given the trajectory $\bar{\sigma}_j^{t-1}$ on vertex j. From messages, one can obtain marginal probabilities of site trajectories to evaluate observables of interest. Equation (2) constructs $\mu_{i\to j}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^t)$ out of the edge messages $\mu_{k\to i}(\bar{\sigma}_k^t|\bar{\sigma}_i^{t-1})$ of the previous time step. This is exact for tree graphs and covers locally treelike graphs in the Bethe-Peierls approximation. Although we have gained a lot in the sense that the computational complexity is now linear in the system size, it is still exponential in time t, if we were to encode the edge messages without any approximation.

III. CANONICAL FORM OF AN MPEM

To circumvent this exponential increase of computation costs, we can exploit the decay of temporal correlations and approximate the exact edge message by a matrix product

$$\mu_{i\to j}\big(\bar{\sigma}_i^t\big|\bar{\sigma}_j^{t-1}\big) = A_{i\to j}^{(0)}\big(\sigma_j^0\big)\left[\prod_{s=1}^{t-1} A_{i\to j}^{(s)}\big(\sigma_i^{s-1}\big|\sigma_j^s\big)\right] \times A_{i\to j}^{(t)}\big(\sigma_i^{t-1}\big)A_{i\to j}^{(t+1)}\big(\sigma_i^t\big). \tag{3}$$

The particular choice of assigning vertex variables $\{\sigma_i^s\}$ and $\{\sigma_j^s\}$ to the $M_s \times M_{s+1}$ matrices $A_{i\to j}^{(s)}(\sigma_i^{s-1}|\sigma_j^s)$ occurring in the matrix product (3) is advantageous for the implementation of the recursion relation (2) for MPEMs, as will become clear in the following. In order for the matrix product to yield a scalar, we set $M_0 = M_{t+2} = 1$.

IV. MPEM EVOLUTION

The time evolution starts at $t = 0$ with $\mu_{i\to j}(\sigma_i^0) = p_i(\sigma_i^0)$. Using the dynamic cavity equation (2), we iteratively build matrix product approximations for edge messages for time $t+1$ from those for time t. It is simple to insert the matrix product ansatz (3) for the edge messages in the dynamic cavity equation but not trivial to bring the resulting edge message again into the canonical MPEM form as required for the subsequent evolution step. The specific assignment of the vertex variables to matrices in Eq. (3) has been chosen such that all contractions (products and sums over vertex variables) occurring in the cavity equation are time-local in the sense that, given MPEMs $\mu_{k\to i}(\bar{\sigma}_k^t|\bar{\sigma}_i^{t-1})$ in canonical form for all neighbors $k \in \partial i \setminus \{j\}$, the resulting $\mu_{i\to j}(\bar{\sigma}_k^{t+1}|\bar{\sigma}_i^t)$ can be written in (noncanonical) matrix product form as

$$\mu_{i\to j}\big(\bar{\sigma}_i^{t+1}\big|\bar{\sigma}_j^t\big) = C_{i\to j}^{(0)}\big(\sigma_i^0\big)\left[\prod_{s=1}^{t+1} C_{i\to j}^{(s)}\big(\sigma_i^s\big|\sigma_j^{s-1}\big)\right]. \tag{4}$$

As depicted in Fig. 2(b), the tensors $C_{i\to j}^{(s)}$ for $1 \leqslant s \leqslant t$ are obtained by contracting the local transition matrix $w_i(\sigma_i^s|\boldsymbol{\sigma}_{\partial i}^{s-1})$ with tensors $A_{k\to i}^{(s)}$ from the time-t MPEMs. This contraction entails a sum over the $z-1$ common indices σ_k^{s-1}, where $z = |\partial i|$ is the vertex degree. Assuming for the simplicity of notation that the matrix dimensions M_s for all time-t MPEMs

