POLITECNICO DI TORINO
Repository ISTITUZIONALE

Matrix Product Belief Propagation for reweighted stochastic dynamics over graphs

Original
Matrix Product Belief Propagation for reweighted stochastic dynamics over graphs / Crotti, S., Braunstein, A.. - In: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. - ISSN 0027-8424. - 120:47(2023), pp. 1-9. [10.1073/pnas.2307935120]

Availability:
This version is available at: 11583/2990678 since: 2024-07-11T15:05:16Z

Publisher:
PNAS

Published
DOI:10.1073/pnas.2307935120

Terms of use:

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

Publisher copyright

(Article begins on next page)

06 November 2026

PNAS RESEARCH ARTICLE | PHYSICS

Matrix Product Belief Propagation for reweighted stochastic dynamics over graphs

Stefano Crotti[a,1,2] and Alfredo Braunstein[a,b,1]

Edited by Cristopher Moore, Santa Fe Institute, Santa Fe, NM; received May 15, 2023; accepted October 2, 2023 by Editorial Board Member Linda R. Petzold

Stochastic processes on graphs can describe a great variety of phenomena ranging from neural activity to epidemic spreading. While many existing methods can accurately describe typical realizations of such processes, computing properties of extremely rare events is a hard task, particularly so in the case of recurrent models, in which variables may return to a previously visited state. Here, we build on the matrix product cavity method, extending it fundamentally in two directions: First, we show how it can be applied to Markov processes biased by arbitrary reweighting factors that concentrate most of the probability mass on rare events. Second, we introduce an efficient scheme to reduce the computational cost of a single node update from exponential to polynomial in the node degree. Two applications are considered: inference of infection probabilities from sparse observations within the SIRS epidemic model and the computation of both typical observables and large deviations of several kinetic Ising models.

stochastic dynamics | nonequilibrium statistical physics | Bayesian inference

Significance

When dealing with stochastic processes on graphs, two elements can render the estimation of observables computationally hard: first, the presence of reweighting terms that bias the probability measure toward rare events and, second, the possibility for variables to return to a previously visited state. While in the past, several works have successfully dealt with either of these two factors, the method presented here addresses both simultaneously. It is a flexible approach, applicable to a wide range of relevant problems, also beyond the two presented in the article: inference in epidemics and analysis of spin dynamics.

The problem of computing observables and marginal probabilities on a complex Markov process on large networks has been addressed extensively in the literature. While Monte-Carlo procedures can be often effective to compute averages approximately, they suffer from two separate issues: large relative sampling errors when computing averages that cancel out at the first order and they are limited to sampling "typical" events, as nontypical ones require an exponential number of samples. To address the first issue, many analytical solutions, mainly based on mean-field methods, have been devised (1–7). A solution that is exact on acyclic graphs is Dynamic Cavity (DC) (8). DC on general processes suffers from one main drawback, the fact that one must be able to represent the joint distribution of a single variable trajectory and a feedback field, and with some exceptions, the space of these trajectories is exponentially large (in the time horizon), and thus the approach becomes impracticable. One of these exceptions is on "nonrecurrent" models, i.e., models in which each variable can only progress sequentially through a finite set of k states, never going back to a previous state. In these cases, the set of trajectories is polynomial in the time horizon [as an example with $q = 3$, a trajectory $(1, 1, 2, 2, 2, 2, 3, 3)$ on epochs $t = 0, \dots, 7$ can be represented by the integer tuple $(2, 6)$ of epochs on which the variable effectively progresses to the next state in the sequence]. Examples of nonrecurrent models are the SI, SIR, SEIR compartmental models in computational epidemics, in which an individual can only transition from Susceptible to Exposed, from Exposed to Infective and from Infective to Recovered. While the use of nonrecurrent models is pervasive, oftentimes a more realistic description demands that reinfections be taken into account. In such cases, "recurrent" models such as the SIS and SIRS are employed. Additionally, important processes in statistical physics such as Glauber dynamics belong to the class of models with recurrence.

In a recent work (9, 10), an interesting DC variant was proposed that exploits the Matrix Product State representation (MPS) to parametrize site trajectories and applied it to the Glauber dynamics on a Random Regular (RR) graph with degree 3. While these results are promising, the scheme suffers from two major limitations: First, it is computationally expensive (the update on a node of degree z is of the order of M^{2z-1} (10) where M is the matrix dimension), making it impractical even for moderately large Erdos-Renyi (ER) random graphs, in which some large-degree vertices are surely present. Second, the scheme is devised to analyze a "free" dynamics without any sort of reweighting, which as we will see is necessary to study atypical trajectories. Matrix Product States, also known as Tensor Trains, are not new in physics and other areas of science, as they have been successfully applied both in many-body quantum systems (11–13), out-of-equilibrium statistical physics (14, 15), machine learning (16, 17) and more.

Author affiliations: [a]Department of Applied Science and Technology, Politecnico di Torino, Turin 10129, Italy; and [b]Italian Institute for Genomic Medicine, Turin 10126, Italy

Author contributions: S.C. and A.B. designed research; performed research; and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. C.M. is a guest editor invited by the Editorial Board.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]S.C. and A.B. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: stefano.crotti@polito.it.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2307935120/-/DCSupplemental.

Published November 14, 2023.

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

We propose an alternative approach, dubbed Matrix Product Belief Propagation (MPBP), based on the Pair Trajectory Belief Propagation formulation which was first introduced in ref. 18. It is closely related to DC but allows naturally to include nonnegative reweighting terms on stochastic trajectories, thus allowing to study large deviations of the system. In practical terms, MPBP consist on a fixed point equation that is solved by iteration, whereas DC is solved sequentially in time, with a number of steps which is equal to the number of epochs of the dynamics. The latter approach is inherently limited to free dynamics: Building trajectories sequentially in time makes it impossible to account for the effect of reweighting terms relative to future epochs.

The Julia code used to implement the method and produce the data presented in this work is publicly accessible at ref. 19.

We describe in the following the models under consideration. Given a graph $G = (V, E)$ with $V = \{1, \dots, N\}$, consider a joint distribution over a set of discrete variables $\boldsymbol{x} = \{x_1, \dots, x_N\}$ throughout T successive epochs of the form

$$p(\bar{\boldsymbol{x}}) = \frac{1}{Z} \prod_{t=0}^{T-1} \prod_{i=1}^{N} f_i^{t+1}\left(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t\right). \quad [1]$$

We use bold letters to indicate multiple variable indices $\boldsymbol{x}_A \equiv \{x_j\}_{j\in A}$ and overbars for multiple times indices $\bar{x} \equiv \{x^t\}_{t=0:T}$. Moreover, we indicate by $\partial i = \{j : (ij) \in E\}$ the set of neighbors of index i.

The form [**1**] includes (but notably is more general than) reweighted Markov dynamics $f_i^{t+1}(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t) = w(x_i^0)^{\delta(t,0)} w(x_i^{t+1}|\boldsymbol{x}_{\partial i}^t, x_i^t)\phi_i^{t+1}(x_i^{t+1})$ with stochastic transitions w and reweighting factors ϕ

$$p(\bar{\boldsymbol{x}}) = \frac{1}{Z} \prod_{i=1}^{N} w(x_i^0) \prod_{t=0}^{T-1} w(x_i^{t+1}|\boldsymbol{x}_{\partial i}^t, x_i^t)\phi_i^{t+1}(x_i^{t+1}). \quad [2]$$

$\delta(y, z)$ is the Kroenecker delta which evaluates to 1 if $y = z$, to 0 otherwise, and $w(x_i^0)$ is the initial state probability, which we take to be factorized over the sites.

