

## Efficient simulations of epidemic models with tensor networks: Application to the one-dimensional susceptible-infected-susceptible model

Wout Merbis<sup>✉,\*</sup>, Clélia de Mulatier, and Philippe Corboz

*Dutch Institute for Emergent Phenomena (DIEP) & Institute for Theoretical Physics (ITFA), University of Amsterdam, 1090 GL Amsterdam, The Netherlands*



(Received 12 April 2023; accepted 20 July 2023; published 8 August 2023)

The contact process is an emblematic model of a nonequilibrium system, containing a phase transition between inactive and active dynamical regimes. In the epidemiological context, the model is known as the susceptible-infected-susceptible model, and it is widely used to describe contagious spreading. In this work, we demonstrate how accurate and efficient representations of the full probability distribution over all configurations of the contact process on a one-dimensional chain can be obtained by means of matrix product states (MPSs). We modify and adapt MPS methods from many-body quantum systems to study the classical distributions of the driven contact process at late times. We give accurate and efficient results for the distribution of large gaps, and illustrate the advantage of our methods over Monte Carlo simulations. Furthermore, we study the large deviation statistics of the dynamical activity, defined as the total number of configuration changes along a trajectory, and investigate quantum-inspired entropic measures, based on the second Rényi entropy.

DOI: [10.1103/PhysRevE.108.024303](https://doi.org/10.1103/PhysRevE.108.024303)

### I. INTRODUCTION

Accurate and efficient mathematical modeling of complex systems, composed of many interacting constituents, poses a formidable challenge in many biophysical and socioecological systems [1–6]. Due to the random nature of biochemical, social, or ecological interactions, models are often defined stochastically in terms of discrete-state Markov chains, describing the continuous-time evolution of a probability vector over the set of all possible system states [7–10]. As the number of system states grows exponentially with system size, exact methods quickly become intractable and one is often forced to rely on Monte Carlo simulation [11] or mean-field approximations [12–14]. In the mean-field approximation, the focus is often to compute average quantities, disregarding variances around these quantities. This can lead to problems in the context in which systems are non-self-averaging, which is often the case for processes with reproduction and annihilation, such as branching processes in epidemic modeling, population dynamics [15,16], or nuclear reactor physics [17]. A concrete and urgent application where this is relevant concerns the modeling and forecasting of contagion spreading [10,18–21], imperative in light of the recent COVID-19 pandemic.

The problem of finding accurate and efficient descriptions of high-dimensional vectors is a familiar one in the study of emergent collective behavior in composite quantum systems. Many-body quantum systems are characterized by an exponentially large Hilbert space, although this is often merely a facade [22]. The subset of physically relevant states is often much smaller (and scaling polynomially in

system size), allowing for efficient representation using tensor networks (TNs) [23–25]. TNs exploit an internal structure of high-dimensional composite wave functions by representing them as a set of lower-dimensional tensors, contracted along internal (bond) indices. The bond dimension of the tensor network represents the size of the effective state space of the system, which is often much smaller than the full Hilbert space. Efficient representations, where the bond dimensions scale at most polynomially with system size, are usually possible when interactions between the constituents are local.

In this regard, stochastic models of composite complex systems display similarities with quantum many-body systems [26–33]. The high-dimensional probability vector is an element of the tensor product space of the probability vectors of the constituents and interactions are defined locally [10]. A natural question is whether and when these high-dimensional probability vectors can be efficiently represented using TNs. In other words, is it possible to use TNs to find an optimal compression of large-dimensional probability distributions, while remaining as close as possible to the original distribution?

TN methods are known to be effective for stochastic kinetic models such as the totally asymmetric simple exclusion process (TASEP) [34–38]. Methods based on the density matrix renormalization group (DMRG) have been applied to stochastic out-of-equilibrium systems in Refs. [39–42]. More recently, methods based on projected entangled pair states (PEPSs) have been applied to stochastic models of nonequilibrium systems [43,44] and matrix product states (MPSs) have been used for studying large deviation statistics of dynamical stochastic systems [45–49]. The majority of work in this direction is, however, based on stochastic systems satisfying local detailed balance, an assumption that does not necessarily

\*w.merbis@uva.nl

hold for all relevant biophysical systems, and certainly not for most epidemic models. In this paper, we investigate the use of TNs for epidemiological modeling by constructing an MPS representation for the one-dimensional (1D) contact process [50–53], known in the epidemiological literature as the susceptible-infected-susceptible (SIS) model [18,21]. The contact process contains sites that may be occupied (infected) or empty (susceptible), and it allows transitions from the empty to the occupied state only if a neighboring site is occupied (infection). In turn, occupied sites may transition to empty regardless of their neighbor’s state, modeling the recovery of individuals. The process has been called “the Ising model of absorbing state transitions,” and it serves as a natural starting point for developing new methods for nonequilibrium problems [54]. The special aspect of this work lies in the fact that the contact process does not satisfy local detailed balance, such that there is no map between the Markov generator and a Hermitian matrix. Hence, known MPS algorithms used for Hermitian time-evolution have to be adapted for the purely stochastic evolution of the contact process, with applications to more generic biophysical and socioecological models.

Successful application of TNs to stochastic models provides an alternative simulation framework, complementary to Monte Carlo (MC) sampling methods, frequently used for analyzing nonequilibrium stochastic systems. The power of TNs lies in providing an efficient description for the  $2^N$ -dimensional probability vector of the model, by finding a lower-dimensional representation that keeps only the relevant physical states. Instead of computing empirical averages over a large number of numerically generated trajectories, the tensor network is capable of directly providing accurate information on observables averaged over the ensemble of all dynamical trajectories. This makes it possible to efficiently compute marginals, conditional probabilities, expectation values, and higher moments for any macroscopic observable one might be interested in, even when variances are large [55]. There is hence great potential for TNs as a modeling and forecasting tool, which will be able to provide more detailed information on correlations and variances than conventional methods. To illustrate the complementary nature of tensor networks and Monte Carlo methods, we showcase three different computations where tensor networks offer advantages over Monte Carlo simulations. These computations involve the efficient calculation of large gaps in the occupancy of the chain, the computation of the scaled cumulant-generating function, and the computation of the Rényi entropy and Rényi mutual information.

First, computing large gaps in the occupancy of the chain constitutes an analysis of extremely rare events in both the absorbing and active phases of the dynamical process. These events are challenging to obtain accurately using Monte Carlo simulations due to their exponentially vanishing probability. However, tensor networks, particularly the MPS representation, provide an efficient means of computing such rare events. While MC methods become increasingly difficult for larger gap sizes, the efficiency of the MPS computation hardly depends on gap size, and hence TNs are capable of providing accurate information on the distribution of rare large gap sizes.

Second, we study the thermodynamics of dynamical trajectories [56–58] and the large deviation theory [59] for the dynamical activity in the 1D contact process using the MPS. The moment-generating function for the dynamical activity (defined as the number of configuration changes per unit time) with dual parameter  $s$  can be obtained as the norm of a tilted probability vector, which evolves in time by a tilted generator [42,56]. This is obtained from the Markov generator by exponentially tilting the off-diagonal terms by  $e^s$ , such that for  $s > 0$  active trajectories become more likely, while  $s < 0$  tilts the dynamics towards inactive trajectories. The scaled cumulant-generating function (SCGF) for the dynamical activity at late times can then be obtained as the leading eigenvalue of the tilted generator [57], which we obtain using variational MPS methods. We find that this leads to a clear kink in the SCGF, indicating a discontinuous phase transition between the active and inactive phases as a function of control parameter  $s$ , in accordance with previous studies using DMRG methods [42]. Furthermore, through the large deviation principle, the SCGF relates to “rare trajectories” which contain activity deviating from the average. Although rare trajectories are generated by the tilted operator, this is not a stochastic operator, and hence these rare trajectories cannot be sampled directly [60].

Lastly, the computation of the Rényi entropy and Rényi mutual information (RMI) [61–63] in the context of the one-dimensional contact process is a task in which tensor networks offer a significant advantage. Classical information measures, such as the full Shannon entropy and mutual information, are computationally demanding to compute accurately over the space of all state configurations, because undersampling can lead to biased results, and convergence to the true values is slow [64,65]. While these measures remain hard to compute using TNs, the MPS representation does enable efficient computation of the second Rényi entropy and related quantities. Here we compute Rényi entropies and RMI for the one-dimensional contact process, and we find that, in accordance with phase transitions in equilibrium models [66], the RMI can be used to detect the absorbing state transition. This demonstrates that the RMI also provides a way to detect nonequilibrium phase transitions independent of a choice of order parameter.

In the following sections, we present the model and discuss different driving protocols to remove the absorbing state, leading the system to a nonequilibrium steady state (NESS) at late times (Sec. II). We introduce the MPS representation used to approximate the complete  $2^N$ -dimensional probability distribution of the chain (Sec. III). The thermodynamics of trajectories and the large deviation principle are reviewed in Sec. IV. We provide details on the variational MPS method, its accuracy, and its efficiency, and we construct the MPS representation of the NESS to study the distribution of gaps in the chain in Sec. V. Results on the scaled cumulant-generating function for the dynamical activity are presented in Sec. VI. Section VII focuses on entropic measures based on Rényi entropies obtained from the MPS representation. Finally, in the discussion section (Sec. VIII), we summarize our findings and explore potential future directions for utilizing tensor networks in the context of complex system modeling. The

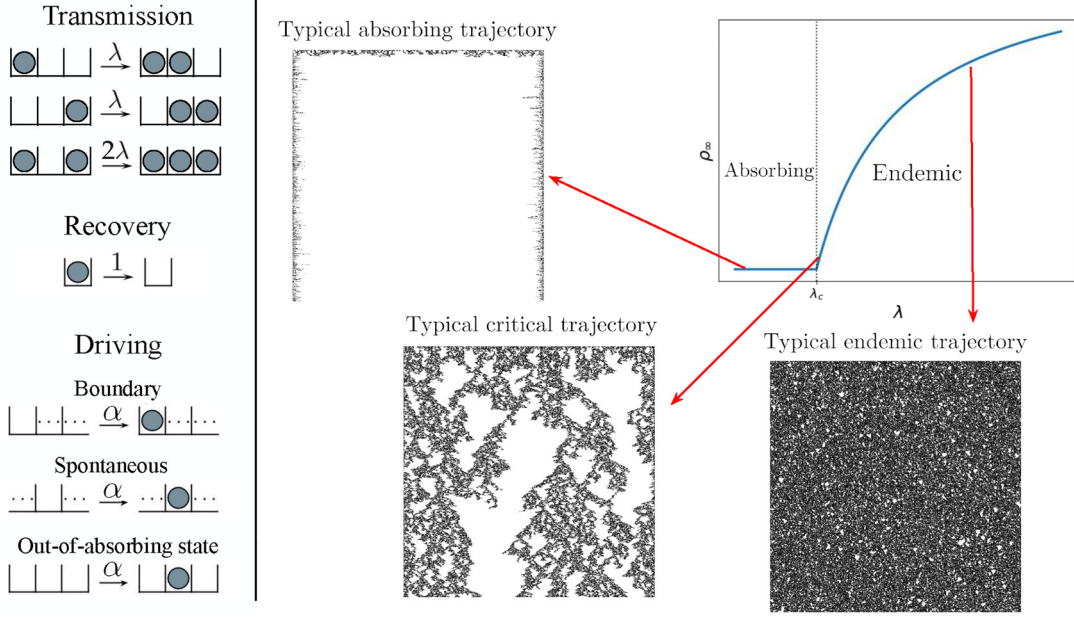


FIG. 1. Left: The possible transitions for the contact process are transmission (each infected site can infect a susceptible neighbor with rate  $\lambda$ ) and recovery (infected sites become susceptible with rate 1). Additionally, to maintain the process out of equilibrium, we consider three possible driving terms: the spontaneous infection of a site either at the boundaries or anywhere on the lattice, or the spontaneous infection of a site only when the absorbing state has been reached. Right: The infinite chain contains a continuous phase transition between an absorbing phase  $\lambda > \lambda_c$  and an endemic phase  $\lambda < \lambda_c$ . Typical (finite-size) trajectories are shown for the boundary-driven system in the absorbing, critical, and endemic regimes, where the horizontal direction corresponds to the length of the chain, and the vertical direction is time.

software developed to perform all computations described in this paper is publicly available in [67].