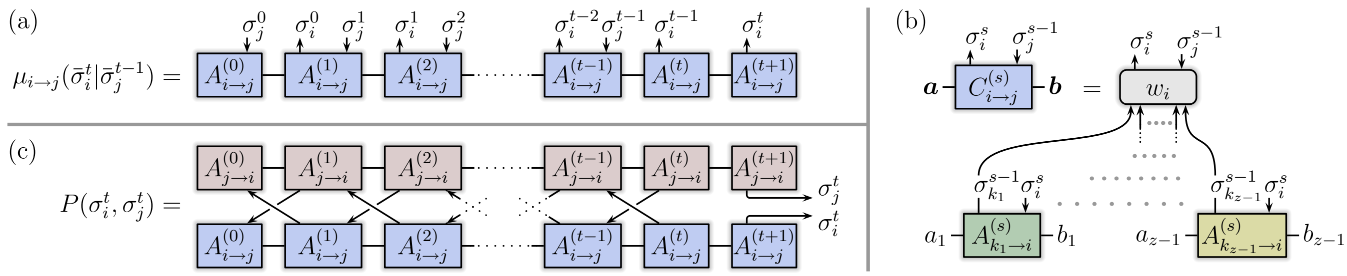

FIG. 2. (a) Graphical representation of a matrix product edge message in canonical form (3). Connecting lines indicate summations over indices. (b) For each time step (2), tensors of the evolved matrix product $\mu_{i\to j}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^t)$ in Eq. (4) are built by contracting the local transition matrix w_i with MPEM tensors of messages $\mu_{k\to i}$, incident to vertex i, where $k\in\partial i\setminus\{j\}=\{k_1,\dots,k_{z-1}\}$, and $\boldsymbol{a}:=(a_1,\dots,a_{z-1})$. (c) Evaluation of probabilities $P(\sigma_i^t,\sigma_j^t)$ as in Eq. (6).

are identical, the resulting matrices

$$C^{(s)}_{i\to j}\left(\sigma_i^s\middle|\sigma_j^{s-1}\right)=\sum_{\boldsymbol{\sigma}^{s-1}_{\partial i\setminus\{j\}}} w_i\left(\sigma_i^s\middle|\boldsymbol{\sigma}^{s-1}_{\partial i}\right)\times\left[\bigotimes_{k\in\partial i\setminus\{j\}} A^{(s)}_{k\to i}\left(\sigma_k^{s-1}\middle|\sigma_i^s\right)\right] \tag{5}$$

have left and right indices of dimensions $\bar{M}_s=(M_s)^{z-1}$ and $\bar{M}_{s+1}=(M_{s+1})^{z-1}$, respectively. The contraction for tensor $C^{(t+1)}$ is very similar and $C^{(0)}_{i\to j}(\sigma_i^0)=p_i(\sigma_i^0)[\bigotimes_{k\in\partial i\setminus\{j\}}A^{(0)}_{k\to i}(\sigma_i^0)]$.

V. MPEM TRUNCATION

In preparation for the next time step, we need to bring the evolved edge message (4) back into canonical form (3). Furthermore, we need to introduce a controlled approximation that reduces the matrix dimensions because they would otherwise grow exponentially in time. Both a reordering of the vertex variables σ_i^s and σ_j^s in the matrix product (4) and a controlled truncation of matrix dimensions can be achieved by sweeping through the matrix product and doing certain singular value decompositions (SVD) [32] of the tensors $C^{(s)}$. The generic idea behind the truncation of a matrix product $\gamma(\boldsymbol{n}):=A_0^{n_0}A_1^{n_1}\dots A_t^{n_t}$ with matrix dimensions $\{M_s\}$ is to part the variables $\boldsymbol{n}$ into two groups $\boldsymbol{n}_L:=(n_0,\dots,n_{r-1})$ and $\boldsymbol{n}_R:=(n_r,\dots,n_t)$ such that $\Gamma_{\boldsymbol{n}_L,\boldsymbol{n}_R}:=\gamma(\boldsymbol{n})$ can be interpreted as a matrix. Its singular value decomposition has the form $\Gamma_{\boldsymbol{n}_L,\boldsymbol{n}_R}=\sum_{k=1}^{M_r}Y_{\boldsymbol{n}_L,k}\lambda_k Z_{k,\boldsymbol{n}_R}$ with isometric matrices Y and Z. Retaining only the $M_r'\leqslant M_r$ largest singular values λ_k, we obtain a controlled approximation $\gamma_{\text{trunc}}(\boldsymbol{n})$ of the original matrix product $\gamma(\boldsymbol{n})$ with 2-norm distance $\|\gamma-\gamma_{\text{trunc}}\|^2=\sum_{k>M_r'}\lambda_k^2$ and decreased matrix dimension M_r' at the (temporal) bond $(r-1,r)$.