Note that $Z = 1$ in the absence of reweighting factors. Two types of reweighted dynamics of the form [**2**] will be used as running examples throughout this work. The first is Bayesian inference on a process of epidemic spreading. The posterior probability of the epidemic trajectory $\bar{\boldsymbol{x}}$ given some independent observations $\{O_i^t\}$ on the system is given by

$$p(\bar{\boldsymbol{x}}|O) = \frac{1}{p(O)} p(\bar{\boldsymbol{x}}) p(O|\bar{\boldsymbol{x}}). \quad [3]$$

Eq. **3** can be seen as a particular case of Eq. **2**, where $p(\bar{\boldsymbol{x}}) = \prod_{i=1}^{N} w(x_i^0) \prod_{t=0}^{T-1} w(x_i^{t+1}|\boldsymbol{x}_{\partial i}^t, x_i^t)$ and corresponds to the distribution of the free dynamics of the chosen epidemiological model, $p(O|\bar{\boldsymbol{x}}) = \prod_i \prod_t p(O_i^t|x_i^t) = \prod_i \prod_t \phi_i^t(x_i^t)$ and $Z = p(O)$.

The simplest among the recurrent epidemiological models is the Susceptible-Infectious-Susceptible (SIS), where each individual starts with a probability γ_i of being infectious at time zero. Then, at each time step, a susceptible node i can be infected by each of its infectious neighbors $j \in \partial i$ with probability λ_{ji}, and an infectious node can recover with probability ρ_i. Observation terms $p(O_i^t|x_i^t)$ are naturally used to model medical tests: O_i^t is the outcome of a test performed on individual i at time t. This formalism allows to incorporate information about the degree of accuracy of tests.

The second example is parallel Glauber dynamics for an Ising model at inverse temperature β with couplings $\{J_{ij}\}$ and external fields $\{h_i\}$. Besides being one of the paradigmatic models in theoretical nonequilibrium statistical physics, Glauber dynamics is employed in the study of neural activity (20, 21). It is defined by transitions

$$\tilde{w}(\sigma_i^{t+1}|\boldsymbol{\sigma}_{\partial i}^t) = \frac{e^{\beta\sigma_i^{t+1}\left(\sum_{j\in\partial i} J_{ij}\sigma_j^t + h_i\right)}}{2\cosh\left[\beta\left(\sum_{j\in\partial i} J_{ij}\sigma_j^t + h_i\right)\right]}. \quad [4]$$

The dynamics does not converge to the equilibrium of the underlying Ising model $p_{J,h}(\boldsymbol{\sigma}) = Z^{-1}\exp[-H_{J,h}(\boldsymbol{\sigma})]$, but it allows to compute observables of interest in some cases (*SI Appendix*).

Moreover, we will allow σ_i to stay in the same state with probability p_0. The transition thus becomes

$$w(\sigma_i^{t+1}|\boldsymbol{\sigma}_{\partial i}^t, \sigma_i^t) = (1 - p_0)\tilde{w}(\sigma_i^{t+1}|\boldsymbol{\sigma}_{\partial i}^t) + p_0\delta(\sigma_i^{t+1}, \sigma_i^t). \quad [5]$$

In the limit $p_0 \to 0$, the stationary distribution converges to $p_{J,h}$ because the dynamics reduces to an asynchronous one (two or more simultaneous state changes happen with probability $\mathcal{O}(p_0^2)$), see also *SI Appendix*.

Additionally, such dynamics can be "tilted" with, e.g., a term $\prod_i \phi_i^T(\sigma_i^T) = \prod_i e^{h\sigma_i^T}$ in order to study atypical trajectories. Note that other models studied in physics such as Bootstrap Percolation can be remapped into Glauber dynamics (22).

Related Work

Mean-field Methods. We briefly review the main features of existing approaches based on the cavity method. Dynamic Message Passing (DMP) (2, 23, 24) and the Cavity Master Equation (6, 7) are simple and fast approximate methods that were originally formulated on continuous time as ODEs for a vector of single-edge quantities (such as cavity magnetizations). Both methods are exact on acyclic graphs on nonrecurrent models (such as SI or SIR), but only approximate on non-non-recurrent ones, and do not allow for atypical trajectories. n-step Dynamic Message Passing (3) makes an n-Markov *ansatz* on messages, exploring mainly $n = 1$; its features are essentially those of DMP, with the difference that it applies to discrete time evolution and describes explicitly interactions at distance n in time. Different flavors of the cluster variational method (5, 25) approximate the dynamics by treating exactly correlations between variables that are close either in time or space. Large deviations have been studied in ref. 26 using a perturbation theory in the particular case of Glauber dynamics on a chain. Table 1 summarizes the features of the methods mentioned above. We take into consideration ability to deal with reweighted dynamics, to deal with recurrent models, and to compute autocorrelations at arbitrary (time) distance.

Monte Carlo. Throughout this work, the performance of algorithms is compared with Monte Carlo simulations. To estimate observables in a reweighted dynamics of the form (**2**) we employ a weighted sampling technique (see, e.g., ref. 27): the posterior average of an observable f is approximated by

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

Table 1. Features of existing analytical methods for the description of stochastic dynamics on graphs, Y for yes, N for no

	Reweighting	Recurrent models	Autocorrelations
BP for nonrecurrent models (18)	Y	N	Y
IBMF (1), DMP (2, 3, 24), CME (6)	N	Y	N
Dynamic Cluster Variational (5)	*	Y	Only two-times
Matrix Product DC (9)	N	Y	Y
Matrix Product Belief Propagation	Y	Y	Y

The asterisks mean that the method could in principle be extended to include the considered feature although this has not, to the best of our knowledge, been done in the literature. IBMF stands for Individual-Based Mean Field, DMP for Dynamic Message Passing, and CME for Cavity Master Equation. We did not include the perturbative approach (26) because it focuses on a very particular setting.

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

$$\hat{f} = \frac{\sum_{\mu=1}^{M} \prod_{i,t} \phi_i^t \left((x_i^t)^{(\mu)} \right) f\left(\bar{\boldsymbol{x}}^{(\mu)} \right)}{\sum_{\mu=1}^{M} \prod_{i,t} \phi_i^t \left((x_i^t)^{(\mu)} \right)}, \qquad [6]$$

where $\left\{ \bar{\boldsymbol{x}}^{(\mu)} \right\}_{\mu}$ are M independent samples drawn from the prior $\prod_{i=1}^{N} w(x_i^0) \prod_{t=0}^{T-1} w(x_i^{t+1} | \boldsymbol{x}_{\partial i}^t, x_i^t)$. Such strategy, however, turns out to be computationally prohibitive whenever the reweighting terms ϕ put most of the probability mass on atypical trajectories, which are (exponentially) unlikely to ever be sampled.

Matrix Product Belief Propagation

For the dynamic version of Belief Propagation (BP), we start with Eq. **1** as a distribution for single site trajectories $\bar{x}_i$. The associated factor graph would present many small loops due to the presence of both $\bar{x}_i$ and $\bar{x}_j$ in factors f_i and f_j. Therefore, we work directly on the so-called dual factor graph where variables are pair of trajectories $(\bar{x}_i, \bar{x}_j)$ living on the edges of the original graph. For more details about this step we refer the reader to ref. 18, figure 3 and equations 8 and 9. The BP equations on the dual factor graph read

$$m_{i\to j}(\bar{x}_i, \bar{x}_j) \propto \sum_{\bar{x}_{\partial i \setminus j}} \prod_{t=0}^{T-1} f_i^{t+1}(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t) \times \prod_{k \in \partial i \setminus j} m_{k\to i}(\bar{x}_k, \bar{x}_i). \qquad [7]$$

Since the number of joint trajectories $(\bar{x}_i, \bar{x}_j)$ is exponentially large in T, an exact representation of the messages is in general computationally unfeasible. Here, similarly to ref. 9, we parametrize messages in terms of matrix product states (11–13), also known as tensor trains in the mathematical literature (28). Following the jargon of tensor networks, in the rest of the paper, we will refer to the size of the matrices as bond dimension. For a wide class of dynamics including Glauber with $J_{ij} = \pm J$ and epidemic spreading with homogeneous infectivity, the computational cost for a single BP iteration is $\mathcal{O}(T|E|M^6)$ where T is the number of epochs, $|E|$ is the number of edges in the graph and M is the bond dimension. In all the applications we considered, small bond dimension (scaling at most polynomially with T) was enough to obtain almost exact results. The full description of the approach is found in *Materials and Methods*.