## II. THE MODEL

The contact process, or SIS model, on a one-dimensional lattice is defined in terms of an array of  $N$  binary variables  $(n_1, \dots, n_N)$ , where  $n_i \in \{0, 1\}$  indicates the occupancy of the lattice site  $i$ . Occupied sites  $n_i = 1$  are considered infected, while empty sites  $n_i = 0$  are susceptible. Infected sites can transition back to being susceptible following a Poisson process with recovery rate  $\gamma$  (see Fig. 1). Infected sites can also infect their direct neighbors with a transmission rate  $\beta$ . As a result, susceptible sites become infected with a rate  $k\beta$ , where  $k = 0, 1$ , or  $2$  is the number of its infected neighbors.

After rescaling time  $t \rightarrow \gamma t$ , the only relevant parameter in the system is the dimensionless ratio:  $\lambda = \beta/\gamma$ . The probability vector over all configurations of the chain  $|P(t)\rangle$  satisfies a master equation  $\partial_t |P(t)\rangle = \hat{W} |P(t)\rangle$  with Markov generator  $\hat{W}$ , which can be written in a quantum notation [32,42,68] as

$$\hat{W} = \lambda \sum_{i=1}^{N-1} (\hat{n}_i \hat{w}_{i+1}^{0 \rightarrow 1} + \hat{w}_i^{0 \rightarrow 1} \hat{n}_{i+1}) + \sum_{i=1}^N \hat{w}_i^{1 \rightarrow 0} + \hat{W}_{\text{driv}}(\alpha), \quad (1)$$

where  $\hat{n}_i$ ,  $\hat{w}_i^{0 \rightarrow 1}$ ,  $\hat{w}_i^{1 \rightarrow 0}$  are local operators on site  $i$ . When represented in terms of a 2D vector space with basis states

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (2)$$

these operators are  $\hat{n} = |1\rangle\langle 1|$ ,  $\hat{w}^{0 \rightarrow 1} = |1\rangle\langle 0| - |0\rangle\langle 0|$ , and  $\hat{w}^{1 \rightarrow 0} = |0\rangle\langle 1| - |1\rangle\langle 1|$ , which corresponds to the matrices:

$$\hat{n} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad (3a)$$

$$\hat{w}^{0 \rightarrow 1} = \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}, \quad (3b)$$

$$\hat{w}^{1 \rightarrow 0} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}. \quad (3c)$$

The subindex  $i$  on the operators in (1) indicates that these operators act only on the local sites  $i$ . Explicit construction of the operators involves a tensor product over all sites, where only site  $i$  is acted upon by the appropriate matrix in (3), and all other sites are multiplied by the two-dimensional identity matrix.

In the absence of the driving term  $\hat{W}_{\text{driv}}(\alpha)$ , the Markov chain defined by  $\hat{W}$  in (1) is a much used example of a system with a nonequilibrium absorbing state transition (see Fig 1). In the limit of infinite system size,  $N \rightarrow \infty$ , there is a critical value  $\lambda_c$  of the control parameter  $\lambda$  below which the system relaxes to the absorbing “healthy” state in which no site is infected. Above the critical threshold  $\lambda > \lambda_c$ , the system goes to a stationary (endemic) state with a nonzero density of infected sites. The critical point of the infinitely long one-dimensional contact process is determined accurately in the literature to be  $\lambda_c = 1.648\,96$  (see [69] and references therein). As shown in Fig. 1, typical trajectories in the critical region display gaps in the density of all sizes.

For finite system sizes, the absorbing state is always reached, even above the critical point. It is therefore common

[42] to include an additional driving process by spontaneous creation of infected sites. There are several possibilities to include this driving term. For instance, one could introduce infected sites only at the boundary of the chain with rate  $\alpha$ . This is done by taking the driving term  $\hat{W}_{\text{driv}}$  to be

$$\hat{W}_{\text{driv}}^{\text{bdy}}(\alpha) = \alpha(\hat{w}_1^{0 \rightarrow 1} + \hat{w}_N^{0 \rightarrow 1}). \quad (4)$$

Alternatively, it is also possible to spontaneously occupy any empty node in the chain by including the driving term

$$\hat{W}_{\text{driv}}(\alpha) = \alpha \sum_{i=1}^N \hat{w}_i^{0 \rightarrow 1}, \quad (5)$$

where taking  $\alpha \sim \frac{1}{N^2}$  ensures the effect of the driving term will vanish in the limit  $N \rightarrow \infty$ . This choice has the advantage of being translation-invariant. A final possibility to remove the absorbing state is to occupy a random empty site only when the absorbing state has been reached, which is implemented by the driving term

$$\hat{W}_{\text{driv}}(\alpha) = \alpha \sum_{i=1}^n \prod_{\substack{j=1 \\ j \neq i}}^N \hat{v}_j \hat{w}_i^{0 \rightarrow 1}. \quad (6)$$

Here  $\hat{v} = |0\rangle\langle 0|$  is the number operator for empty sites.

The presence of a driving term removes the absorbing state and drives the system into a nonequilibrium steady state (NESS) at late times. It would therefore make more sense to talk about inactive and active regimes in the driven contact process, instead of the absorbing and endemic phases. No known analytical solutions exist for the NESS of the driven contact process, but it is possible to obtain accurate and efficient representations of it using variational matrix product state methods, as we will demonstrate in this paper.

### III. MPS REPRESENTATION OF THE MANY-BODY PROBABILITY VECTOR

The probability vector over all configuration  $|P(t)\rangle$  is a  $2^N$ -dimensional vector, with components equal to the probability  $P(\mathbf{x}, t)$  of being in a certain configuration  $\mathbf{x} \in \{0, 1\}^N$  at time  $t$ . This vector can be expanded in the product basis of  $N$  two-dimensional vectors  $|n_i\rangle$ :

$$|P(t)\rangle = \sum_{n_1, \dots, n_N} P(\mathbf{x} = n_1, \dots, n_N, t) |n_1, \dots, n_N\rangle, \quad (7)$$

with

$$|n_1, \dots, n_N\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \dots \otimes |n_N\rangle, \quad (8)$$

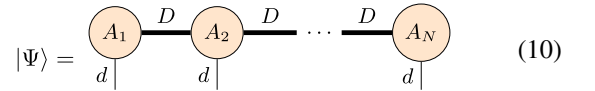
where  $\otimes$  denotes the tensor product. We wish to approximate the high-dimensional object  $|P(t)\rangle$  by a lower-dimensional MPS  $|\Psi\rangle$ , while retaining as much information on the state of the system as possible. An MPS is a vector composed out of an array of tensors  $A_i^{n_i}$ ,

$$|\Psi\rangle = \sum_{n_1, \dots, n_N \in \{0, 1\}} A_1^{n_1} A_2^{n_2} \dots A_N^{n_N} |n_1, \dots, n_N\rangle. \quad (9)$$

Each  $A_i$  is a  $(d \times D \times D)$ -dimensional rank-3 tensor, where  $d$  is the physical dimension of the variables  $n_i$  ( $d = 2$  in our case), and  $D$  denotes the *bond dimension* of the tensor. The

tensor  $A_i$  is multiplied to the left by the tensor  $A_{i-1}$  along  $D$  dimensions and to the right by  $A_{i+1}$  along its other  $D$  dimensions. The  $d = 2$  dimensions remaining correspond to the two states of the physical variable  $|n_i\rangle$ . In the case of open boundary conditions, as we will use here, the first and last tensors are rank-2 (matrices) of dimension  $2 \times D$ . The MPS hence has of the order of  $N \times 2D^2$  parameters (the total number of coefficients in the tensors), which can be much smaller than the original  $2^N$  parameters of  $|P(t)\rangle$ , depending on the choice of  $D$ . If  $D \sim 2^{N/2}$ , the system is represented with all  $2^N$  parameters, and no information is lost. The goal is to compress the representation of  $|P(t)\rangle$  by reducing the bond dimension  $D$  [and therefore the number of parameters used to represent  $|P(t)\rangle$ ], while retaining as much information as possible about the system.

Graphically, the MPS is often depicted as a network, where the nodes represent the tensors, and links are the indices. Two connected nodes are contracted over these indices. Equation (9) is graphically represented by



$$|\Psi\rangle = \begin{array}{c} \text{---} A_1 \text{---} D \text{---} A_2 \text{---} D \text{---} \dots \text{---} D \text{---} A_N \text{---} \\ | \quad d \quad | \quad d \quad | \quad \quad \quad | \quad d \quad | \end{array} \quad (10)$$

The open indices represent the two-dimensional physical basis of states  $n_i$  at each site. The thicker internal links are the  $D$ -dimensional bond indices, which are contracted over in Eq. (9).

For quantum states, the bond dimension  $D$  limits the amount of entanglement between sites in the chain, as the entanglement entropy  $S_A(L)$  of a connected subset  $A$  of  $L$  sites is bounded by  $S_A \leq 2 \log D$ . Ground states of gapped, local Hamiltonians in 1D are known to obey the area law of entanglement [70], implying that  $S_A$  does not scale with the volume of  $A$ , but only with the size of the boundary of  $A$ , which in 1D is a constant. For these cases, the MPS provides an efficient representation of the many-body wave function, even for infinite systems. In the case of critical states in 1D, there is a logarithmic correction to the area law,  $S_A(L) \sim \log L$ , and in this case  $D$  scales polynomially with system size. If  $D$  is taken of order  $2^{N/2}$ , any many-body vector can be represented as an MPS. The other limiting case of  $D = 1$  corresponds to the mean-field approximation, where no information about correlations in the chain is taken into account. Hence, by varying the bond dimension between 1 and  $2^{N/2}$ , one systematically goes beyond the mean-field approximation.

In this work, we will use the MPS ansatz (9) to find efficient approximations for many-body probability distributions over the ensemble of dynamical trajectories for the contact process. In that case, the bond dimension  $D$  can be thought of as the effective number of relevant configurations in the model, which might be much lower than the total number of configurations in the system. Any many-body vector can be put in MPS form by a sequence of singular value decompositions (SVD) [24], and  $D$  represents the number of singular values kept between the bonds after truncating the smaller singular values. The SVD constitutes a linear transformation on the basis states (in this case, the configurations of the chain), and by truncating the bonds to the largest singular values, only the most relevant linear combinations of basis states are kept.



#### IV. THERMODYNAMICS OF DYNAMICAL TRAJECTORIES AND LARGE DEVIATION STATISTICS

The time-evolving stochastic process traces out a *trajectory*  $\omega_t = \{\mathbf{x}_0(t_0), \mathbf{x}_1(t_1), \dots, \mathbf{x}_k(t_k)\}$  by hopping to configuration  $\mathbf{x}_l$  at hopping time  $t_l$ . To study the statistics of the trajectories in the system, we consider the dynamical activity  $K(\omega_t)$ , defined as the total number of configuration changes along a trajectory  $\omega_t$  of duration  $t$  [56,71]. The dynamical activity is an order parameter for the phase transition between active and inactive trajectories; this transition is dynamical in nature. Below, we will review how the moment-generating function for the dynamical activity, in the ensemble of all possible trajectories, can be expressed as the norm of a tilted probability vector.

##### A. Dynamical activity

The dynamical activity  $K(\omega_t)$  is a trajectory-dependent random variable, which forces us to track the probability vector  $|P(k, t)\rangle = \sum_{\mathbf{x}} P(\mathbf{x}, k, t) |\mathbf{x}\rangle$  that the system is in any configuration  $\mathbf{x}$  and has activity  $k$  at time  $t$ . Thus  $P(\mathbf{x}, k, t) = \langle \mathbf{x} | P(k, t) \rangle$ . The moment-generating function  $\mathcal{Z}_K(s)$  for the dynamical activity is

$$\mathcal{Z}_K(s) = \langle e^{sK(\omega_t)} \rangle_{\omega_t} = \sum_{k=0}^{\infty} \langle \mathbf{1} | e^{sk} | P(k, t) \rangle. \quad (11)$$

Here the brackets  $\langle \rangle_{\omega_t}$  denote the ensemble average over all dynamical trajectories  $\omega_t$  leading up to time  $t$  from an initial configuration  $\mathbf{x}_0 \in \{0, 1\}^N$ . The unit column vector  $\langle \mathbf{1} |$  is the *flat state*, such that the contraction of  $\langle \mathbf{1} |$  with a many-body vector returns the  $L^1$ -norm of that vector.