Following this principle, the truncation of all matrix dimensions of the time-evolved MPEM (4) can be done by sequential SVDs of tensors in the matrix product. In a first sweep, starting with a decomposition of the rightmost tensor $C^{(t+1)}_{i\to j}(\sigma_i^{t+1}|\sigma_j^t)\overset{\text{SVD}}{=:}U^{(t+1)}\Lambda^{(t+1)}\tilde{C}^{(t+1)}_{i\to j}(\sigma_i^{t+1}|\sigma_j^t)$ and progressing iteratively to the left with

$$C^{(s)}_{i\to j}\left(\sigma_i^s\middle|\sigma_j^{s-1}\right)U^{(s+1)}\Lambda^{(s+1)}\overset{\text{SVD}}{=:}U^{(s)}\Lambda^{(s)}\tilde{C}^{(s)}_{i\to j}\left(\sigma_i^s\middle|\sigma_j^{s-1}\right),$$

the new tensors $\tilde{C}$ are isometries. In a second sweep from left to right, matrix dimensions can be truncated from $\bar{M}_s$ to something smaller in an SVD. In two further sweeps, the indices $\{\sigma_i^s\}$ and $\{\sigma_j^s\}$ of the vertex variables can be rearranged to get back to the canonical form (3). After executing these steps for all edge messages, the next evolution step from $t+1$ to $t+2$ can follow. More details of this procedure are described in the appendices.

VI. EVALUATION OF OBSERVABLES

The joint probability of trajectories $\bar{\sigma}_i^t$ and $\bar{\sigma}_j^{t-1}$ for the vertices of an edge (i,j) is given by the product of the two corresponding edge messages. After marginalization, one obtains, for example, the probability for the edge state (σ_i^t,σ_j^t) at time t as

$$P\left(\sigma_i^t,\sigma_j^t\right)=\sum_{\bar{\sigma}_i^{t-1},\bar{\sigma}_j^{t-1}}\mu_{i\to j}\left(\bar{\sigma}_i^t\middle|\bar{\sigma}_j^{t-1}\right)\mu_{j\to i}\left(\bar{\sigma}_j^t\middle|\bar{\sigma}_i^{t-1}\right). \tag{6}$$

In the MPEM approach, this can be evaluated efficiently, as indicated in Fig. 2(c), by executing the contractions sequentially from left ($s=0$) to right ($s=t-1$). Similarly, one can for example also compute temporal correlators $\langle\sigma_i^t\sigma_i^s\rangle$ from probabilities $P(\sigma_i^t,\sigma_i^s)$.

VII. NONEQUILIBRIUM GLAUBER-ISING DYNAMICS

We have used the described MPEM algorithm to study Glauber dynamics of the kinetic Ising model, introduced in Glauber's seminal paper Ref. [31]. Figure 3 displays the results in comparison to MC simulations and to the one-step Markov approximation [17]. Specifically, we have Ising spins interacting ferromagnetically on $z=3$ random regular graphs, with local transition matrices $w_i(\sigma_i^{t+1}|\boldsymbol{\sigma}_{\partial i}^t)=\exp(\beta\sum_{j\in\partial i}\sigma_i^{t+1}\sigma_j^t)/Z$. In the initial state, all spins have magnetization $\langle\sigma_i^0\rangle=1/2$, i.e., $p_i(\uparrow)=3/4$. Besides being applicable for single instances of finite graphs, the MPEM approach gives also direct access to the thermodynamic limit. For disordered systems, this can be done in a population dynamics scheme. The homogeneous case, considered here, is particularly simple as all edges of the graph are equivalent in the thermodynamic limit. Hence, one can work with a single MPEM.