Results

In this section, we illustrate the effectiveness of MPBP applied to dynamics of epidemic spreading and of the kinetic Ising model. We first focus on free dynamics, showing results that are at least comparable with the existing methods, often more accurate. Then, we move to reweighted processes, where our approach really represents an innovation.

Risk Assessment in Epidemics. As examples of free dynamics, we estimate the marginal probability of an individual being in the infectious state under the SIS model, in several settings (Fig. 1). On a random tree and on a diluted random graph, both of size $N = 1{,}000$, MPBP shows almost no discrepancy with Monte Carlo averages (Fig. 1 *A* and *B*). In the former case, a single node was picked as the sole infectious individual at time zero, in the latter a uniform probability $\gamma_i \equiv \gamma$ was put on each node. As a comparison, we report the curves obtained using a discrete-time version of Dynamic message Passing (DMP) (24), Individual-Based Mean Field (IBMF) (1), and Cavity Master Equation (CME) (6), which were originally devised for continuous time evolution (more details in *SI Appendix*). We evaluate the accuracy of each method by considering the average absolute error with respect to a Monte Carlo simulation $\frac{1}{N} \sum_{i=1}^{N} \left| p(x_i^t = I) - p^{MC}(x_i^t = I) \right|$ (*Insets* of Fig. 1). The same analysis is repeated on Zachary's karate club graph (29) (Fig. 1*C*), the same benchmark used in (7, 24). It must be pointed out that although MPBP shows by far the best performance in these comparisons, the other considered methods are significantly simpler. None of the analytic methods is devised to analyze reweighted dynamics. Finally, we compare MPBP against three continuous-time methods, DMP, IBMF, and CME, on the karate club graph (Fig. 1*D*). The comparison is made by multiplying the transmission and recovery rates for the continuous setting λ, ρ by the time-step Δt (in this case $\Delta t = 1$) to turn them into probabilities to be handled by MPBP. MPBP gives the best overall prediction across the considered window.

Moving to reweighted processes, Fig. 2 shows the efficacy of MPBP when performing inference of trajectories given some observations. On a small ($N = 23$) random graph, a 10-step trajectory $\bar{\boldsymbol{y}}$ was sampled from a SIS prior distribution with $\lambda = 0.15, \rho = 0.12, \gamma = 0.13$. We then observed the state of a random half $I \subset V$ of the nodes, added the corresponding reweighting factors $\prod_{i \in I} \phi_i^T(x_i^T) = \prod_{i \in I} \delta(y_i^T, x_i^T)$ and performed inference using Eq. **3**.

The MPBP estimate for the posterior marginals, obtained with matrices of size 3, agrees almost perfectly with Monte Carlo simulations. This is good indication that MPBP applied to sparse problems will keep giving accurate results even when on larger and/or more constrained instances where Monte Carlo methods fail, leaving little to compare against.

Realistic scenarios are often better described by the Susceptible-Infectious-Recovered-Susceptible (SIRS) model where transmis-

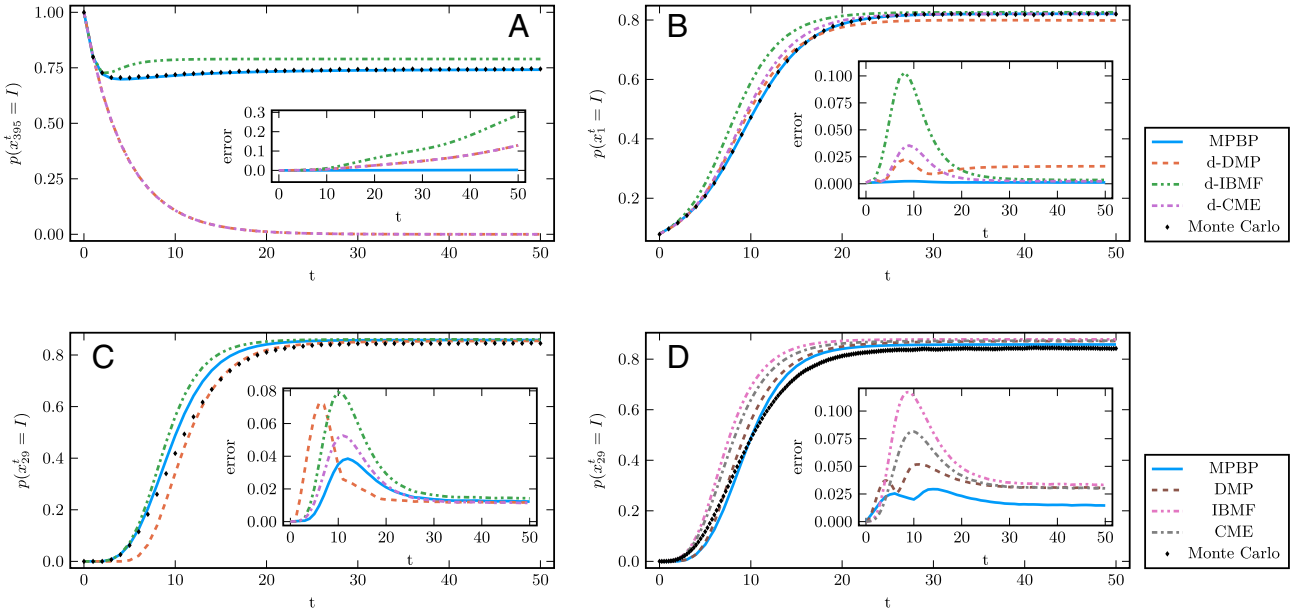

Fig. 1. Marginal probabilities of free dynamics under the SIS model, comparison with models mentioned in the text. The main panels correspond to marginals for a single node of the graph, *Insets* show the average absolute error over all nodes with respect to Monte Carlo simulations. Panels (*A*–*C*) compare against discretized versions of DMP, IBMF, and CME (here with a "d-" prefix) and the Monte Carlo strategy reported in the text, panel (*D*) against regular continuous-time versions and a Gillespie-like Monte Carlo simulation. (*A*) Marginal of node 395, the most connected one of a random tree with $N = 1{,}000$ nodes, $\lambda = 0.3$, $\rho = 0.2$. Node 395 is the only infectious at time zero. Bond dimension 12. (*B*) Marginal of node 1 of a ER graph with $N = 1{,}000$ nodes, average connectivity $c = 5$, $\lambda = 0.1$, $\rho = 0.05$, $\gamma = 0.08$. Bond dimension 10. (*C*) Marginal of node 29 (zero-based numbering to match previous works) of Zachary's karate club network, $N = 34$ nodes, $\lambda = 0.1$, $\rho = 0.05$, node 0 is the only infectious at time zero. Bond dimension 10. (*D*) Same as (*C*) but the comparison is with continuous-time methods, with the addition of CME.

sion of infections is analogous to the SIS case, but an infectious node i can recover with probability ρ_i and a recovered become susceptible again with probability σ_i. From a practical point of view, extending the SIR to SIRS in the MPBP framework takes little effort: It suffices to enrich the factors with the new transition $R \to S$. Fig. 3 shows the performance of MPBP at estimating the posterior trajectories for a single realization of an epidemic drawn from a prior whose parameters $\lambda, \rho, \sigma, \gamma$ are homogeneous and known.