The moment-generating function  $\mathcal{Z}_K(s)$  is more conveniently computed as the  $L^1$ -norm of a *tilted probability vector*  $|P(s, t)\rangle$ , defined as the Laplace transform of  $|P(k, t)\rangle$  [57]:

$$\mathcal{Z}_K(s) = \langle \mathbf{1} | P(s, t) \rangle, \quad (12a)$$

$$|P(s, t)\rangle = \sum_k e^{sk} |P(k, t)\rangle. \quad (12b)$$

The tilted probability vector evolves in time by the action of a tilted generator  $\tilde{W}(s)$ , which can be obtained from the Markov generator (1) by multiplying all off-diagonal elements with  $e^s$  [41,42,56,57,71],

$$\partial_t |P(s, t)\rangle = \tilde{W}(s) |P(s, t)\rangle, \quad (13)$$

with

$$\begin{aligned} \tilde{W}(s) = & \lambda \sum_{i=1}^{N-1} (\hat{n}_i \tilde{w}_{i+1}^{0 \rightarrow 1}(s) + \tilde{w}_i^{0 \rightarrow 1}(s) \hat{n}_{i+1}) \\ & + \sum_{i=1}^N \tilde{w}_i^{1 \rightarrow 0}(s) + \tilde{W}_{\text{driv}}(\alpha, s), \end{aligned} \quad (14)$$

such that

$$\tilde{w}^{0 \rightarrow 1}(s) = \begin{pmatrix} -1 & 0 \\ e^s & 0 \end{pmatrix}, \quad \tilde{w}^{1 \rightarrow 0}(s) = \begin{pmatrix} 0 & e^s \\ 0 & -1 \end{pmatrix}. \quad (15)$$

At  $s = 0$ , the tilted generator reduces to the Markov generator (1). Away from  $s = 0$ , the tilted generator is neither Hermitian nor stochastic, so it will not conserve total probability. For

stochastic processes with local detailed balance, there does exist a similarity transformation to a Hermitian matrix [46,57]. In the case of the contact process, the infection transmission requires nearest-neighbor interactions, while recovery is a purely local process, and hence the process does not satisfy local detailed balance.

The exponential tilting weighs trajectories by their activity. When  $s > 0$ , active trajectories are exponentially favored, while for  $s < 0$  they are exponentially suppressed. In the case of the boundary driven system, the tilted generator becomes diagonal in the extreme inactive limit  $s \rightarrow -\infty$ , with leading eigenvalue  $-2\alpha$ , due to the boundary driving term. The corresponding eigenvector is the completely healthy state: the product state with all sites in the  $|0\rangle$  state.  $|P(s, t)\rangle$  reduces to the time-dependent probability vector  $|P(t)\rangle$  at  $s = 0$ , which at late times becomes the NESS of the driven system. This is not a product state, but contains correlations between the sites in the chain. No known analytical solution for the NESS exists in the literature to the best of our knowledge.

##### B. Large deviation principle

For the driven contact process, the absorbing state is eliminated and the system is driven into a NESS. In that case, the probability distribution  $P(k, t)$  of observing activity  $k$  at time  $t$  satisfies a large deviation (LD) principle, such that at late times,  $P(k, t) \sim e^{-t\phi(k/t)}$ , with  $\phi(k)$  the LD rate function [59]. The moment-generating function  $\mathcal{Z}_K(s)$  similarly satisfies a LD principle:

$$\mathcal{Z}_K(s) \sim e^{t\Theta(s)}, \quad \Theta(s) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \langle e^{sK(\omega_t)} \rangle. \quad (16)$$

The scaled cumulant-generating function (SCGF)  $\Theta(s)$  plays the role of dynamical free energy and is related to the LD rate function  $\phi(k)$  by Legendre transform [59],

$$\phi(k) = \sup_s (sk - \Theta(s)). \quad (17)$$

As the tilted probability vector is  $|P(s, t)\rangle = e^{\tilde{W}(s)t} |x_0\rangle$ , the SCGF becomes the leading eigenvalue  $\lambda_0$  of the tilted generator  $\tilde{W}(s)$ ,

$$\Theta(s) = \lambda_0(\tilde{W}(s)). \quad (18)$$

Here, the similarity between the tilted generator and a many-body Hamiltonian comes into play. Using the MPS ansatz for the many-body probability vector  $|P\rangle$ , we can obtain the leading eigenvalue of  $\tilde{W}(s)$  by a variational search over MPS states:

$$\Theta(s) = \max_{|\Psi\rangle} \frac{\langle \Psi | \tilde{W}(s) | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \quad (19)$$

where  $|\Psi\rangle$  is a many-body vector and  $\langle \Psi |$  its complex conjugate. In the coming sections, we will use variational MPS algorithms to find efficient representations for the leading eigenvector of  $\tilde{W}(s)$  and use this to compute the NESS, the SCGF, and other observables directly in the ensemble average over all trajectories.

## V. NESS COMPUTATION BY VARIATIONAL MPS

Here we show how the NESS of the 1D contact process can be found by a variational search over MPS states (9). We comment on the difference with sampling-based (Monte-Carlo) methods, and we demonstrate the accuracy of our approximations. To illustrate the effectiveness of the MPS representation, we compute the distribution of gaps of length  $k$  in the NESS, including the expectation values for large gaps, which are difficult to obtain accurately using sampling-based methods.

### A. Variational MPS methods

The NESS is the leading right eigenvector of the Markov generator (1), which corresponds to an eigenvalue of 0. The eigenvector is found in MPS form by sequentially optimizing for a single tensor  $A_i$  in (9), while keeping all other tensors constant, and iteratively sweeping over the chain until convergence. There are many good reviews detailing variational MPS algorithms for many-body quantum systems, such as [24,25,72]. Here, we will focus mainly on the differences with the usual MPS algorithms arising from the fact that we are optimizing for a probability distribution and not a complex-valued wave function.

The Markov generator  $\hat{W}$  is not a Hermitian matrix, and hence its left and right eigenvectors are not related to each other by complex conjugation. This does not, however, pose any serious problem, as long as we consistently update the MPS using the right eigenvectors of  $\hat{W}$ . The Markov generator is first represented as a matrix product operator (MPO), with constant bond dimension  $D_W = 4$  (see [24] for a review on the construction of MPOs). Schematically, this is done by representing  $\hat{W}$  as a chain of rank-4 tensor of dimensionality (2,2,4,4), except at the boundaries of the chain, where the tensors are of rank 3. In the case of the boundary driven system, the MPO is constructed as follows:

$$\hat{W} = LM \dots MR$$

$$= \begin{array}{c} 2 \\ | \\ \text{---} L \text{---} M \text{---} \dots \text{---} M \text{---} R \text{---} \\ | \\ 2 \end{array} \quad \begin{array}{c} 2 \\ | \\ \text{---} M \text{---} \\ | \\ 2 \end{array} \quad \dots \quad \begin{array}{c} 2 \\ | \\ \text{---} M \text{---} \\ | \\ 2 \end{array} \quad \begin{array}{c} 2 \\ | \\ \text{---} R \text{---} \\ | \\ 2 \end{array} \quad (20)$$

with

$$M = \begin{pmatrix} \mathbb{1} & 0 & 0 & 0 \\ \lambda \hat{w}^{0 \rightarrow 1} & 0 & 0 & 0 \\ \lambda \hat{n} & 0 & 0 & 0 \\ \hat{w}^{1 \rightarrow 0} & \hat{n} & \hat{w}^{0 \rightarrow 1} & \mathbb{1} \end{pmatrix} \quad (21)$$

and

$$L = (\hat{w}^{1 \rightarrow 0} + \alpha \hat{w}^{0 \rightarrow 1} \quad \hat{n} \quad \hat{w}^{0 \rightarrow 1} \quad \mathbb{1}), \quad (22)$$

$$R = \begin{pmatrix} \mathbb{1} \\ \lambda \hat{w}^{0 \rightarrow 1} \\ \lambda \hat{n} \\ \hat{w}^{1 \rightarrow 0} + \alpha \hat{w}^{0 \rightarrow 1} \end{pmatrix}. \quad (23)$$

The single tensor optimization problem for site  $i$  is then constructed by removing the local tensors  $A_i$  and  $A_i^\dagger$  from the contraction  $\langle \Psi | \hat{W} | \Psi \rangle$ , which creates a  $D^2 d$ -dimensional

matrix  $\hat{W}_{\text{eff}}^i$ . For instance, for  $N = 5$  and  $i = 3$ , the effective generator  $\hat{W}_{\text{eff}}$  is

$$\hat{W}_{\text{eff}} = \begin{array}{c} \text{---} A_1 \text{---} A_2 \text{---} \text{---} A_4 \text{---} A_5 \text{---} \\ | \quad | \quad | \quad | \quad | \\ L \quad M \quad M \quad M \quad R \\ | \quad | \quad | \quad | \quad | \\ \text{---} A_1 \text{---} A_2 \text{---} \text{---} A_4 \text{---} A_5 \text{---} \end{array} \quad (24)$$

The tensor  $A_i$  in the MPS is then replaced with the leading right eigenvector of  $\hat{W}_{\text{eff}}^i$ . This vector is found by the implicitly restarted Arnoldi method [73], iteratively optimizing for the eigenvector corresponding to the largest real eigenvalue. In practice, it is computationally more efficient to contract  $\hat{W}_{\text{eff}}^i$  with the local tensor  $A_i$  in the iterative eigenvector search directly, instead of explicitly constructing the  $dD^2$ -dimensional matrix (24). In this way, the eigenvector search comes at a computational cost which scales as  $D^3$ . It is also beneficial to form the optimization problem for two sites simultaneously. The resulting effective generator is then a  $d^2 D^2$ -dimensional matrix, and its eigenvector can be brought back to the MPS form by performing a singular value decomposition (SVD). After the SVD, we truncate the singular value spectrum by dropping all singular values below a fixed threshold  $\sim 10^{-16}$ . This procedure dynamically adjusts the bond dimension and keeps only singular values that are relevant. In practice, we need to provide also an upper bound on the bond dimension, usually taken to be around  $D_{\text{max}} \sim 200$ , to prevent the bond dimension from becoming too large. There are hence two ways to dynamically control the bond dimension of the MPS, one by truncating the SV spectrum on each bond, and another by controlling the maximal bond dimension for each bond.

One iterative sweep across the MPS starts with sequentially replacing pairs of local tensors  $A_i, A_{i+1}$ , starting from site (1,2) to the end, followed by sequential pairwise updates in the other direction. The backward sweep is done for complementary pair combinations with regard to the forward sweep, i.e., if in the forward sweep sites (1, 2), (3, 4), ... are paired together, the backward sweep pairs the sites ..., (5, 4), (3, 2) followed with a single site update of the first site. By repeatedly sweeping across the chain in this fashion, the MPS ansatz converges to the leading right eigenvector  $|\Psi_0\rangle$ , which is the NESS of the Markovian process.

Once the MPS representation  $|\Psi_0\rangle$  is found, we can use it to compute observables. It is important to note that  $|\Psi_0\rangle$  is directly corresponding to a probability distribution (and not its  $L^2$  norm), but the optimization procedure does not produce a normalized vector. Therefore, the probability distribution  $P(\mathbf{x})$  over all configurations  $\mathbf{x}$  in the NESS is given as

$$P(\mathbf{x}) = \frac{1}{\langle \mathbf{1} | \Psi_0 \rangle} |\Psi_0\rangle. \quad (25)$$

Likewise, we compute observables in the  $L^1$  norm, while dividing by the  $L^1$  norm of the total vector  $\langle \mathbf{1} | \Psi_0 \rangle$ . For instance, the expected density in the chain is computed as

$$\langle n \rangle = \frac{\langle \mathbf{1} | \frac{1}{N} \sum_{i=1}^N \hat{n}_i | \Psi_0 \rangle}{\langle \mathbf{1} | \Psi_0 \rangle}, \quad (26)$$

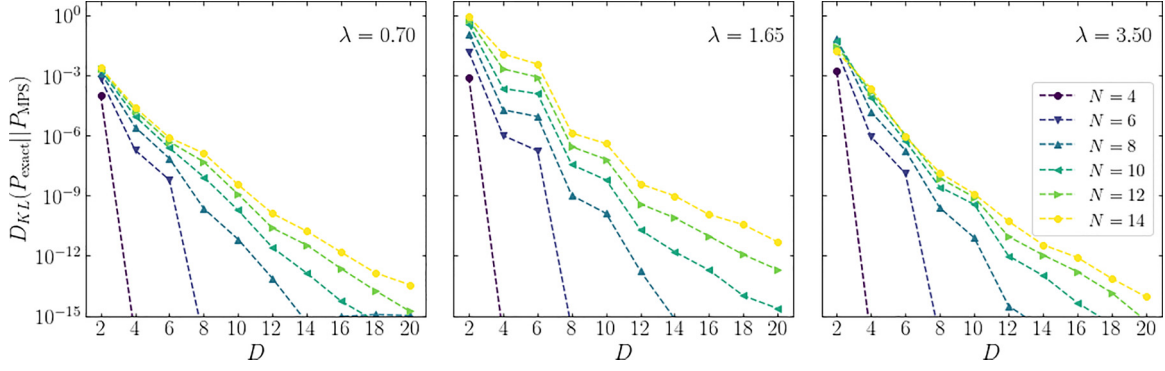


FIG. 2. The Kullback-Leibler divergence between the NESS found as the leading eigenvector of the explicit construction of the  $2^N$ -dimensional Markov generator  $P_{\text{exact}}$  and the MPS representation of the NESS  $P_{\text{MPS}}$  for various values of  $D$  and  $N$  and  $\lambda = 1$  (left),  $\lambda \simeq \lambda_c$  (middle), and  $\lambda = 3$  (right). Parameters for this plot are  $\epsilon = 10^{-8}$  and  $\alpha = \frac{1}{N^2}$  with spontaneous infection at all empty sites.

and similarly for other observables. The contraction is implemented efficiently by representing the observable of interest as an MPO, and performing a horizontal contraction of the MPO with the state  $|\Psi_0\rangle$  on top and a direct product of local flat states  $\langle \mathbf{1} | = (1, 1)$  on the bottom. If one first contracts the lower physical index of the MPO with the flat state, followed by the contraction of the resulting tensor with the local tensors  $A_i$ , this can be done with a computational cost scaling as  $O(ND_{\text{MPO}}D^2)$ , where  $D_{\text{MPO}}$  is the bond dimension of the MPO.