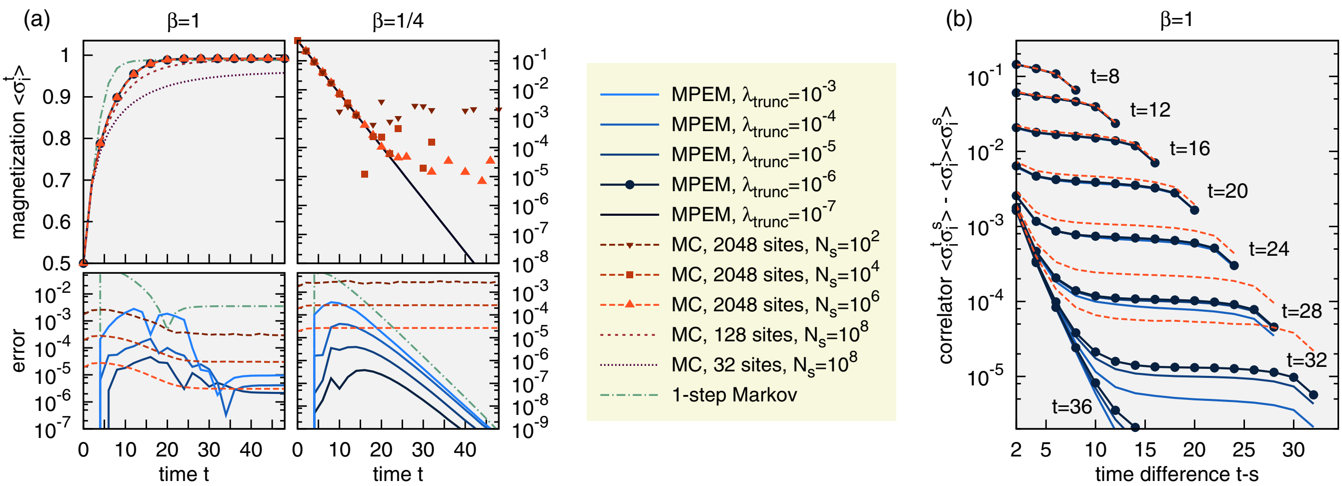

FIG. 3. (a) Magnetization and (b) connected temporal correlations for Glauber dynamics on $z=3$ random regular graphs of different sizes for MC and in the thermodynamic limit for the MPEM and one-step Markov approaches. Because of odd-even effects in the dynamics, only even time steps are shown. For MC (with N_s samples), the errors of the magnetization [lower panels in (a)] are quantified by the standard deviation of the magnetization, i.e., under ignorance of remaining finite-size effects. For MPEM and one-step Markov, errors are quantified by the deviation from the result of the most accurate (quasi-exact) MPEM simulation (truncation threshold $\lambda_{\mathrm{trunc}}=10^{-6}$ for $\beta=1$ and $\lambda_{\mathrm{trunc}}=10^{-7}$ for $\beta=1/4$). In plot (b), the three MPEM curves for $\lambda_{\mathrm{trunc}}=10^{-4},10^{-5},10^{-6}$ overlap up to time $t=24$.

Figure 3(a) shows the equilibration of the magnetization. In the ferromagnetic phase ($\beta=1$), it approaches a finite equilibrium value, whereas it decays exponentially to zero in the paramagnetic phase ($\beta=1/4$). As shown for $\beta=1$, MC simulations contain finite-size effects which become small for the system with 2048 sites. MC errors decrease slowly when increasing the number of samples N_s as $1/\sqrt{N_s}$. This is problematic for observables with small absolute values where cancellation effects make it difficult to obtain a precise estimate. This is, e.g., apparent in the magnetization decay for $\beta=1/4$ which, in contrast, is very accurately captured with MPEM. In these simulations, we control the MPEM accuracy by keeping only singular values λ_k above a threshold, specified by $\lambda_k/(\sum_{k'}\lambda_{k'}^2)^{1/2}>\lambda_{\mathrm{trunc}}$. Decreasing λ_{trunc}, increases accuracy and computation costs. The one-step Markov approximation [17] is not suited to handle temporal correlations. At long times it performs well for $\beta=1/4$ and fairly good for $\beta=1$, but deviates rather strongly at earlier times.

Figure 3(b) shows the connected temporal correlation function $\langle\sigma_i^t\sigma_i^s\rangle-\langle\sigma_i^t\rangle\langle\sigma_i^s\rangle$ for the ferromagnetic regime $\beta=1$ as a function of $t-s$ for several times t. After an initial exponential decay in $t-s$ to a value that decreases exponentially with t, the correlator continues to drop but now much more slowly, becoming almost constant. Its decay behavior can be difficult to impossible to capture with MC. In the example, MC deviations are often orders of magnitude above those of the numerically cheaper MPEM simulations. MC data for $t>32$ have been suppressed due to very big errors.

This decay of temporal correlations is also reflected in the matrix dimensions $\{M_s\}$ in MPEMs $\mu_{i\to j}(\bar{\sigma}_i^t|\bar{\sigma}_j^{t-1})$ of predefined approximation accuracy. We observe that $M:=\max_s M_s$ increases rapidly at small times t but then converges to a value that depends on the system parameters and the truncation threshold λ_{trunc}. As every iteration $t\to t+1$ requires a few sweeps through matrix products of length t, this implies that computation costs are $O(t^2)$, i.e., quadratic instead of exponential in t.