The state of a random 75% of the system was observed at an intermediate time (colored dots). We see good agreement between the true infection times (black lines) and the marginal probabilities of being Infectious (in yellow). Nodes are sorted in increasing order of true first infection time.

Kinetic Ising. As examples of free dynamics we consider the evolution of magnetization $\langle \sigma_i^t \rangle$ and time autocovariance $\langle \sigma_i^t \sigma_i^s \rangle - \langle \sigma_i^t \rangle \langle \sigma_i^s \rangle$ for pairs of epochs (t, s), on ferromagnetic, Random Field and spin-glass Ising Models (Fig. 4), under the stochastic transition Eq. **5**. First, we consider a model with uniform couplings $J_{ij} \equiv J$ on an infinite Random Regular Graph like the one studied in ref. 9 but with degree 8 instead of 3. We then apply our method to an infinite Erdos-Renyi graph, again with uniform couplings and in the ferromagnetic phase, using a population dynamics approach. Next, we study a Random Field Ising Model (RFIM) with uniform couplings and random external fields $h_i = \pm h$ on a large graph. In all three cases, the system is initialized in a magnetized state and the fraction of up spins grows or decreases monotonically until it reaches a stationary

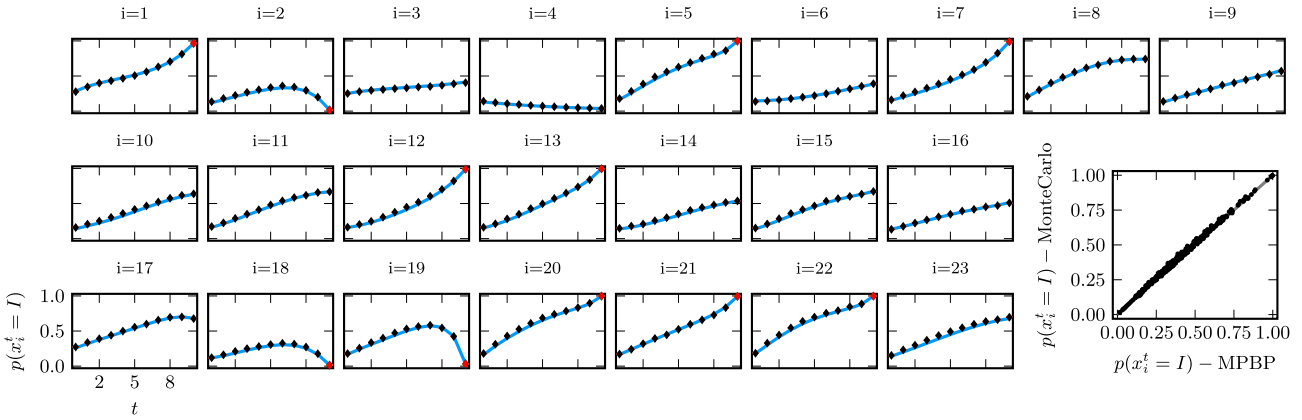

Fig. 2. MPBP (solid line) with bond dimension 3 correctly computes marginals of an SIS model defined on an Erdos-Renyi graph with 23 nodes and average connectivity 4, $\lambda = 0.15$, $\rho = 0.12$, $\gamma = 0.13$. The state of a random half of the variables was observed at final time $T = 10$ and used to reweight the distribution (red dots). Black dots are the average over 10^6 Monte Carlo simulations. (*Bottom-right*) Comparison of all points from the previous plots, the Pearson correlation coefficient is 0.9986.

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

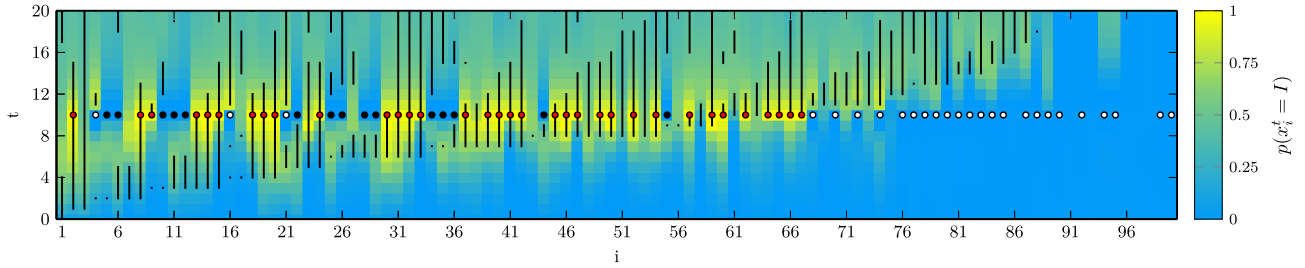

Fig. 3. Inference on a single epidemic outbreak sampled from a SIRS model on an Erdos-Renyi graph with average connectivity $c = 2.5$, $N = 100$. Bond dimension $M = 3$. The process to be inferred was drawn from a SIRS prior with $\lambda = 0.4, \rho = \sigma = 0.15, \gamma = 0.01$, the same parameters were used for the inference. The state of 75% of the nodes was observed at time 10 (white = S, red = I, black = R) and used to reweight the distribution. Black lines correspond to true infection periods.

value. For these second and third models, we picked the same settings as in ref. 25. Finally, we consider an antiferromagnetic model with $J = -1$ at zero temperature ($\beta = \infty$), focusing on the nearest-neighbor correlation $\langle \sigma_i^t \sigma_j^t \rangle$, $(i,j) \in E$ rather than the magnetization, which is null at steady state. Above the critical inverse temperature $\beta_c = \log\left(1+\sqrt{2}\right)$ (30), the underlying Ising system is in a glassy phase. For this model, we used the modified version of the dynamics reported in Eq. **5** with $p_0 = 0.25$.

Finally, we study the large deviation behavior of a free dynamic $W(\overline{\boldsymbol{\sigma}}) = \prod_{i=1}^{N} w(\sigma_i^0) \prod_{t=0}^{T-1} w(\sigma_i^{t+1} | \boldsymbol{\sigma}_{\partial i}^t)$ by tilting it with an external field at final time $\prod_i \phi_i^T(\sigma_i^T) = \prod_i e^{h\sigma_i^T}$. In the thermodynamic limit $N \to \infty$ this allows to select a particular value for the magnetization at final time $m = \frac{1}{N}\sum_i \sigma_i^T$. The Bethe Free Energy computed via MPBP is an approximation for

$$f(h) = -\frac{1}{N} \log \sum_{\{\sigma_i^t\}_{i,t}} W(\overline{\boldsymbol{\sigma}}) e^{h\sum_i \sigma_i^T}, \qquad [8]$$

$$= -\frac{1}{N} \log \sum_{m} e^{-N[g(m)-hm]}, \qquad [9]$$

$$\xrightarrow{N\to\infty} \min_m \left\{ g(m) - hm \right\}, \qquad [10]$$

$$= g(m(h)) - hm(h), \qquad [11]$$

where $g(m) = -\frac{1}{N} \log \sum_{\{\sigma_i^t\}_{i,t}} W(\overline{\boldsymbol{\sigma}}) \delta\left(Nm, \sum_i \sigma_i^T\right)$, and $m(h) = \arg\min_m \left\{ g(m) - hm \right\}$. In regions where $g(m)$ is convex, the Legendre transform (10) can be inverted to obtain a large deviation law for the probability of observing the system at final time with magnetization m

$$p(m) \sim e^{-N[f(h(m))+mh(m)]} \qquad [12]$$

where $h(m)$ is the inverse of $m(h)$. Fig. 5 shows the estimate of $g(m)$ for a ferromagnetic Ising model on an infinite random graph initialized at magnetization $m^0 = 0.1$ and evolving for $T = 10$ epochs. $p(m)$ has a minimum at $m \approx 0.145$ which corresponds to the free dynamics $h = 0$.