The local optimization process is repeated until the density in the chain has converged, meaning that the relative difference between  $\langle n \rangle_k$  after the  $k$ th sweep and  $\langle n \rangle_{k-1}$  has dropped below a threshold value:  $|\frac{\langle n \rangle_{k-1} - \langle n \rangle_k}{\langle n \rangle_k}| < \epsilon$ , with  $\epsilon = 10^{-6}$ – $10^{-8}$  depending on system size. To demonstrate the accuracy of the MPS representation as a function of the bond dimension  $D$ , we show in Fig. 2 the Kullback-Leibler (KL) divergence between  $P_{\text{exact}}$ , obtained as the leading right eigenvector of the explicit construction of the  $2^N$ -dimensional Markov generator, and  $P_{\text{MPS}}$  obtained from the contraction of the MPS representation, for small system sizes and various values of the maximal bond dimension  $D$  and transmission rate  $\lambda$ . With bond dimensions  $D = 2^{N/2}$ , the MPS is exact and the KL divergence with the exact solution becomes  $O(10^{-16})$ , within machine precision. For smaller bond dimension the MPS representation is still very close to the exact solution, however larger system sizes do require a larger bond dimension to achieve the same accuracy, especially near the critical transmission rate  $\lambda_c$ .

In Fig. 3 we investigate the growth of the bond dimension with system size  $N$  for the three different driving protocols and three values of the transmission rate  $\lambda$ . We let the algorithm dynamically adjust the bond dimensions for each bond by truncating the singular values below a lower threshold, taken to be  $10^{-14}$ . The figure displays the maximal bond dimension in the chain as a function of system size, and we observe that in the active and inactive region, the bond dimension tends to a constant value for most driving protocols. At the critical point, the bond dimension is expected to grow linearly in  $N$  due to the emergent scale invariance [74]. In Fig. 3 the orange curves are close to criticality, but not exactly at the critical point due to finite size and driving effects, hence a scaling slower than linear is observed here. In any case, a growth of  $D$  which is at most linear in  $N$  ensures the

efficiency of the MPS representation of the NESS, compared to the maximal exponential growth of  $D$  with system size. By comparing with Fig. 2, we see that the KL divergence with the exact solution is extremely small ( $< 10^{-12}$ ) for the smaller system sizes ( $N \simeq 8$ – $14$ ). This shows that the MPS representation is also accurate.

### B. Distribution of gaps in the NESS

By searching for an efficient MPS representation, we are approximating directly the full  $2^N$ -dimensional probability vector in the NESS. Once the optimization process has finished to the desired accuracy, we can, in principle, compute efficiently any observable we want directly in the ensemble average over all configurations. This includes global and local densities, variances thereof, but also all of the higher moments, the complete moment-generating function for the observable, or marginals, by projecting the probability vector on specific microscopic configurations.

This is a fundamentally different approach from more common sampling-based methods, frequently employed for the analysis of such systems. Using these methods, one typically needs to generate a large number of trajectories and then compute the observables of interest as the empirical averages over the set of generated trajectories. This makes the accurate computation of expectation values of rare events very demanding, as it requires one to generate a large set of trajectories to even encounter the rare event once. Using the MPS method, computing expectation values for events which occur very infrequently is not more difficult than any other event, as we have access to an approximation of the full probability distribution over all possible configurations of the system.

To illustrate this point with an example pertaining to the current dynamical system, we use our MPS representation of the NESS to compute the distribution of gaps of length  $k$  in a chain of total length  $N$ . A gap of length  $k$  is defined as a configuration with  $k$  empty sites in a row, preceded and followed by an occupied site. The expected number of gaps of length  $k$  in a chain of length  $N$  can be computed from the MPS representation of the NESS by contracting it with an MPO of bond dimension  $k + 3$ , constructed as follows:

$$\hat{G}_k = L_k M_k \cdots M_k R_k \quad (27)$$

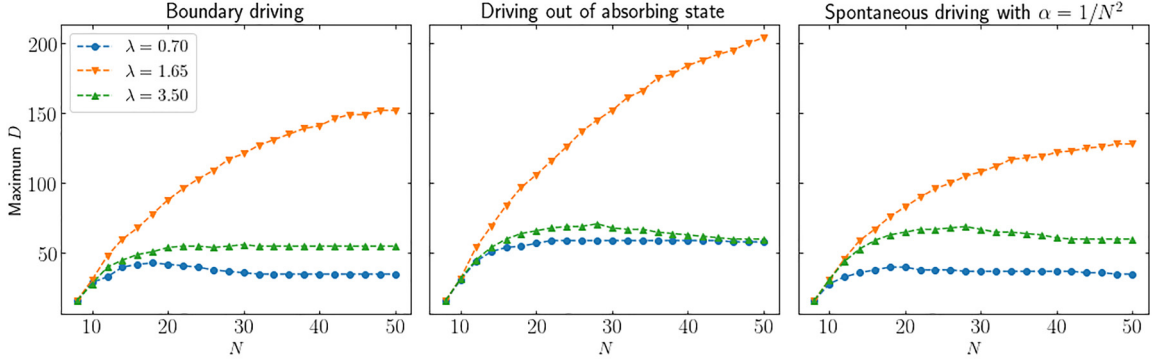


FIG. 3. The dynamically obtained maximal bond dimension  $D$  of the MPS representation of the NESS, obtained by truncating singular values below the threshold value of  $10^{-14}$ , as a function of system size  $N$  for three values of  $\lambda$  and three different driving protocols. The three values for  $\lambda$  were chosen to lie in the absorbing phase ( $\lambda = 0.7$ ), the active phase ( $\lambda = 3.5$ ), and in the critical region  $\lambda = \lambda_c = 1.649$ . The critical value of  $\lambda = \lambda_c$  is the critical value for the system in the thermodynamic limit without driving. The center figure is the closest to that system and displays the fastest increase in bond dimension close to criticality. The convergence criterion on the MPS was set to  $\epsilon = 10^{-8}$  for these plots.

with

$$M_k = \begin{pmatrix} \mathbb{1} & \hat{n} & 0 & \cdots & 0 \\ 0 & 0 & \hat{v} & & \\ & & & \ddots & \vdots \\ \vdots & & & & \hat{v} & 0 \\ 0 & & \cdots & 0 & \hat{n} & \mathbb{1} \end{pmatrix} \quad (28)$$

and

$$L_k = (\mathbb{1} \ \hat{n} \ 0 \ \cdots \ 0), \quad R_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \hat{n} \\ \mathbb{1} \end{pmatrix}. \quad (29)$$

When explicitly contracting the MPO (27), one obtains

$$\hat{G}_k = \sum_{i=1}^{N-1-k} \hat{n}_i \hat{v}_{i+1} \cdots \hat{v}_{i+k+1} \hat{n}_{i+k+2}. \quad (30)$$

This operator exactly projects the state on all possible positions where a gap of length  $k$  could be located in the chain.

In Fig. 4 we show the distribution of gaps in a chain of length  $N = 50$  computed by first optimizing an MPS for the NESS at a given  $\lambda$ , and then contracting the MPO (27) with the MPS. For this figure, we have used a driving protocol only on the left boundary of the chain. We notice that for sub- and supercritical values of  $\lambda$ , the expectation values for gaps drop exponentially with the length of the gap. For  $\lambda \sim \lambda_c$  the expected number of gaps remains larger for large  $k$  and follows a power law initially. For larger gaps, the expected value drops again exponentially, which is due to finite-size effects, as the gap size starts to become close to the length of the chain. The right panel of Fig. 4 shows that the expected number of gaps over  $N$  at  $\lambda = \lambda_c$  fits a power law  $E[g(k)] \sim k^{-\nu}$ , which is characteristic for scale-invariance at criticality. We find an

exponent of  $\nu = 1.71(6)$ , which corresponds to the fractal dimension [75,76] for the critical contact process.

In the left panel of Fig. 4, we compare the expected number of gaps with ensemble averages over trajectories generated by Markov chain Monte Carlo methods. The solid points with error bars are data obtained by averaging the observed number of gaps in  $2 \times 10^6$  configurations generated by Gillespie's algorithm [77]. The sampling-based algorithm gives excellent agreement with the MPS results for small gap sizes. Large gaps (of the order of 10–20 sites) are, however, extremely rare in both the absorbing and the active phase, making it difficult to obtain accurate statistics using sampling-based methods. In Fig. 4, the probability to find a gap of length  $k = 17$  is of the order of  $10^{-7}$  for both phases, which means that to compute this probability one must simulate at least  $10^7$  MC trajectories. More generally, because the probability of finding a gap decreases exponentially with gap length, the number of samples needed to derive accurate statistics on these rare events will grow exponentially with gaps size. The MPS, however, has no problem in providing an accurate prediction in the expected number of gaps in these cases. Here, the main computational cost comes from optimizing the MPS at the desired accuracy, which, as outlined above, scales as  $O(ND^3)$ . After that, the computation of the expectation value of a gap of length  $k$  amounts to contracting a MPO with bond dimension  $D_{\text{MPO}} = k + 3$ , which scales as  $O(ND^2k + NDk^2)$ . We find that these results are robust under varying the maximal bond dimension  $D$ , such that finite  $D$  effects are irrelevant for this observable.