VIII. DISCUSSION

The described MPEM algorithm, based on matrix product approximations of edge messages allows for an efficient and accurate solution of the dynamic cavity equations. Besides lifting restrictions of earlier approaches for the simulation of stochastic nonequilibrium dynamics in networks, mentioned in the introduction, it gives direct access to the thermodynamic limit, and its error scaling is favorable to that of MC simulations. It allowed us to obtain quasiexact solutions of the cavity equations for Glauber-Ising dynamics. We think that this approach is a very valuable tool, particularly as it yields temporal correlations and other decaying observables with unprecedented accuracy as demonstrated in the example. It hence gives access to low-probability events. This opens a door for the study of diverse dynamic processes and inference or dynamic optimization problems for physical, technological, biological, and social networks.

ACKNOWLEDGMENTS

We thank G. Del Ferraro for providing data to benchmark our one-step Markov simulation for Fig. 3 and acknowledge support by the Marie Curie Training Network NETADIS (FP7, Grant No. 290038).

APPENDIX A: TRUNCATING MATRIX PRODUCTS

Let us explain the notion of truncations at the example of a matrix product

$$\gamma(\boldsymbol{n}):=A_0^{n_0}A_1^{n_1}\dots A_t^{n_t}, \tag{A1}$$

where $A_s^{n_s}$ is an $M_s \times M_{s+1}$ matrix and $M_0 = M_{t+1} = 1$. Our goal is to reduce in a controlled way, e.g., the left matrix dimension M_r of $A_r^{n_r}$. First, let us part the variables $\boldsymbol{n}$ into two groups $\boldsymbol{n}_L := (n_0, \dots ,n_{r-1})$ and $\boldsymbol{n}_R := (n_r, \dots ,n_t)$. For the truncation, we suggest employing a singular value decomposition (SVD) [32] of the matrix product such that

$$\gamma(\boldsymbol{n}) =: \Gamma_{\boldsymbol{n}_L,\boldsymbol{n}_R} \overset{\text{SVD}}{=} \sum_{k=1}^{M_r} Y_{\boldsymbol{n}_L,k}\lambda_k Z_{k,\boldsymbol{n}_R}. \tag{A2}$$

Y and Z are isometric matrices; i.e., they obey

$$Y^\dagger Y = \mathbb{1} \quad \text{and} \quad ZZ^\dagger = \mathbb{1}. \tag{A3}$$

Now, truncating some of the singular values $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{M_r} \geqslant 0$, such that only the M_r' largest are retained, we obtain the controlled approximation

$$\gamma_{\text{trunc}}(\boldsymbol{n}) := \sum_{k\leqslant M_r'} Y_{\boldsymbol{n}_L,k}\lambda_k Z_{k,\boldsymbol{n}_R}$$

$$\text{with error} \quad \|\gamma - \gamma_{\text{trunc}}\|^2 = \sum_{k>M_r'} \lambda_k^2. \tag{A4}$$

Note that this truncation scheme yields the minimum possible norm loss $\|\Delta\gamma\| \equiv [\sum_{\boldsymbol{n}} \Delta\gamma^2(\boldsymbol{n})]^{1/2}$ for the given new matrix dimension M_r'.

While it is very desirable to discard unimportant information and control the growth of computation cost through such truncations, the SVD (A2) appears to be an insurmountable task. Assuming that each variable n_s can take d different values and that $2r \leqslant t+1$, the cost for the SVD would scale exponentially in time like d^{t+r+1}. This is because the SVD of an $M \times N$ matrix with $M \leqslant N$ has a computation cost $O(M^2N)$ [32]. However, the beauty of matrix products is that such an SVD can in fact be done sequentially with linear costs of order tdM^3 as follows. Here, $M := \max_s M_s$ is the maximum matrix dimension in Eq. (A1).