Such an analysis could not have been carried out by means of Monte Carlo methods since the probability of sampling a trajectory ending at m is infinitesimal, as is clear from the large deviation law in Fig. 5.

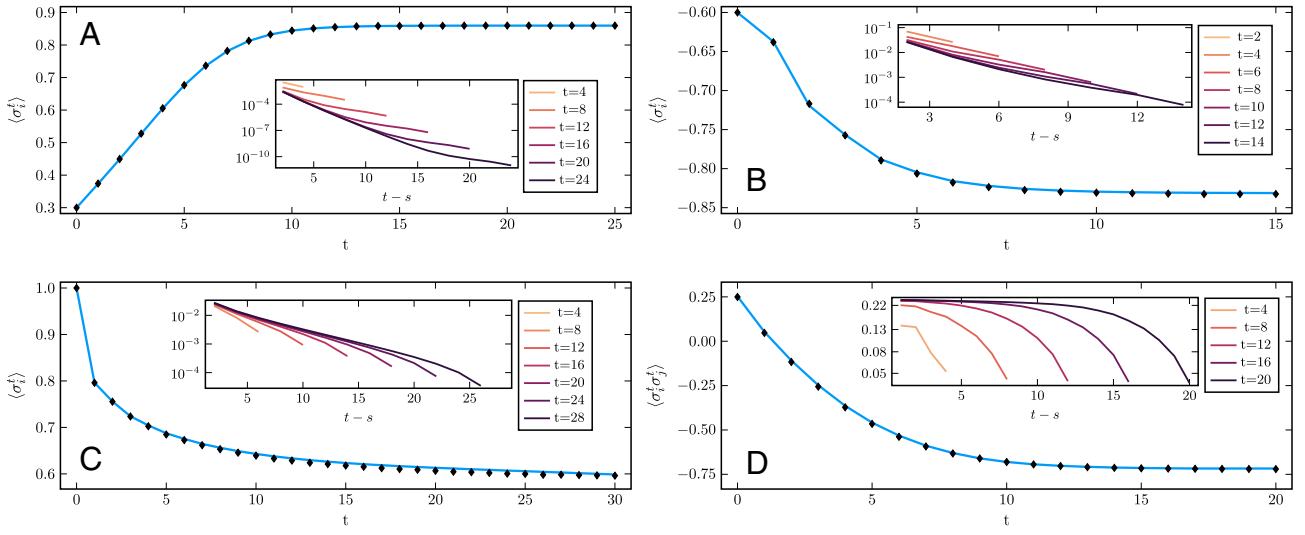

Fig. 4. Magnetization $\langle \sigma_i^t \rangle$ (*A*–*C*) or nearest-neighbor correlation $\langle \sigma_i^t \sigma_j^t \rangle$ (*D*) as a function of time for different Ising models. Solid lines are MPBP, dots are Monte Carlo simulations on graphs of size N_{MC}, dashed horizontal lines are the equilibrium values (*A*–*C*) or 1RSB prediction (*D*) for the corresponding static versions of the models. Insets show autocovariances $\langle \sigma_i^t \sigma_i^s \rangle - \langle \sigma_i^t \rangle \langle \sigma_i^s \rangle$, only even epochs are shown for panels (*A*–*C*) because of odd-even effects in the dynamics of ferromagnetic models (as in refs. 9 and 25). (*A*) Infinite 8-Random Regular Graph, $\beta J = 0.2$, $N_{MC} = 5{,}000$, bond dimension 25. (*B*) Infinite Erdos-Renyi graph with mean connectivity $c = 4$, $\beta J = 0.5$, $N_{MC} = 5{,}000$, bond dimension 18. (*C*) Random Field Ising Model on Erdos-Renyi graph with mean connectivity $c = 3$, $\beta J = 2/c$, $N = N_{MC} = 1{,}000$ and $\beta h_i = \pm 0.6$ sampled uniformly, matrix size 10. (*D*) Antiferromagnetic Ising Model on infinite 3-Random Regular Graph with $J = -1$, $\beta = \infty$, $N_{MC} = 5{,}000$, bond dimension 23.

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

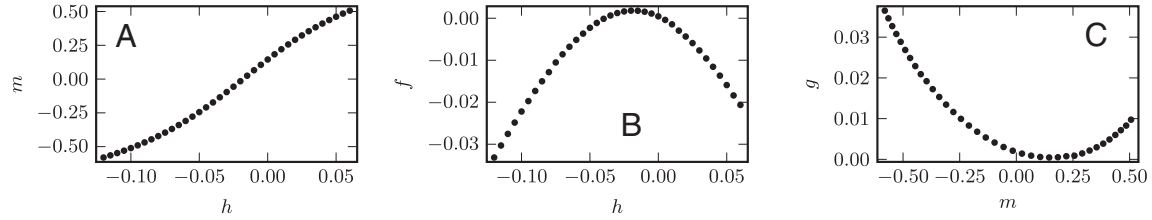

Fig. 5. Large deviation study of Glauber dynamics on an infinite 3-Random Regular Graph. Free dynamics with $\beta J = 0.6$, $T = 10$, magnetization at time zero $m^0 = 0.1$, zero external field, reweighted with an external field at final time $\prod_i \phi_i^T(\sigma_i^T) = \prod_i e^{h\sigma_i^T}$. (*A*) Magnetization vs. reweighting field, (*B*) Bethe Free Energy vs. reweighting field, (*C*) Magnetization-constrained free energy $g(m)$ vs. magnetization. Bond dimension 25.

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

Discussion

It is often the case that stochastic processes which can be described accurately, be it by analytical or Monte-Carlo methods, become computationally difficult as soon as the dynamics is biased by some reweighting factor. This constitutes a massive limitation since reweighting is essential whenever one is interested in describing atypical trajectories, an emblematic example being inference in epidemic models. As of today, there exists, to the best of our knowledge, no analytic method able to describe reweighted complex dynamics on networks except for the simple case of nonrecurrent models. In this article we adopted the matrix-product parametrization, inspired by techniques used originally in quantum physics and recently applied to classical stochastic dynamics in ref. 9, to devise the Matrix Product Belief Propagation method. We used it to describe reweighted Markov dynamics on graphs and applied it to epidemic spreading and a dynamical Ising models. With respect to the important work in refs. 9 and 10, which we recall that applies only to free dynamics, our contribution is twofold.

First, we develop for MPBP a general scheme to render the computation time linear in the node degree rather than exponential on a wide class of models, allowing us to compare it extremely favorably with existing methods on standard benchmark examples (which typically include vertices with large degrees). The bottleneck of the whole computation in the final scheme is due to the SVD factorization, which are cubic in the bond dimension M: larger matrices give a better approximation, but require a greater computational effort. The overall cost per iteration, assuming the bond dimension constant, is $O(T|E|)$, i.e.n linear in the number of edges of the graph. A small number of iterations is normally sufficient for approximate convergence to a fixed point. A strategy we found to be effective is to start with bond dimension M very small, say 4 or 5, iterate to convergence, then repeat with increasingly larger M until the increment in accuracy is small enough or the computation time becomes too large. In any case, the final value for M is specified in all figures and is relatively modest. It is fair to point out, however, that although linear, depending on the target accuracy of the approximation controlled by the parameter M, the method may be substantially more computationally intensive than the others used for comparison.

Second and more importantly, the MPBP approach allows to include reweighting factors. In particular, the approach proposed in refs. 9 and 10 is iterated forward in (dynamical) time and thus allows no backward flow of information which is necessary with reweighting factors. Reweighting factors are necessary to analyze conditioned dynamics and rare events.