## VI. LD STATISTICS FROM VARIATIONAL MPS

In this section, we will compute the scaled cumulant-generating function (SCGF) (or the dynamical free energy)  $\Theta(s)$  by a variational search over MPS states (9), while dynamically adjusting the bond dimension for each bond in the chain. We will use this to study the SCGF and other observables of interest, such as the dynamical activity at finite  $s$  and the susceptibility  $\chi(s) = \partial_s^2 \Theta(s)$ , whose peak characterizes the active/inactive transition for finite system sizes. The



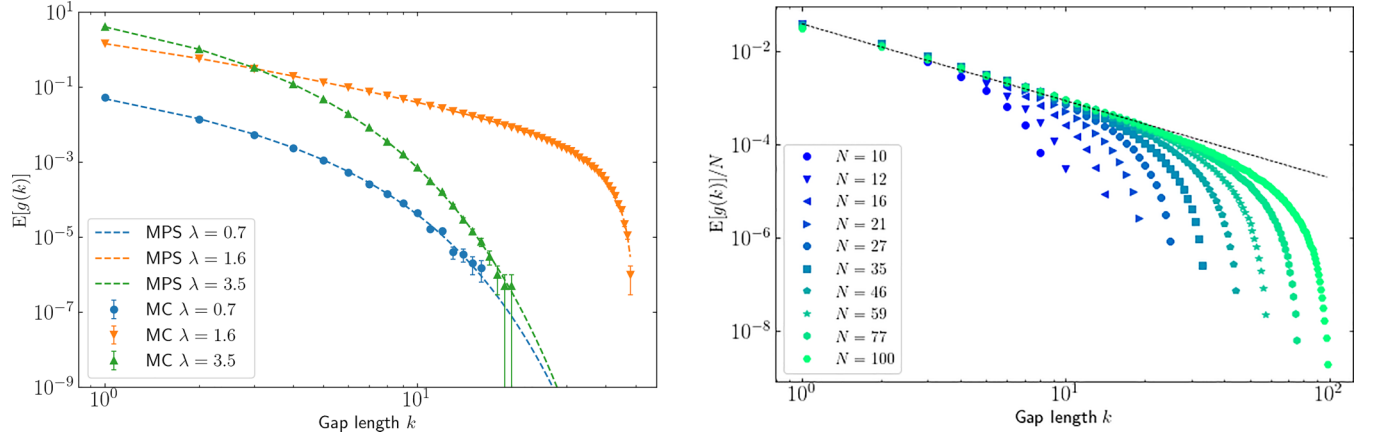


FIG. 4. Left: The distribution of gaps of length  $k$  in a chain of length  $N = 50$  at three different values of  $\lambda$ . The dotted lines are obtained from first finding an MPS representation of the NESS, followed by the contraction of this MPS with an MPO which gives the expected number of gaps of length  $k$ . Solid points with error bars are obtained from dynamical Monte Carlo sampling of over  $2 \times 10^6$  trajectories of the physical process, recorded once the system has reached stationarity. Right: The distribution of gaps at  $\lambda_c$  for various system sizes. The dotted line is a power-law fit for small  $k$ .

results show a limiting behavior consistent with a discontinuous phase transition in dynamical activity between the active and inactive phases, as a function of  $s$ , for any value of  $\lambda$ . The dynamical system hence contains both a continuous phase transition in the density of the stationary state (while varying  $\lambda$ ) and a discontinuous phase transition in the dynamical activity of trajectories (while varying  $s$ ).

#### A. MPS methods for the SCGF

To obtain an MPS representation that is both accurate and efficient, we have to find the optimal bond dimension  $D$ . This should be just big enough to give accurate predictions, but as small as possible to optimize computational efficiency. The goal in this section is to obtain the leading eigenvector  $\Theta(s)$  of the tilted generator from (19). We do this by dynamically adjusting the bond dimension for each bond, as described in the previous section. In this case, however, the convergence criteria on the MPS is based on the variance in  $\Theta(s)$ , given by  $\text{var}[\Theta(s)] = \langle \mathbf{1} | \tilde{W}(s)^2 | \Psi(s) \rangle - \Theta(s)^2$ . When this drops below a threshold value, taken to be  $10^{-8}$ , we can assure the accuracy of the MPS. After having found the leading right eigenvector of  $\tilde{W}(s)$  in MPS form for a specific value of  $s$ , this state is used as the initial state for the next value of  $s$ .

We investigate the variance as a function of the maximal bond dimension  $D_{\text{max}}$  for  $N = 50$  and various values of  $\lambda$  and  $s$  in Fig. 5. Here, the SVD cutoff was removed, such that all bonds in the MPS have bond dimension  $D_{\text{max}}$ . This shows that even small bond dimensions give a good approximation in most cases, particularly when  $\lambda$  and  $s$  are chosen such that the process is in the inactive phase.

#### B. Critical threshold and finite-size scaling

In Fig. 6 we show  $\Theta(s)/N$  for various system sizes at transmission rate  $\lambda = 2 > \lambda_c$ . We see that for trajectories tilted towards the active region ( $s > 0$ ), all lines lie perfectly on top of each other. The curves all pass through the origin, as at this point the tilted generator becomes the Markov generator

with leading eigenvalue equal to zero. Around  $s = 0$ , the SCGF grows linearly in both  $s$  and  $N$ , i.e.,  $\Theta(s) \sim sN$ . For large negative  $s$ , the SCGF converges to  $-2$ , as for inactive trajectories the only activity comes from the boundary driving terms. At a critical value for  $s = s_c(N)$ , a kink in the free energy appears, leading to a jump in the activity  $K(s) = \partial_s \Theta(s)$ , illustrated in the middle panel of Fig. 6. This is characteristic of a discontinuous phase transition. For  $\lambda > \lambda_c$ , the critical value  $s_c(N)$  converges to zero in the large- $N$  limit. The activity  $K(s)$  gives the expectation value of the dynamical activity  $K$  at  $s = 0$ , and away from  $s = 0$  it represents the activity in the tilted trajectories. The lower panel shows the rate function  $\phi(k)$ , computed by performing the Legendre transform (17). Its value gives the rate of the exponential decay of  $P(t, k)$  at late times, such that its zero gives the expected value for the activity per unit time in the NESS. The probability of observing trajectories with activity smaller than average decays at a rate independent of  $N$ , as the inactive regime is dominated by the boundary driving. In Appendix we display similar plots for a value of  $\lambda = 0.7 < \lambda_c$  in the absorbing phase.

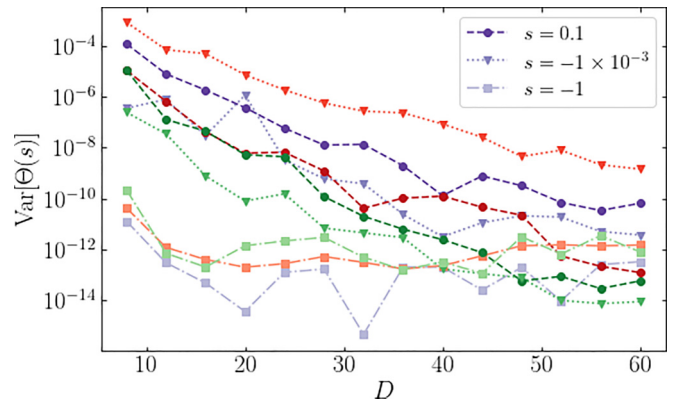


FIG. 5. The variance in  $\Theta(s)$  as a function of the bond dimension  $D$  for various values of  $s$ . Different colors indicate different  $\lambda$  values: Purple lines correspond to  $\lambda = 1$ , reds to  $\lambda = \lambda_c$ , and greens to  $\lambda = 3.5$ . This plot was obtained for  $N = 50$  using boundary driving.

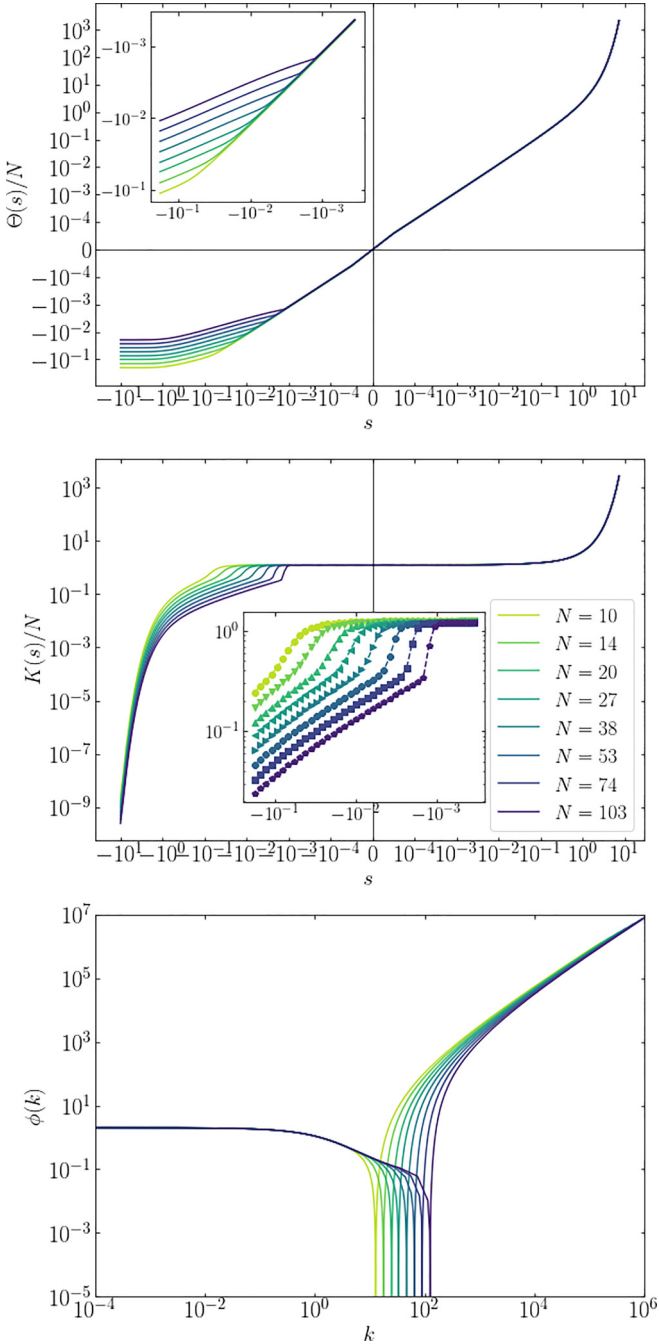


FIG. 6. The scaled cumulant-generating function (top) normalized by system size for various  $N$  at  $\lambda = 2 > \lambda_c$  using the boundary driving protocol. The middle panel shows the activity per site as a function of  $s$ , obtained as  $K(s)/N = \partial_s \Theta(s)/N$ . The rate function (bottom) clearly shows zeros at the expected value for the activity. Due to the boundary driving, inactive trajectories (small  $k$ ) all have the same constant decay rate  $\phi(k)$ .

The susceptibility  $\chi(s) = \partial_s^2 \Theta(s)$  is shown in the left panel of Fig. 7. It shows a peak at the critical value  $s_c(N)$  where the SCGF kinks and the activity jumps. Similar to [46], we find that the behavior of the susceptibility for  $s < s_c(N)$  is almost universal: leading into the peak, the susceptibility shows a scaling behavior  $\chi \sim (-s)^{-\gamma}$  with exponent  $\gamma = 1.4498(6)$

regardless of system size. This exponent does depend on the value of  $\lambda$ .

The location of  $s_c(N)$  can be estimated by the location of the peak in  $\chi(s)$ . This reveals that  $s_c(N)$  in turn satisfies a scaling law  $s_c(N) \sim -N^{-\alpha}$  whenever  $\lambda > \lambda_c$ . We show the critical values  $s_c(N)$ , as obtained from the peak in the susceptibility as a function of  $1/N$  in the middle panel of Fig. 7. The dashed lines are linear fits used to obtain  $\alpha$ , which is shown against the infection parameter  $\lambda$  in the right panel of Fig. 7.

### C. Density of infected sites at finite $s$

From the optimized tilted probability vector  $|\Psi_0(s)\rangle$ , we can obtain information on the density of infected sites at finite values of the dual parameter  $s$ . This gives the expectation value for the density on trajectories weighted by their activity; for  $s < 0$ , inactive trajectories are more likely, while  $s > 0$  corresponds to more active trajectories.

The density in the chain is computed efficiently by contracting an MPO as (26), now at finite dual parameter  $s$ . Likewise, the variance in density is computed by contracting the square of the density MPO with the obtained tilted probability vector. In Fig. 8, we show the density of infected sites and the variance in density  $N(\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2)$  in the chain at various lengths as a function of the tilting parameter  $s$  at  $\lambda = 2 > \lambda_c$ . It is clear from the plot that less active trajectories (lower  $s < 0$ ) correspond to less dense configurations, while more active trajectories are hardly any denser than those of the Markovian processes at  $s = 0$ . At the critical points  $s_c(N)$ , the expected density drops abruptly and the variance in density shows a peak. The heat plot in the right panel of Fig. 8 shows the density per site in the chain for  $N = 103$ . For  $s > s_c$ , the density is uniform over the whole chain. The active-inactive transition at  $s_c$  (indicated by the red dashed line) clearly corresponds to a sharp drop in density in the middle of the chain. Below  $s_c$ , one can observe the diffusion from the boundaries due to the driving terms (4): only the region close to the boundary stays infected.

## VII. ENTROPIC MEASURES

The formulation of the tilted many-body probability distribution as an MPS makes it possible to efficiently compute a variety of entropic measures for the stochastic process. As we are optimizing for a probability distribution, the Shannon entropy  $H(\mathbf{x})$  would be the first quantity to look at, followed closely by the mutual information between two halves of the chain. However, computing  $H(\mathbf{x}) = -\sum_{\mathbf{x}} P(\mathbf{x}) \log P(\mathbf{x})$  over all configurations  $\mathbf{x}$  would require taking the logarithm of an MPS, for which there are no efficient methods known to us at the moment. Fortunately, entropic measures based on (second) Rényi entropy can be computed efficiently [61,63].