First, we do an exact transformation of the matrix product (A1) to bring it to the orthonormalized form

$$\gamma(\boldsymbol{n}) = Y_0^{n_0}\dots Y_{r-1}^{n_{r-1}}\tilde{A}_r^{n_r}Z_{r+1}^{n_{r+1}}\dots Z_t^{n_t}, \tag{A5}$$

where tensors Y_s and Z_s obey the left and right orthonormality constraints

$$\sum_n \left(Y_s^n\right)^\dagger Y_s^n = \mathbb{1} \quad \text{and} \quad \sum_n Z_s^n\left(Z_s^n\right)^\dagger = \mathbb{1}, \tag{A6}$$

respectively. This is achieved through a sequence of SVDs. It starts with the SVD $A_0^{n_0} =: Y_0^{n_0}\Lambda_0 V_0$, where Λ_0 is a diagonal matrix containing the singular values, V_0 is isometric according to $V_0V_0^\dagger = \mathbb{1}$, and Y_0 obeys Eq. (A6). The sweep continues with the SVD $\Lambda_0V_0A_1^{n_1} =: Y_1^{n_1}\Lambda_1V_1$ and so on until the computation of Y_{r-1}. Analogously, we do a second sequence of SVDs starting from the right with $A_t^{n_t} =: U_t\Lambda_tZ_t^{n_t}$, where $Z_t^{n_t}$ obeys Eq. (A6), and continue with $A_{t-1}^{n_{t-1}}U_t\Lambda_t =: U_{t-1}\Lambda_{t-1}Z_{t-1}^{n_{t-1}}$ and so on until Z_{r+1} has been computed with $A_{r+1}^{n_{r+1}}U_{r+2}\Lambda_{r+2} =: U_{r+1}\Lambda_{r+1}Z_{r+1}^{n_{r+1}}$. Finally, we define the central tensor as $\tilde{A}_r^{n_r} := \Lambda_{r-1}V_{r-1}A_r^{n_r}U_{r+1}\Lambda_{r+1}$, and we have thus determined all matrices in Eq. (A5). After this preparation, we can do the actual truncation, based on the SVD $\tilde{A}_r^{n_r} =: U_r\Lambda Z_r^{n_r}$ with the same singular values $\lambda_1 \geqslant \cdots \geqslant \lambda_{M_r}$ as in Eq. (A2). With the $M_r \times M_r'$ matrix $[\Lambda_{\text{trunc}}]_{kk'} := \delta_{kk'}\lambda_k$, the truncated matrix product (A4) takes the form

$$\gamma_{\text{trunc}}(\boldsymbol{n}) = Y_0^{n_0}\dots Y_{r-2}^{n_{r-2}}\left(Y_{r-1}^{n_{r-1}}U_r\Lambda_{\text{trunc}}\right)Z_r^{n_r}\dots Z_t^{n_t}.$$

APPENDIX B: PROCESSING EVOLVED MPEMS

In the evolution step described in the main text, matrix dimensions are increased to $\bar{M}_s$ and the evolved edge message (4) is in a noncanonical form. Here we discuss how to apply the truncation as described in Appendix A to compress the evolved MPEM and bring it back to canonical form (3).

In a first sweep from right ($s = t+1$) to left ($s = 0$), using SVDs, we can sequentially impose the right orthonormality constraints [see Eq. (A6)] on the C tensors. In a subsequent sweep from left to right, again based on SVDs, at each step, the MPEM is in orthonormalized form [see Eq. (A5)] and we can now truncate the tensors to decrease bond dimensions from $\bar{M}_s$ to something smaller. According to the triangle inequality, the norm distance of the original edge message $\mu_{i\to j}(\bar{\sigma}_i^{t+1}|\bar{\sigma}_j^t)$ and the resulting truncated MPEM are bounded from above by the sum of errors (A4) of the individual truncations.

What remains is to reorder the indices $\{\sigma_i^s\}$ and $\{\sigma_j^s\}$ of the vertex variables. In a sweep from right to left, we go from the noncanonical variable assignment $(\sigma_i^0)(\sigma_i^1|\sigma_j^0)\dots(\sigma_i^{t+1}|\sigma_j^t)$ in the truncated and orthonormalized version $\tilde{C}_{i\to j}^{(0)}(\sigma_i^0)\prod_{s=1}^{t+1}\tilde{C}_{i\to j}^{(s)}(\sigma_i^s|\sigma_j^{s-1})$ of the MPEM (4) to the assignment $(\sigma_i^0\sigma_j^0)\dots(\sigma_i^t|\sigma_j^t)(\sigma_i^{t+1})$, yielding the matrix product

$$\mu_{i\to j}\left(\bar{\sigma}_i^{t+1}\middle|\bar{\sigma}_j^t\right) \overset{\text{trunc}}{\approx} \left[\prod_{s=0}^{t} D_{i\to j}^{(s)}\left(\sigma_i^s\middle|\sigma_j^s\right)\right] D_{i\to j}^{(t+1)}\left(\sigma_i^{t+1}\right).$$