MPBP, like many other statistical physics-inspired approaches to stochastic dynamics, is based on the cavity approximation. The Belief Propagation formalism gives access to the thermodynamic limit for certain ensembles of random graphs, provides an approximation to the partition function through the Bethe Free Energy, and allows to compute time autocorrelations. The limits of validity of MPBP are inherited from those of the cavity approximation: using the jargon of disordered systems, the approximation is accurate as long as the problem is in a *Replica Symmetric* (RS) phase. In the case of epidemic inference presented in Fig. 3 this is surely the case, since the trajectory to be inferred was sampled from the same prior used for the inference. This amounts to working on the Nishimori line, where it is known that no replica symmetry-breaking takes place (31). A study of the performance in regimes where replica symmetry is broken is left for future investigation.

On graphs with short loops, the performance of BP degrades substantially. In the static case, this issue can sometimes be overcome by resorting to higher order approximations (32). We argue that the same ideas can be translated to dynamics, for example by describing explicitly the joint trajectory of quadruples of neighboring variables on a square lattice.

Software implementing the method is available at ref. 19 and can be used to directly reproduce the results in the article. The framework is flexible and accommodates for the inclusion of new models of dynamics.

As a final remark, we recall that the method applies more in general to any distribution of the type [**1**], where t need not be interpreted as a time index but could, for instance, span a further spatial direction. Investigation along this line is left for future work.

Materials and Methods

As anticipated, messages are parametrized in terms of matrix products

$$m_{i\to j}(\bar{x}_i, \bar{x}_j) \propto \prod_{t=0}^{T} A^t_{i\to j}(x^t_i, x^t_j), \qquad [13]$$

where, for any (x^t_i, x^t_j), $A^t_{i\to j}(x^t_i, x^t_j)$ is a real-valued matrix. We set A^0 to have one row and A^T to have one column, so that the whole product gives a scalar. Plugging the *ansatz* [**13**] into the RHS of the BP equation [**7**] gives

$$m_{i\to j}(\bar{x}_i, \bar{x}_j) \propto \prod_{t=0}^{T} B^t_{i\to j}(x^{t+1}_i, x^t_i, x^t_j), \qquad [14]$$

with

$$B^t_{i\to j}(x^{t+1}_i, x^t_i, x^t_j) = \sum_{\{x^t_k\}_{k\in\partial i\setminus j}} f^{t+1}_i(x^{t+1}_i, \boldsymbol{x}^t_{\partial i}, x^t_i) \times \left[\bigotimes_{k\in\partial i\setminus j} A^t_{k\to i}(x^t_k, x^t_i)\right]. \qquad [15]$$

Two steps are missing in order to close the BP equations under a matrix product *ansatz*, as discussed in ref. 9. First, matrices must be recast into the form [**13**]. Second, if incoming A matrices have bond dimension M, B matrices will have bond dimension $M^{|\partial i|-1}$ and thus will keep growing indefinitely throughout the iterations. Both issues are solved by means of two successive sweeps of Singular Value Decompositions (SVD). SVD decomposes a real-valued matrix A as $A_{ij} = \sum_{k,l=1}^{M} U_{ik}\Lambda_{kl}V_{jl}$ where $\Lambda_{kl} = \lambda_k\delta_{k,l}$ is the diagonal matrix of singular values $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$ and $U^\dagger U = VV^\dagger = \mathbb{1}$ (we use the dagger symbol for matrix transpose to avoid confusion with the time labels t, T, but all matrices are real-valued). By retaining only the largest M' singular values and setting the others to zero, one can approximate A_{ij} with $\widetilde{A}_{ij} := \sum_{k=1}^{M'} U_{ik}\lambda_k V_{jk}$ making an error $\|A - \widetilde{A}\|_F^2 = \sum_{ij}\left(A_{ij} - \widetilde{A}_{ij}\right)^2 = \sum_{k=M'+1}^{M}\lambda_k^2$. As a result, both U and V are smaller in size.

The first sweep is done from left to right $t = 0, 1, 2, \ldots, T-1$ by performing an SVD decomposition

$$B^t_{i\to j}(x_i^{t+1}, x_i^t, x_j^t) = C^t_{i\to j}(x_i^t, x_j^t)\Lambda^t\left[V^t(x_i^{t+1})\right]^\dagger, \qquad [16]$$

then redefine $B^{t+1}_{i\to j}(x_i^{t+2}, x_i^{t+1}, x_j^{t+1})$ as $\Lambda^t\left[V^t(x_i^{t+1})\right]^\dagger B^{t+1}_{i\to j}(x_i^{t+2}, x_i^{t+1}, x_j^{t+1})$. The decomposition in Eq. **16** is performed by incorporating x_i^t, x_j^t as row indices and x_i^{t+1} as column index (see the *SI Appendix* for more details). At the end of this first sweep, the message looks like

$$m_{i\to j}(\bar{x}_i, \bar{x}_j) = \prod_{t=0}^{T} C^t_{i\to j}(x_i^t, x_j^t), \qquad [17]$$

where, thanks to the properties of the SVD, it holds that

$$\sum_{x_i^t x_j^t}\left[C^t_{i\to j}(x_i^t, x_j^t)\right]^\dagger C^t_{i\to j}(x_i^t, x_j^t) = \mathbb{1}. \qquad [18]$$

At this point the form [**13**] is recovered: The BP equations are closed under a matrix product *ansatz*. All that is left to do is perform a second sweep of SVD, this time discarding the smallest singular values to obtain matrices of reduced size. Going right to left $t = T, T-1, \ldots, 1$, incorporating (x_i^t, x_j^t) as column indices:

$$\begin{aligned} C^t_{i\to j}(x_i^t, x_j^t) &\overset{\text{SVD,trunc}}{=:} U^t\Lambda^t A^t_{i\to j}(x_i^t, x_j^t) \\ C^{t-1}_{i\to j}(x_i^{t-1}, x_j^{t-1}) &\leftarrow C^{t-1}_{i\to j}(x_i^{t-1}, x_j^{t-1})U^t\Lambda^t. \end{aligned} \qquad [19]$$

The errors made during the truncations are controlled: consider a generic step t in the sweep from right to left. The MPS is in the so-called mixed-canonical form (33):

$$C^0\cdots C^t A^{t+1}\cdots A^T, \qquad [20]$$

with $C^0\cdots C^{t-1}$ left-orthogonal ($C^\dagger C = \mathbb{1}$) and $A^{t+1}\cdots A^T$ right-orthogonal ($AA^\dagger = \mathbb{1}$). C^t is neither.

Canonical forms are a useful tool to perform controlled truncations (28, 33). The error in replacing C^t by $\widetilde{C}^t$ which retains only M' of the M singular values is

$$\begin{aligned} &\|C^0\cdots C^t A^{t+1}\cdots A^T - C^0\cdots\widetilde{C}^t A^{t+1}\cdots A^T\|_F^2 \\ &= \|C^t - \widetilde{C}^t\|_F^2 = \sum_{k=M'+1}^{M}\lambda_k^2, \end{aligned} \qquad [21]$$

where the first equality holds thanks to the orthonormality of C and A matrices. Keeping the MPS in canonical form ensures that the global error on the matrix product reduces to the local error on C^t.

As a side remark, we point out that there exist techniques to compute directly the SVD truncated to the M' largest singular values (34, 35). Such strategies can be advantageous for large M and small M'.

The results in this work were obtained by fixing the number of retained singular values, and hence the bond dimension. Alternatively, given a target threshold ε, one can select M' adaptively such that, e.g., $\frac{\lambda_{M'}}{\sqrt{\sum_k\lambda_k^2}} < \varepsilon$, as in ref. 9. We find the approach with fixed bond dimension better suited for an iterative solver such as BP, where messages are computed and then overwritten many times before convergence is reached. During the first iterations, a coarse approximation with small bond dimension is sufficient and helps to keep the computation time under control. Then, as messages approach a fixed point, one can refine the estimate by either increasing the bond dimension or switching to a threshold-based truncation method.