The *second Rényi entropy*, or the *collision entropy*  $H_2(\mathbf{x})$ , is defined as

$$H_2(\mathbf{x}) = -\log_2 \sum_{\mathbf{x}} P(\mathbf{x})^2. \quad (31)$$

This can be obtained efficiently in terms of the  $L^2$ -norm of the distribution  $|\Psi_0(s)\rangle$ , as we can express the collision

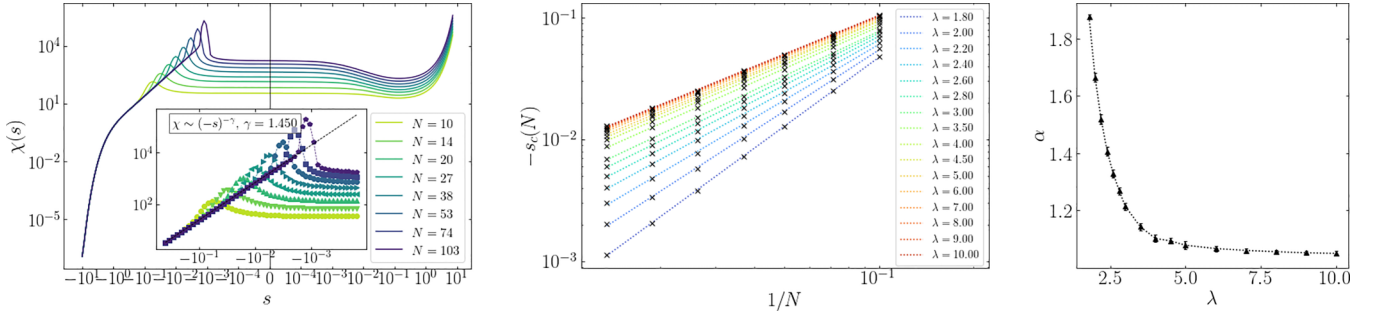


FIG. 7. Left: The susceptibility  $\chi(s) = \partial_s^2 \Theta(s)$  shows a peak for finite  $N$  at a critical value of  $s = s_c(N)$  which becomes more narrow for increasing  $N$ . Middle: The critical values  $s_c(N)$  at different values for the infection rate  $\lambda > \lambda_c$ . The results show a scaling of  $s_c(N) \sim N^{-\alpha}$  with  $\alpha > 1$ . Right: The exponents  $\alpha$  as a function of infection rate  $\lambda$ .

entropy as

$$H_2(\mathbf{x}) = -\log_2 \left( \frac{\langle \Psi_0 | \Psi_0 \rangle}{\langle \mathbf{1} | \Psi_0 \rangle^2} \right). \quad (32)$$

Here  $\langle \Psi_0 | \Psi_0 \rangle$  is computed efficiently by contracting the following tensor network site by site:

$$\langle \Psi_0 | \Psi_0 \rangle = \begin{array}{c} \text{---} A_1 \text{---} A_2 \text{---} \dots \text{---} A_N \text{---} \\ | \\ \text{---} A_1 \text{---} A_2 \text{---} \dots \text{---} A_N \text{---} \end{array}. \quad (33)$$

By computing the Rényi entropy in this way, we can make use of the MPS approximation of the full  $2^N$ -dimensional probability distribution, without actually ever having to compute it explicitly. In this way, biases in entropic measures due to undersampling [64,65] can be avoided.

Considering a partition of the chain in two regions,  $L$  and  $R$ , the *second Rényi entanglement entropy*  $H_2(L)$  of the region  $L$  is defined as the second Rényi entropy of the states distribution marginalized over the complementary region  $R$ . In terms of tensor networks, this can be computed by first contracting all sites in region  $R$  with the local flat state  $\langle \mathbf{1} | = (1 \ 1)$ , followed by contracting the marginal distribution over  $L$  with a copy of itself. Hence, we have

$$H_2(L) = -\log_2 \left( \frac{\langle \Psi_0(L) | \Psi_0(L) \rangle}{\langle \mathbf{1} | \Psi_0 \rangle^2} \right) \quad (34)$$

with

$$\langle \Psi_0(L) | \Psi_0(L) \rangle = \begin{array}{c} \text{---} A_1 \text{---} \dots \text{---} A_L \text{---} A_R \text{---} \dots \text{---} A_N \text{---} \\ | \quad \quad \quad | \quad \quad \quad | \\ \text{---} A_1 \text{---} \dots \text{---} A_L \text{---} A_R \text{---} \dots \text{---} A_N \text{---} \end{array}. \quad (35)$$

Here the green nodes denote local flat states  $\langle \mathbf{1} |$ . The second Rényi entanglement entropy  $H_2(R)$  for region  $R$ , defined as the complement of  $L$ , is computed similarly, by marginalizing over the sites in region  $L$  followed by computing the second Rényi entropy of region  $R$ .

Using the above quantities, one can compute the second Rényi mutual information (RMI)  $H_2(L; R)$  between the two subregions  $L$  and  $R$ :

$$H_2(L; R) = H_2(L) + H_2(R) - H_2(L, R). \quad (36)$$

Here  $H_2(L, R) = H_2(\mathbf{x})$  is defined above in (32), and  $H_2(L)$  and  $H_2(R)$  are the second Rényi entanglement entropies of regions  $L$  and  $R$ , respectively.

The RMI is particularly interesting since it provides a measure of the amount of information contained in a subregion of the system about the rest of the system. Conventional measures based on the correlation length are susceptible to

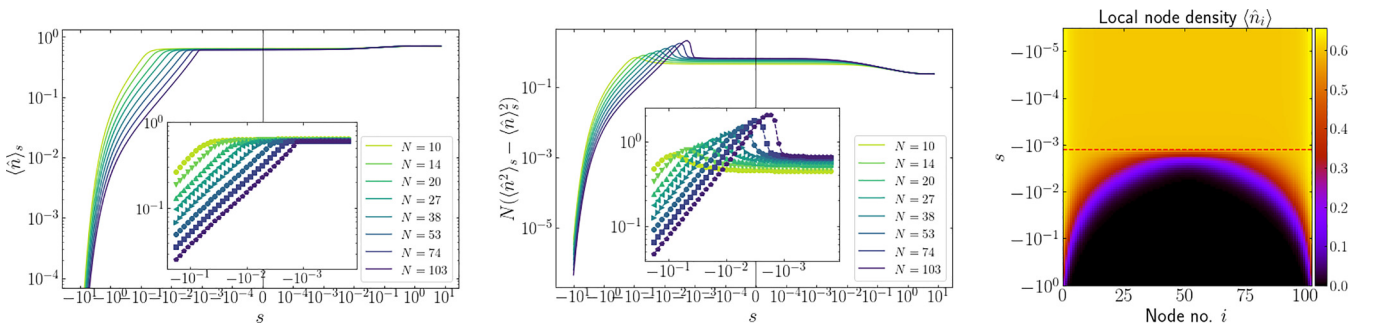


FIG. 8. The left panel shows the density of infected sites in the chain of various lengths at  $\lambda = 2$  as a function of  $s$ . Trajectories with lower activity clearly correspond to those with less dense configurations. The middle panel displays the variance in global density, which peaks when the expected density shows a sudden drop. The right panel shows the local node density as a function of  $s$  and the node position in the chain for  $N = 103$ , respectively. The dashed red line indicates the critical value  $s_c(N = 103)$  as determined from the peak in susceptibility.

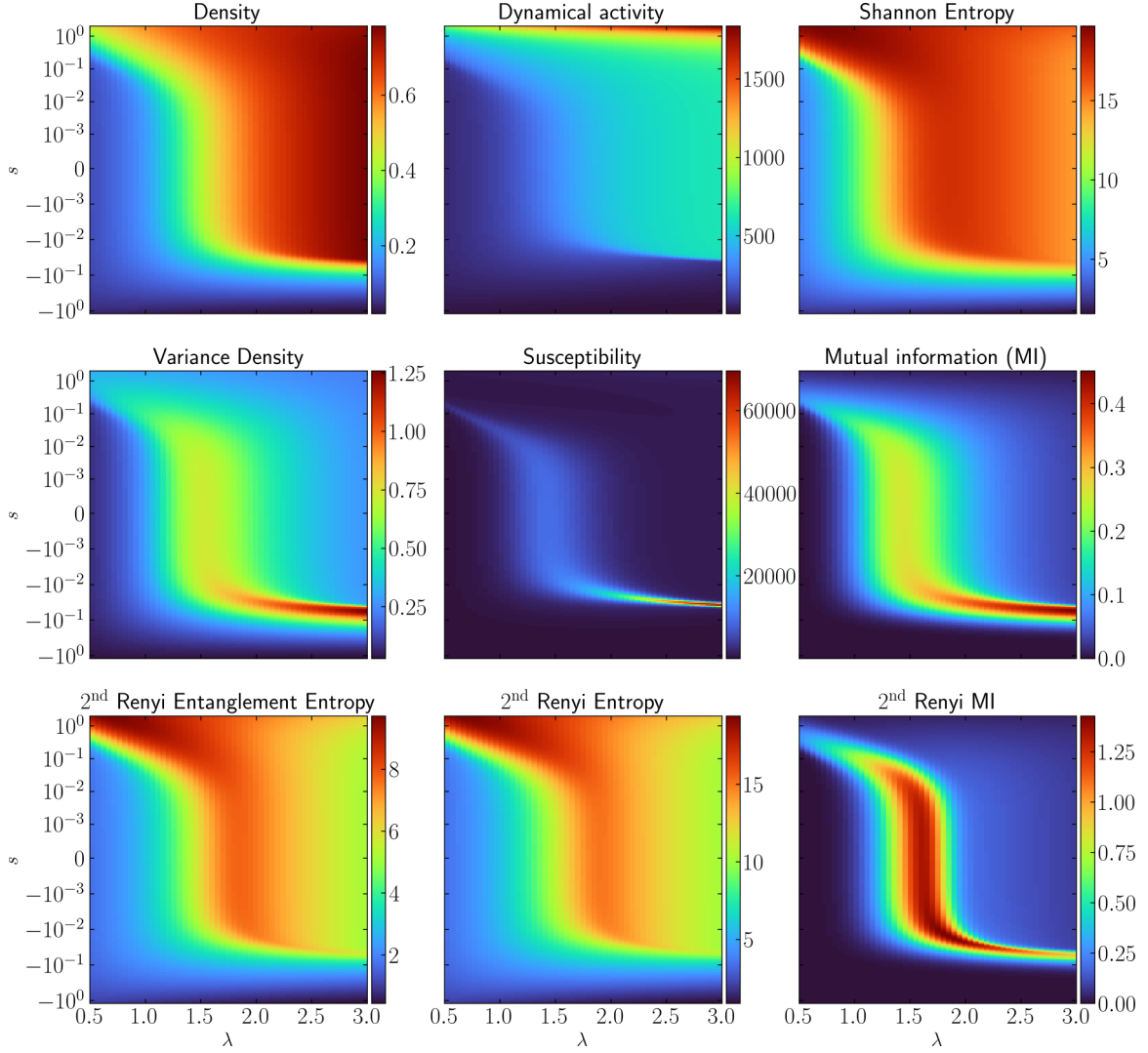


FIG. 9. Various observables for the contact process in the chain of length  $N = 20$  in the  $(\lambda, s)$  plane.

overlooking hidden correlations, whereas the RMI is based on the total amount of information of a subsystem about the complementing region, without overlooking hidden correlations [61]. Furthermore, it has been shown in [66] that the second RMI provides a way to detect phase transitions in classical equilibrium systems, without any prior knowledge of an order parameter or other thermodynamic quantities. Below, we demonstrate that the second RMI can be computed efficiently from the MPS representation of the NESS for the contact process, and we use it to identify the absorbing state transition by a finite-size scaling argument.

In Fig. 9, we show and compare various quantities in the  $(\lambda, s)$ -plane for a chain of length  $N = 20$ . For this chain length, it is still computationally feasible to compute the complete Shannon entropy and mutual information between two halves of the chain. These plots have been added for comparison. The division into subregions in this plot is always taken to be the half-chain, and driving at both boundaries was used here.

We find that the second Rényi mutual information shows a clear peak close to the region where the susceptibility and mutual information is maximal. However, a closer inspection shows that the peak in  $H_2(L; R)$  is shifted slightly towards higher values of  $\lambda$ . This shows that the second Rényi mutual information is maximal just above the critical transmission rate  $\lambda_c$ .