At the right boundary, we start with an SVD and controlled truncation $\tilde{C}_{i\to j}^{(t+1)}(\sigma_i^{t+1}|\sigma_j^t) \approx: U^{(t+1)}(\sigma_j^t)\Lambda_{\text{trunc}}^{(t+1)}D_{i\to j}^{(t+1)}(\sigma_i^{t+1})$, and continue with

$$\begin{aligned}\tilde{C}_{i\to j}^{(t)}&\left(\sigma_i^t\middle|\sigma_j^{t-1}\right)U^{(t+1)}\left(\sigma_j^t\right)\Lambda_{\text{trunc}}^{(t+1)}\\ &\approx: U^{(t)}\left(\sigma_j^{t-1}\right)\Lambda_{\text{trunc}}^{(t)}D_{i\to j}^{(t)}\left(\sigma_i^t\middle|\sigma_j^t\right)\end{aligned}$$

and so on until ending at $s = 0$. In an analogous final sweep from left to right, we change to the canonical variable assignment $(\sigma_j^0)(\sigma_i^0|\sigma_j^1)\dots(\sigma_i^{t-1}|\sigma_j^t)(\sigma_i^t)(\sigma_i^{t+1})$ as in Eq. (3). After executing these steps for all edge messages, the next evolution step from $t+1$ to $t+2$ can follow.

[1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, New York, 2008).

[2] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, UK, 2010).

[3] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, New York, 2009).

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

[5] M. Mézard, G. Parisi, and R. Zecchina, Analytic and algorithmic solution of random satisfiability problems, Science **297**, 812 (2002).

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

[7] T. Rogers, I. P. Castillo, R. Kühn, and K. Takeda, Cavity approach to the spectral density of sparse symmetric random matrices, Phys. Rev. E **78**, 031116 (2008).

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

[9] B. Karrer and M. E. J. Newman, Message passing approach for general epidemic models, Phys. Rev. E **82**, 016101 (2010).

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

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

[12] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Lage-Castellanos, and R. Zecchina, Bayesian Inference of Epidemics on Networks via Belief Propagation, Phys. Rev. Lett. **112**, 118701 (2014).

[13] M. Shrestha and C. Moore, Message-passing approach for threshold models of behavior in networks, Phys. Rev. E **89**, 022805 (2014).

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

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

[16] P. Zhang, Inference of kinetic Ising model on sparse graphs, J. Stat. Phys. **148**, 502 (2012).

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

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

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

[20] A. Pelizzola and M. Pretti, Variational approximations for stochastic dynamics on graphs, J. Stat. Mech. (2017) 073406.

[21] Y. Roudi and J. Hertz, Dynamical TAP equations for non-equilibrium Ising spin glasses, J. Stat. Mech. (2011) P03031.

[22] L. Bachschmid-Romano, C. Battistin, M. Opper, and Y. Roudi, Variational perturbation and extended Plefka approaches to dynamics on random networks: The case of the kinetic Ising model, J. Phys. A **49**, 434003 (2016).

[23] L. Accardi, Topics in quantum probability, Phys. Rep. **77**, 169 (1981).

[24] M. Fannes, B. Nachtergaele, and R. F. Werner, Finitely correlated states on quantum spin chains, Commun. Math. Phys. **144**, 443 (1992).

[25] S. Rommer and S. Östlund, A class of ansatz wave functions for 1D spin systems and their relation to DMRG, Phys. Rev. B **55**, 2164 (1997).

[26] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).

[27] I. V. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. **33**, 2295 (2011).

[28] S. R. White, Density-Matrix Formulation for Quantum Renormalization Groups, Phys. Rev. Lett. **69**, 2863 (1992).

[29] S. R. White, Density-matrix algorithms for quantum renormalization-groups, Phys. Rev. B **48**, 10345 (1993).

[30] U. Schollwöck, Density-matrix renormalization-group, Rev. Mod. Phys. **77**, 259 (2005).

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

[32] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Baltimore, MD, 1996).