Bond Dimension. Issues may arise whenever excessive truncations turn the matrix product into an ill-defined probability distribution taking negative values. This is to be expected and indeed was encountered in the experiments we run. Rerunning BP with larger bond dimension invariably solved the problem. Fig. 6 shows the effect of varying the bond dimension in two of the settings shown in the previous plots. Instead, truncating too much may lead to unreasonable results such as negative probability values.

Turning to the expressive power of the MPS *ansatz*, it is reasonable to expect that truncating conservatively, i.e., allowing large bond dimension, will lead to better and better approximations. Indeed, matrix products with arbitrarily large bond dimension can represent exactly any distribution. However, it is hard to make quantitative statements about the relationship between bond dimension and the complexity that can be captured. Based on the discussion in the context of quantum mechanics (see, e.g., ref. 33, section 4.2.2), it is plausible to assume that strong and long-range (here in time, in the quantum context these are usually in space) correlations need large matrices to be captured accurately. However, this cannot possibly be the whole story, since there exists a simple counterexample: any trajectory of the SI epidemic model can be represented using MPS of finite bond dimension despite featuring infinite-range correlations. More details are found in the *SI Appendix*.

Convergence. The BP equations are iterated until convergence to a fixed point. We opted for an asynchronous update scheme because it tends to feature better convergence properties with respect to a synchronous one. Nevertheless, the two can be used interchangeably. As usual with BP, the procedure naturally lends itself to parallelization, to a larger extent with the synchronous approach.

As a criterion for convergence to a fixed point, we computed the marginal distributions at all nodes and epochs $b_i^t(x_i^t)$ (Eq. **23**) and checked whether, for an iteration *it* and the successive one,

$$\max_{i\in\{1,\ldots,N\}}\max_{t\in\{0,\ldots,T\}}\max_{x_i^t}\left|\left[b_i^t(x_i^t)\right]^{(it+1)} - \left[b_i^t(x_i^t)\right]^{(it)}\right| < \varepsilon, \qquad [22]$$

for some small threshold ε. A stricter criterion can be considered by computing $\max_{(i,j)\in E}\sum_{\bar{x}_i,\bar{x}_j}\|m_{i\to j}(\bar{x}_i,\bar{x}_j)^{(it+1)} - m_{i\to j}(\bar{x}_i,\bar{x}_j)^{(it)}\|_F$. The two criteria lead to similar outcomes [results not shown, see implementation (19)].

It is worth noting that in the case of free dynamics, one can build the messages incrementally from time 0 to time T as in DC (see, e.g., ref. 9), with no need to iterate until convergence. Because each sweep of SVD over t matrices takes linear time in t, the total computational cost when using such scheme scales quadratically with T. Instead, initializing messages for all T epochs and then doing N_{iter} iterations as in our method takes $\mathcal{O}(N_{iter}T)$. The two are essentially equivalent since we observed that typically the number of iterations needed to converge is of the order of T.

It is worth noting that, up to the errors introduced by the truncations, which we showed to be controlled, MPBP is exact on acyclic graphs.

Observables. On a fixed point of the BP equations, single-node marginal distributions, "beliefs," are given by

$$b_i(\bar{x}_i) \propto \sum_{\bar{x}_{\partial i}}\prod_{t=0}^{T-1} f_i^{t+1}(x_i^{t+1}, \boldsymbol{x}_{\partial i}^t, x_i^t)\times\prod_{k\in\partial i} m_{k\to i}(\bar{x}_k, \bar{x}_i). \qquad [23]$$

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

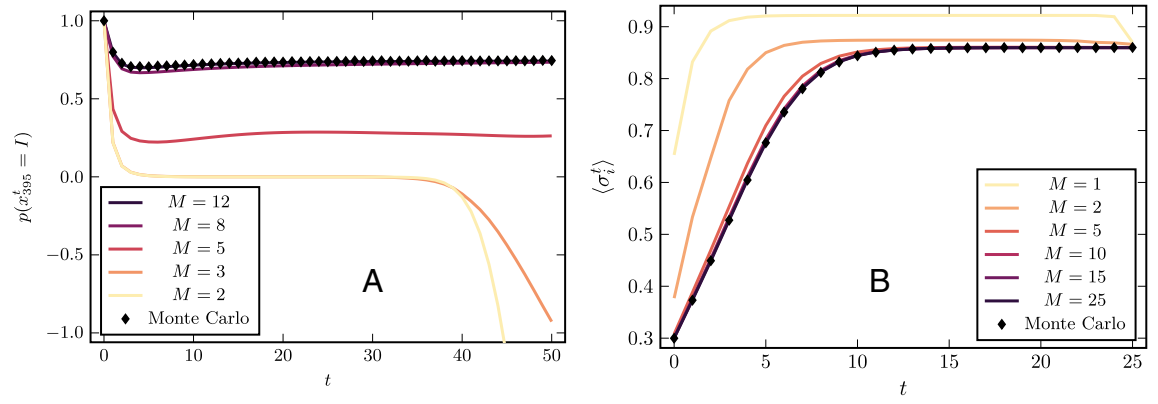

Fig. 6. Effect of varying the bond dimension M on the accuracy of the approximation. (*A*) SIS model on a tree, the same settings as Fig. 1*A*. Too small bond dimension gives unreasonable results. (*B*) Glauber dynamics on an infinite random regular graph of degree 8, same settings as Fig. 4*A*.

Single-variable and pair marginal distributions as well as time autocorrelations can be computed efficiently on a fixed point of BP by means of standard tensor network contraction techniques (for details, see *SI Appendix* or ref. 28). The BP formalism also gives access to the Bethe Free Energy, an approximation to (minus the logarithm of) the normalization of Eq. **2**, which can be interpreted as the likelihood of the parameters of the dynamics (e.g., infection rates, temperature,...). In cases where such parameters are unknown, they can be learned via a maximum-likelihood procedure.

Thermodynamic Limit. Just like standard BP, MPBP lends itself to be extended to infinite graphs. In the case of random regular graphs with homogeneous properties (e.g., $\lambda_{ij} \equiv \lambda, \rho_i \equiv \rho$ for epidemic models, $J_{ij} \equiv J, h_i \equiv h$ for Glauber dynamics), a single message is sufficient to represent the distribution in the thermodynamic limit. For graph ensembles with variable degree and/or parameters distributed according to some disorder, we adopt a population dynamics approach (more details in *SI Appendix*).

A Family of Models with Linear Computational Cost. As mentioned before, in the scheme proposed in ref. 9, matrices before truncation have size M^{z-1} where M is the size of matrices in the incoming messages and z is the degree. The bottleneck is the sweeps of SVDs which yield a computational cost $\mathcal{O}(M^{3z-3})$ for a single BP message. Although in a later work [(10), section 6] it was shown that such cost can be reduced to $\mathcal{O}(M^{2z-1})$, the exponential dependence on the degree still represents an issue even for graphs of moderately large connectivity. Here we show an improved scheme that, for a wide class of models including many in epidemics and kinetic Ising, performs the computation in $\mathcal{O}(M^6)$. The dependence on z is only polynomial and depends on the details of the model.

It is enough to notice that in many cases transition probabilities $w(x_i^{t+1}|\boldsymbol{x}^t_{\partial i}, x_i^t)$ depend on $\boldsymbol{x}^t_{\partial i}$ only through some intermediate variable which incorporates the aggregate interaction with all the neighbors. In epidemic models like SI, SIR, and SIRS, the transition probability only depends on the event that at least one of the neighbors has infected node i. In the case of kinetic Ising the transition probability only depends on the local field, which is a weighted sum of neighboring spins.