In [66], the second RMI was used to detect phase transitions in classical equilibrium models. The analysis relied on a scaling argument, which shows the RMI contains a constant contribution (i.e., independent of subsystem area) which changes sign across the critical temperature  $T_c$ . This causes the finite-size RMI curves to “fan out” at  $T_c$  for different system sizes, producing a crossing of the curves. In our case, our system is driven out-of-equilibrium, however the RMI curves still show a crossing close to the critical infection rate  $\lambda_c$ . This is demonstrated in Fig. 10, which displays the RMI in the NESS for various system sizes as a function of  $\lambda$ . The right panel of Fig. 10 displays the crossing points of the



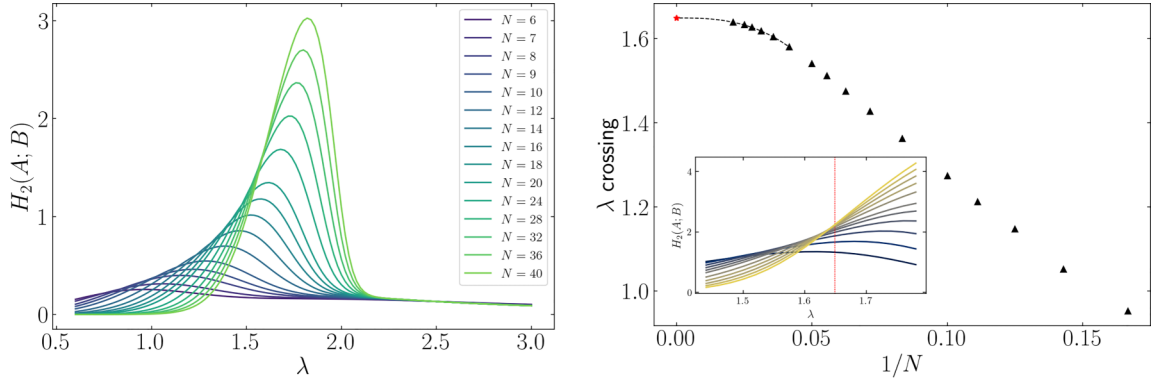


FIG. 10. The second Rényi mutual information (RMI) (left) in the NESS as a function of  $\lambda$  (with boundary driving). The right panel shows the locations  $\lambda(N)$  where the RMI curves for  $N$  and  $2N$  cross, as a function of  $1/N$ . The red star marks the critical value of the 1D contact process  $\lambda_c = 1.6489$ . The inset shows a detail of the crossing points in the RMI with  $N = 20, 24, 28, 32, 36, 40, 48, 56, 64, 72, 80$ . The dotted line is an extrapolation to  $N \rightarrow \infty$  based on Eq. (37).

RMI at system sizes  $N$  and  $2N$  for increasing  $N$ . For small system sizes, there is a considerable finite-size effect, as the intersections are not occurring at the same locations. As  $N$  increases, the location of the intersection creeps closer to  $\lambda_c$ . For chains of length  $\sim 80$ , the crossings already occur very close to the critical infection rate.

We extrapolate the data for the largest available values of  $N$  using the scaling ansatz,

$$\lambda_c(N) = \lambda_c \left( 1 + \frac{b}{N^\chi} + \dots \right). \quad (37)$$

This allows us to derive a critical transmission rate  $\lambda_c = 1.6487(6)$ , independently from using any order parameter. The obtained value is consistent with the known critical value for  $\lambda$ . Additionally, we obtain a value for the subleading scaling with  $N$ :  $\chi = 2.77(5)$ . This analysis demonstrates the ability to detect dynamical phase transitions on the basis of the RMI alone, without the need for an order parameter.

## VIII. CONCLUSIONS AND DISCUSSION

In this paper, we have used the matrix product state to study the one-dimensional contact process. We have found that efficient representations for the many-body probability vector in the thermodynamic limit  $t \rightarrow \infty$  are obtained in the active and inactive regimes, while at the critical threshold for the active/inactive phase transition, the bond dimension of the MPS grows at most linear with system size. We have illustrated how the MPS can be used for efficient computation of the distribution of gaps in a chain, even for large gap sizes corresponding to extremely rare events. We used the MPS to obtain accurate approximations of the scaled cumulant-generating function for the dynamical activity, which is dual to the large deviation rate function by Legendre transform. This analysis sheds light on how trajectories with different dynamical activity contribute to the properties of the system in the late time regime.

Our approach enables one to study entropic measures in a straightforward way, which is challenging to do with other approaches. In particular, we compute the second Rényi mutual information for the contact process. Previous methods

obtained the second RMI for equilibrium models relying on Monte Carlo simulations of the replicated partition function (see, for instance, [66]). In our case, we are lacking a Boltzmann distribution as we are looking at an out-of-equilibrium problem, so it is not clear how to obtain the RMI from applying the replica trick on a partition function. Instead, we obtained the second RMI directly from the MPS representation of the NESS, and we demonstrated its use to detect dynamical phase transitions.

The MPS methods described in this paper are complementary to the more frequently used Monte Carlo methods. The MPS allows one to approximate the full probability distribution over all configurations, which allows for the efficient computation of quantities where detailed knowledge of the full probability distribution is required. The main computational cost of the MPS methods comes from the optimization procedure, after which even higher-order observables may be computed efficiently. Monte Carlo methods still provide advantages over the methods described in this paper, especially in terms of ease of use, generalization to complex network interactions, and the ability to easily add more complexity to the microscopic interactions. A precise performance comparison between tensor networks and Monte Carlo methods depends on the details of the simulated systems and on the types of questions being addressed, i.e., what observables are computed, the distance to criticality, the desired accuracy, and the type of observables in which one is interested. It would be interesting to perform such a detailed performance comparison in future work.

This work opens the doors to tensor network applications in mathematical epidemiology. Obvious generalizations are the extension of these methods to the infinitely long chain with infinite-MPS techniques [78], to two- and higher-dimensional arrays using projected-entangled-pair states (PEPS) [79–82], or to (tree) networks with tree tensor networks [83–85] or with the multiscale entanglement renormalization ansatz [86]. One could also imagine formulating a tensor network state for more involved compartmental models of epidemiology along the lines illustrated in this paper, but with a closer connection to real-world infectious disease outbreaks. The tensor network could then be used to obtain probability

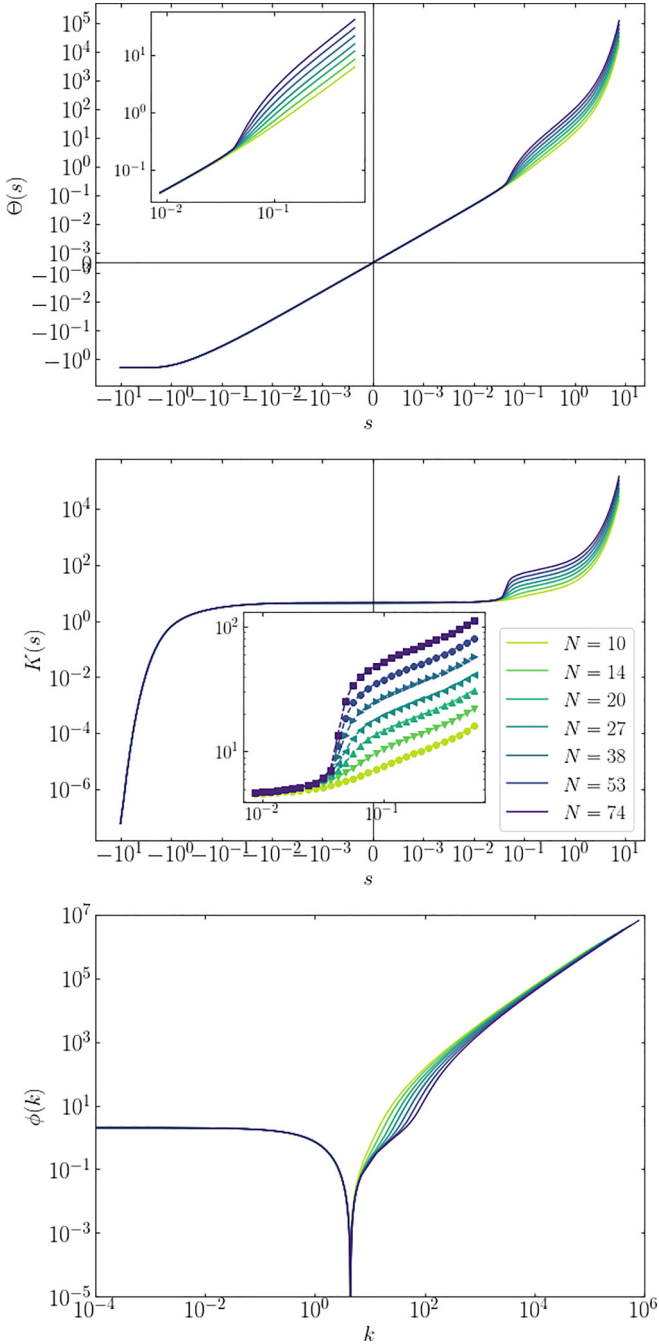


FIG. 11. The scaled cumulant-generating function (top) for various  $N$  at  $\lambda = 0.7 < \lambda_c$  using the boundary driving protocol. The middle panel shows the activity per unit time as a function of  $s$ , obtained as  $K(s) = \partial_s \Theta(s)$ . The rate function (bottom) shows zeros at the expected value for the activity. Due to the boundary driving, inactive trajectories (small  $k$ ) all have the same constant decay rate  $\phi(k)$ .

distributions over observables of relevance to the tracking and forecasting of the epidemic state through a population of interest. Measurements on the populace, such as test and tracing policies, could then be implemented into the tensor network by projecting certain sites on the measurement outcome. In this way, the state will become closer to the actual state of the population, while retaining the information of past correlations between sites not projected on.

While the one-dimensional contact process remains an interesting system for statistical physicists, in mathematical epidemiology one is more often interested in time-dependent spreading processes on complex networks [21]. It is therefore of great interest to explore how the methods described in this paper can be used to predict and forecast epidemic spreading in realistic scenarios. Major steps in this direction would be an extension to study the dynamical evolution by time-evolving block decimation algorithms (TEBD) [87] and to extend the methods to spreading processes on real-world, complex contact networks. We hope to address such questions in future work.

#### ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 101001604). W.M. acknowledges the support from the NWO Klein grant awarded to NWA route 2 in 2020.

#### APPENDIX: SCALED CUMULANT-GENERATING FUNCTION FOR $\lambda < \lambda_c$

Here we display our results for the SCGF  $\Theta(s)$  for values of  $\lambda$  in the absorbing phase. Figure 11 displays the SCGF  $\Theta(s)$  for  $\lambda = 0.7 < \lambda_c$ . Note that we plot here the SCGF and not the SCGF per site, as was plotted in Fig. 6 for  $\lambda > \lambda_c$ . We see that in this case for any  $N$ , the SCGF is approximately linear around  $s = 0$  with the same  $N$ -independent slope. This implies that the activity is independent of  $N$ , as illustrated as well by the middle panel of Fig. 11. Only for trajectories tilted towards the active region ( $s > 0$ ) is a jump in the dynamical activity visible.

The lower panel of Fig. 6 displays the rate function for the dynamical activity when  $\lambda = 0.7 < \lambda_c$ . The plot shows that the likelihood of observing trajectories with a dynamical activity that is average or lower decays exponentially with a rate independent of  $N$ , which is an effect of the boundary driving term. For trajectories with activity larger than average, the decay rate does depend on  $N$ , such that larger  $N$  corresponds to a lower rate function.

- [1] A. L. Lloyd and R. M. May, How viruses spread among computers and people, *Science* **292**, 1316 (2001).
- [2] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).

- [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.* **424**, 175 (2006).
- [4] C. Castellano, S. Fortunato, and V. Loreto, Statistical physics of social dynamics, *Rev. Mod. Phys.* **81**, 591 (2009).