More formally, consider intermediate scalar variables y^t_A with $A \subseteq \partial i$ encoding information about $\boldsymbol{x}^t_A$. By definition of conditional probability

$$p\left(x_i^{t+1}|\boldsymbol{x}^t_{\partial i}, x_i^t\right) = \sum_{y_{\partial i}} p\left(x_i^{t+1}|y^t_{\partial i}, x_i^t\right) p\left(y^t_{\partial i}|\boldsymbol{x}^t_{\partial i}, x_i^t\right). \quad [24]$$

If it holds that

$$p\left(y^t_{A\cup B}|x^t_{A\cup B}, x_i^t\right) = \sum_{y_A, y_B} p\left(y^t_{A\cup B}|y^t_A, y^t_B, x_i^t\right) \times p\left(y^t_A, y^t_B|x^t_{\partial i}, x_i^t\right), \quad [25]$$

$$= \sum_{y_A, y_B} p\left(y^t_{A\cup B}|y^t_A, y^t_B, x_i^t\right) \times p\left(y^t_A|x^t_A, x_i^t\right) p\left(y^t_B|x^t_B, x_i^t\right), \quad [26]$$

for $A \cup B \subseteq \partial i$ (i.e., that the y of disjoint index sets are independent given the x's), then it suffices to provide:

1. $p\left(y^t_j|x^t_j, x^t_i\right)$
2. $p\left(y^t_{A\cup B}|y^t_A, y^t_B, x^t_i\right)$

to be able to compute the set of outgoing messages from a node in a recursive manner. This is more efficient than the naive implementation provided that the number of values that each y can assume does not grow exponentially with the number of x's it incorporates. More details of the computation can be found in *SI Appendix*.

Data, Materials, and Software Availability. Source code data have been deposited in https://github.com/stecrotti/MatrixProductBP.jl (19).

ACKNOWLEDGMENTS. This study was carried out within the Future Artificial Intelligence Research and received funding from the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza–Missione 4 Componente 2, Investimento 1.3–D.D. 1555 11/10/2022, PE00000013). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

1. P. Van Mieghem, J. Omic, R. Kooij, Virus spread in networks. *IEEE/ACM Trans. Networking* **17**, 1–14 (2008).
2. B. Karrer, M. E. J. Newman, Message passing approach for general epidemic models. *Phys. Rev. E* **82**, 016101 (2010).
3. G. Del Ferraro, E. Aurell, Dynamic message-passing approach for kinetic spin models with reversible dynamics. *Phys. Rev. E* **92**, 010102 (2015).
4. A. Pelizzola, Variational approximations for stationary states of Ising-like models. *Eur. Phys. J. B* **86**, 1–8 (2013).
5. A. Pelizzola, M. Pretti, Variational approximations for stochastic dynamics on graphs. *J. Stat. Mech: Theory Exp.* **2017**, 073406 (2017).
6. E. Aurell, G. Del Ferraro, E. Domínguez, R. Mulet, Cavity master equation for the continuous time dynamics of discrete-spin models. *Phys. Rev. E* **95**, 052119 (2017).
7. E. Ortega, D. Machado, A. Lage-Castellanos, Dynamics of epidemics from cavity master equations: Susceptible-infectious-susceptible models. *Phys. Rev. E* **105**, 024308 (2022).
8. I. Neri, Désiré. Bollé,, The cavity approach to parallel dynamics of Ising spins on a graph. *J. Stat. Mech: Theory Exp.* **2009**, P08009 (2009).
9. T. Barthel, C. De Bacco, S. Franz, Matrix product algorithm for stochastic dynamics on networks applied to nonequilibrium Glauber dynamics. *Phys. Rev. E* **97**, 010104 (2018).
10. T. Barthel, The matrix product approximation for the dynamic cavity method. *J. Stat. Mech.: Theory Exp.* **2020**, 013217 (2020).

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.

11. D. Perez-Garcia, F. Verstraete, M. M. Wolf, J. I. Cirac, Matrix product state representations. *Quantum Inf. Comput.* **7**, 401–430 (2007).
12. J. Frank Verstraete, C. Ignacio, Matrix product states represent ground states faithfully. *Phys. Rev. B* **73**, 094423 (2006).
13. M. Fannes, B. Nachtergaele, R. F. Werner, Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144**, 443–490 (1992).
14. B. Derrida, M. R. Evans, V. Hakim, V. Pasquier, Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A: Math. Gen.* **26**, 1493 (1993).
15. M. C. Bañuls, J. P. Garrahan, Using matrix product states to study the dynamical large deviations of kinetically constrained models. *Phys. Rev. Lett.* **123**, 200601 (2019).
16. Z.-Y. Han, J. Wang, H. Fan, L. Wang, P. Zhang, Unsupervised generative modeling using matrix product states. *Phys. Rev. X* **8**, 031012 (2018).
17. E. Stoudenmire, D. J. Schwab, Supervised learning with tensor networks. *Adv. Neural. Inf. Process. Syst.* **29** (2016).
18. F. Altarelli, A. Braunstein, L. Dall'Asta, R. Zecchina, Large deviations of cascade processes on graphs. *Phys. Rev. E* **87**, 062115 (2013).
19. S. Crotti, A. Braun-stein, MatrixProductBP.jl. GitHub. https://github.com/stecrotti/MatrixProductBP.jl. Accessed 27 October 2023.
20. A. Renart *et al*., The asynchronous state in cortical circuits. *Science* **327**, 587–590 (2010).
21. Y. Roudi, J. Hertz, Mean field theory for nonequilibrium network reconstruction. *Phys. Rev. Lett.* **106**, 048702 (2011).
22. H. Ohta, S. Sasa, A universal form of slow dynamics in zero-temperature random-field Ising model. *Europhys. Lett.* **90**, 27008 (2010).
23. P. Van Mieghem, The n-intertwined SIS epidemic network model. *Computing* **93**, 147–169 (2011).
24. M. Shrestha, S. V. Scarpino, C. Moore, Message-passing approach for recurrent-state epidemic models on networks. *Phys. Rev. E* **92**, 022821 (2015).
25. E. D. Vázquez, G. Del Ferraro, F. Ricci-Tersenghi, A simple analytical description of the non-stationary dynamics in Ising spin systems. *J. Stat. Mech: Theory Exp.* **2017**, 033303 (2017).
26. G. Del Ferraro, E. Aurell, Perturbative large deviation analysis of non-equilibrium dynamics. *J. Phys. Soc. Jpn.* **83**, 084001 (2014).
27. N. Antulov-Fantulin, A. Lančić, T. Šmuc, H. Štefančić, M. Šikić, Identification of patient zero in static and temporal networks: Robustness and limitations. *Phys. Rev. Lett.* **114**, 248701 (2015).
28. I. V. Oseledets, Tensor-train decomposition. *SIAM J. Sci. Comput.* **33**, 2295–2317 (2011).
29. J. Kunegis. Zachary Karate Club (2023). http://konect.cc/networks/ucidata-zachary/. Accessed 14 January 2023.
30. A. Coja-Oghlan, P. Loick, B. F. Mezei, G. B. Sorkin, The Ising antiferromagnet and max cut on random regular graphs. *SIAM J. Discret. Math.* **36**, 1306–1342 (2022).
31. Y. Iba, The Nishimori line and Bayesian statistics. *J. Phys. A: Math. Gen.* **32**, 3875 (1999).
32. J. S. Yedidia, W. Freeman, Y. Weiss, Generalized belief propagation. *Adv. Neural. Inf. Process. Syst.* **13**, 1–7 (2000).
33. U. Schollwöck, The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96–192 (2011).
34. R. M. Larsen, "Lanczos bidiagonalization with partial reorthogonalization" (DAIMI Report Series No. 537, 1998).
35. J. Baglama, L. Reichel, Augmented implicitly restarted Lanczos bidiagonalization methods. *SIAM J. Sci. Comput.* **27**, 19–42 (2005).

Downloaded from https://www.pnas.org by SISTEMA BIBL POLITECN TORINO BIBL CENTRALE INGEGNERIA on January 22, 2024 from IP address 130.192.10.104.