- [5] A. Vespignani, Modelling dynamical processes in complex socio-technical systems, *Nat. Phys.* **8**, 32 (2012).
- [6] O. Artime and M. De Domenico, From the origin of life to pandemics: Emergent phenomena in complex systems, *Philos. Trans. R. Soc. A* **380**, 20200410 (2022).
- [7] J. P. Gleeson, Binary-State Dynamics on Complex Networks: Pair Approximation and Beyond, *Phys. Rev. X* **3**, 021004 (2013).
- [8] N. Masuda, M. A. Porter, and R. Lambiotte, Random walks and diffusion on networks, *Phys. Rep.* **716-717**, 1 (2017).
- [9] A. Ghavasieh, C. Nicolini, and M. De Domenico, Statistical physics of complex information dynamics, *Phys. Rev. E* **102**, 052304 (2020).
- [10] W. Merbis and M. de Domenico, Emergent information dynamics in many-body interconnected systems, *Phys. Rev. E* **108**, 014312 (2023).
- [11] V. Privman, *Finite Size Scaling and Numerical Simulation of Statistical Systems* (World Scientific, Singapore, 1990).
- [12] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Accuracy of mean-field theory for dynamics on real-world networks, *Phys. Rev. E* **85**, 026106 (2012).
- [13] J. Gao, B. Barzel, and A.-L. Barabási, Universal resilience patterns in complex networks, *Nature (London)* **530**, 307 (2016).
- [14] I. Z. Kiss, J. C. Miller, P. L. Simon *et al.*, *Mathematics of Epidemics on Networks*, Interdisciplinary Applied Mathematics (Springer, Cham, 2017), Vol. 46.
- [15] W. R. Young, A. J. Roberts, and G. Stuhne, Reproductive pair correlations and the clustering of organisms, *Nature (London)* **412**, 328 (2001).
- [16] B. Houchmandzadeh, Theory of neutral clustering for growing populations, *Phys. Rev. E* **80**, 051920 (2009).
- [17] E. Dumonteil, F. Malvagi, A. Zoia, A. Mazzolo, D. Artusio, C. Dieudonné, and C. De Mulatier, Particle clustering in monte carlo criticality simulations, *Ann. Nucl. Energy* **63**, 612 (2014).
- [18] H. W. Hethcote, The mathematics of infectious diseases, *SIAM Rev.* **42**, 599 (2000).
- [19] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, *Phys. Rev. Lett.* **86**, 3200 (2001).
- [20] F. D. Sahneh, C. Scoglio, and P. Van Mieghem, Generalized epidemic mean-field model for spreading processes over multilayer complex networks, *IEEE/ACM Trans. Netw.* **21**, 1609 (2013).
- [21] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.* **87**, 925 (2015).
- [22] D. Poulin, A. Qarry, R. Somma, and F. Verstraete, Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space, *Phys. Rev. Lett.* **106**, 170501 (2011).
- [23] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [24] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, *Ann. Phys.* **326**, 96 (2011).
- [25] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states, *Ann. Phys.* **349**, 117 (2014).
- [26] M. Schönberg, Application of second quantization methods to the classical statistical mechanics, *Il Nuovo Cimento (1943-1954)* **9**, 1139 (1952).
- [27] M. Doi, Second quantization representation for classical many-particle system, *J. Phys. A* **9**, 1465 (1976).
- [28] P. Grassberger and M. Scheunert, Fock-space methods for identical classical objects, *Fortschr. Phys.* **28**, 547 (1980).
- [29] L. Peliti, Path integral approach to birth-death processes on a lattice, *J. Phys.* **46**, 1469 (1985).
- [30] G. M. Schütz, Reaction-diffusion processes of hard-core particles, *J. Stat. Phys.* **79**, 243 (1995).
- [31] H. Hinrichsen, Non-equilibrium critical phenomena and phase transitions into absorbing states, *Adv. Phys.* **49**, 815 (2000).
- [32] G. Schütz, *Exactly Solvable Models for Many-body Systems Far from Equilibrium* (Academic, 2001), pp. 1–251.
- [33] W. Merbis and I. Lodato, Logistic growth on networks: Exact solutions for the susceptible-infected model, *Phys. Rev. E* **105**, 044303 (2022).
- [34] B. Derrida and J. L. Lebowitz, Exact Large Deviation Function in the Asymmetric Exclusion Process, *Phys. Rev. Lett.* **80**, 209 (1998).
- [35] T. H. Johnson, S. R. Clark, and D. Jaksch, Dynamical simulations of classical stochastic systems using matrix product states, *Phys. Rev. E* **82**, 036702 (2010).
- [36] J. de Gier and F. H. L. Essler, Large Deviation Function for the Current in the Open Asymmetric Simple Exclusion Process, *Phys. Rev. Lett.* **107**, 010602 (2011).
- [37] A. Lazarescu and K. Mallick, An exact formula for the statistics of the current in the tasep with open boundaries, *J. Phys. A* **44**, 315001 (2011).
- [38] M. Gorissen, A. Lazarescu, K. Mallick, and C. Vanderzande, Exact Current Statistics of the Asymmetric Simple Exclusion Process with Open Boundaries, *Phys. Rev. Lett.* **109**, 170601 (2012).
- [39] T. Nishino, Density matrix renormalization group method for 2d classical models, *J. Phys. Soc. Jpn.* **64**, 3598 (1995).
- [40] E. Carlon, M. Henkel, and U. Schollwöck, Density matrix renormalization group and reaction-diffusion processes, *Eur. Phys. J. B* **12**, 99 (1999).
- [41] M. Gorissen, J. Hooyberghs, and C. Vanderzande, Density-matrix renormalization-group study of current and activity fluctuations near nonequilibrium phase transitions, *Phys. Rev. E* **79**, 020101(R) (2009).
- [42] J. Hooyberghs and C. Vanderzande, Thermodynamics of histories for the one-dimensional contact process, *J. Stat. Mech.* (2010) P02017.
- [43] P. Helms and G. K.-L. Chan, Dynamical Phase Transitions in a 2D Classical Nonequilibrium Model via 2D Tensor Networks, *Phys. Rev. Lett.* **125**, 140601 (2020).
- [44] L. Causser, M. C. Bañuls, and J. P. Garrahan, Optimal Sampling of Dynamical Large Deviations in Two Dimensions Via Tensor Networks, *Phys. Rev. Lett.* **130**, 147401 (2023).
- [45] P. Helms, U. Ray, and G. K.-L. Chan, Dynamical phase behavior of the single-and multi-lane asymmetric simple exclusion process via matrix product states, *Phys. Rev. E* **100**, 022101 (2019).
- [46] M. C. Bañuls and J. P. Garrahan, Using Matrix Product States to Study the Dynamical Large Deviations of Kinetically Constrained Models, *Phys. Rev. Lett.* **123**, 200601 (2019).
- [47] L. Causser, M. C. Banuls, and J. P. Garrahan, Finite Time Large Deviations via Matrix Product States, *Phys. Rev. Lett.* **128**, 090605 (2022).

- [48] N. E. Strand, H. Vroylandt, and T. R. Gingrich, Using tensor network states for multi-particle brownian ratchets, *J. Phys. Chem.* **156**, 221103 (2022).
- [49] J. P. Garrahan and F. Pollmann, Topological phases in the dynamics of the simple exclusion process, [arXiv:2203.08200](https://arxiv.org/abs/2203.08200).
- [50] T. E. Harris, Contact interactions on a lattice, *Ann. Probab.* **2**, 969 (1974).
- [51] P. Grassberger, On the critical behavior of the general epidemic process and dynamical percolation, *Math. Biosci.* **63**, 157 (1983).
- [52] I. Jensen and R. Dickman, Time-dependent perturbation theory for nonequilibrium lattice models, *J. Stat. Phys.* **71**, 89 (1993).
- [53] J. Hooyberghs and C. Vanderzande, One-dimensional contact process: Duality and renormalization, *Phys. Rev. E* **63**, 041109 (2001).
- [54] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 2005).
- [55] T. H. Johnson, T. J. Elliott, S. R. Clark, and D. Jaksch, Capturing Exponential Variance Using Polynomial Resources: Applying Tensor Networks to Nonequilibrium Stochastic Processes, *Phys. Rev. Lett.* **114**, 090602 (2015).
- [56] V. Lecomte, C. Appert-Rolland, and F. Van Wijland, Thermodynamic formalism for systems with markov dynamics, *J. Stat. Phys.* **127**, 51 (2007).
- [57] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, First-order dynamical phase transition in models of glasses: An approach based on ensembles of histories, *J. Phys. A* **42**, 075007 (2009).
- [58] R. Chetrite and H. Touchette, Nonequilibrium markov processes conditioned on large deviations, in *Annales Henri Poincaré* (Springer, 2015), Vol. 16, pp. 2005–2057.
- [59] H. Touchette, The large deviation approach to statistical mechanics, *Phys. Rep.* **478**, 1 (2009).
- [60] L. Causser, M. C. Bañuls, and J. P. Garrahan, Optimal sampling of dynamical large deviations via matrix product states, *Phys. Rev. E* **103**, 062144 (2021).
- [61] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Area Laws in Quantum Systems: Mutual Information and Correlations, *Phys. Rev. Lett.* **100**, 070502 (2008).
- [62] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Entropy Scaling and Simulability by Matrix Product States, *Phys. Rev. Lett.* **100**, 030504 (2008).
- [63] S. O. Scalet, Á. M. Alhambra, G. Styliaris, and J. I. Cirac, Computable rényi mutual information: Area laws and correlations, *Quantum* **5**, 541 (2021).
- [64] L. Paninski, Estimation of entropy and mutual information, *Neural Comput.* **15**, 1191 (2003).
- [65] S. Panzeri, R. Senatore, M. A. Montemurro, and R. S. Petersen, Correcting for the sampling bias problem in spike train information measures, *J. Neurophysiol.* **98**, 1064 (2007).
- [66] J. Iaconis, S. Inglis, A. B. Kallin, and R. G. Melko, Detecting classical phase transitions with renyi mutual information, *Phys. Rev. B* **87**, 195134 (2013).
- [67] W. Merbis, StochasticTN: Stochastic tensor networks for the contact process (v0.1), Zenodo, 2023, <https://doi.org/10.5281/zenodo.8139314>.
- [68] W. Merbis, Exact epidemic models from a tensor product formulation, [arXiv:2102.11708](https://arxiv.org/abs/2102.11708).
- [69] J. R. G. de Mendonça, Precise critical exponents for the basic contact process, *J. Phys. A* **32**, L467 (1999).
- [70] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277 (2010).
- [71] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses, *Phys. Rev. Lett.* **98**, 195702 (2007).
- [72] F. Verstraete, V. Murg, and J. I. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143 (2008).
- [73] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods* (SIAM, 1998).
- [74] P. Calabrese and J. Cardy, Entanglement entropy and quantum field theory, *J. Stat. Mech.* (2004) P06002.
- [75] J.-P. Hovi, A. Aharony, D. Stauffer, and B. B. Mandelbrot, Gap Independence and Lacunarity in Percolation Clusters, *Phys. Rev. Lett.* **77**, 877 (1996).
- [76] F. Ginelli, H. Hinrichsen, R. Livi, D. Mukamel, and A. Politi, Directed percolation with long-range interactions: Modeling nonequilibrium wetting, *Phys. Rev. E* **71**, 026121 (2005).
- [77] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [78] G. Vidal, Classical Simulation of Infinite-Size Quantum Lattice Systems in One Spatial Dimension, *Phys. Rev. Lett.* **98**, 070201 (2007).
- [79] F. Verstraete and J. I. Cirac, Renormalization algorithms for quantum-many body systems in two and higher dimensions, [arXiv:cond-mat/0407066](https://arxiv.org/abs/cond-mat/0407066).
- [80] Y. Nishio, N. Maeshima, A. Gendiar, and T. Nishino, Tensor product variational formulation for quantum systems, [arXiv:cond-mat/0401115](https://arxiv.org/abs/cond-mat/0401115).
- [81] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac, Classical Simulation of Infinite-Size Quantum Lattice Systems in Two Spatial Dimensions, *Phys. Rev. Lett.* **101**, 250602 (2008).
- [82] P. C. G. Vlaar and P. Corboz, Simulation of three-dimensional quantum systems with projected entangled-pair states, *Phys. Rev. B* **103**, 205137 (2021).
- [83] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network, *Phys. Rev. A* **74**, 022320 (2006).
- [84] L. Tagliacozzo, G. Evenbly, and G. Vidal, Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law, *Phys. Rev. B* **80**, 235127 (2009).
- [85] T. Felser, S. Notarnicola, and S. Montangero, Efficient Tensor Network Ansatz for High-Dimensional Quantum Many-Body Problems, *Phys. Rev. Lett.* **126**, 170603 (2021).
- [86] G. Vidal, Class of Quantum Many-Body States That Can Be Efficiently Simulated, *Phys. Rev. Lett.* **101**, 110501 (2008).
- [87] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, *Phys. Rev. Lett.* **93**, 040502 (2004).