Non-equilibrium dynamics and large deviations in stochastic lattice models via tensor networks

University of Nottingham
UK | CHINA | MALAYSIA

Luke Causer

A thesis presented for the degree of Doctor of Philosophy

Supervisors: Prof. Juan P. Garrahan, Dr. Stephen Powell

School of Physics & Astronomy
The University of Nottingham
United Kingdom

December 2022

Abstract

Over the last few decades, numerical tensor networks have revolutionized the study of quantum many-body systems. Despite this success, their application to classical stochastic problems has not yet been extensively explored. This thesis investigates how tensor network methods can be applied to studying the slow dynamics and the dynamical fluctuations of kinetically constrained models used in the modelling of structural glasses.

This thesis is divided into three parts. It first gives a brief introduction to stochastic dynamics, and explains how the statistics of dynamical observables can be understood through the framework of large deviations. Various approaches to calculating the dynamical large deviations are explained, including the estimation of leading eigenvectors of deformed Markov generators, and trajectory path sampling. It is then followed by an overview of tensor networks in one and two dimensions, which can be used to extract extremal eigenvectors from stochastic generators and simulate time evolution.

The second part then investigates two kinetically constrained models: the "XOR-Fredrickson-Andersen" model, inspired by Rydberg atoms in their anti-blockade regime, and a stochastic Fredkin model, a direct stochastic generalization of the quantum spin model. Their steady-state properties and non-equilibrium dynamics are studied through theoretical and numerical techniques, including tensor networks and Monte Carlo sampling. Both models display slow and glassy dynamics, motivating the study of their dynamical large deviations. This is done to a high precision via tensor network methods, uncovering and detailing first-order dynamical phase transitions for each model.

The final part of this thesis aims to further develop the application of tensor networks to dynamical fluctuations in classical stochastic dynamics. To this end, a novel method of directly sampling the rare trajectories associated with the large deviations in one-dimensional stochastic dynamics is developed. This is then accompanied by a method which directly simulates the evolution of the master equation using time evolution methods with matrix product states, allowing for the study of biased dynamics at arbitrary times. The development of these new approaches allow for detailed characterizations of dynamical phase transitions of kinetically constrained models. This is demonstrated for the East, Fredrickson-Andersen and symmetric simple exclusion process, where the spatial and temporal finite-size scalings of their first-order phase transitions are determined. Finally, the methods are extended to two-dimensions via projected entangled-pair states.

Acknowledgements

This thesis is dedicated to my partner, Rhia, who has endlessly supported me over the last three years. I could not have done it without you.

I am extremely grateful to my supervisor, Juan P. Garrahan, for his excellent guidance and support throughout my studies. Much of what I now know is owed to you.

I would like to thank Mari Carmen Bañuls for her constant collaboration on much of the work included in this thesis, and also for the opportunity to visit her research group at MPQ, an invaluable learning experience.

Thank you to the many occupants of C18 for the interesting discussions and entertainment over the years. I am particularly grateful to Jamie and Konstantinos for many insightful chats, and also for their careful proof reading of this document.

Finally, and most importantly, I would like to thank my cat, Mia, for making the struggles and road-blocks of scientific research the *second most* stressful thing in my life.

Contents

Acronyms

1D one dimension / one dimensional

2D two dimensions / two dimensional

ASEP asymmetric simple exclusion process

CTMC continuous-time Monte Carlo

DMRG density matrix renormalization group

DW domain wall

FA Fredrickson-Andersen

KCM kinetically constrained model

LD large deviation

MCMC Markov chain Monte Carlo

MERA multi-scale entanglement renormalization ansatz

MGF moment generating function

MPO matrix product operator

MPS matrix product state

PDF probability distribution function

PEPS projected entangled-pair state

SCGF scaled cumulant generating function

SEP simple exclusion process

SSEP symmetric simple exclusion process

SVD singular value decomposition

TN tensor network

TPS transition path sampling

TTN tree tensor network

vMPS variational matrix product states

XORFA XOR-Fredrickson-Andersen

Chapter 1

Introduction

The connection between statistical mechanics and quantum many-body problems is a fruitful one. On one hand, methods of statistical mechanics can be exploited to classically simulate and investigate quantum mechanics. A key example of this is quantum Monte Carlo algorithms, whereby the properties of the ground states [1], thermal states [2, 3] and the spectrum [4, 5] of quantum Hamiltonians can be estimated. On the other hand, quantum mechanical phenomena can be conscripted to improve the sampling methods of classical statistical mechanics [6, 7], such as exploiting quantum tunneling to overcome large barriers in the energy landscape [8]. This thesis aims to further this connection, with a focus on the connection between ideas in classical stochastic dynamics and quantum many-body problems. Specifically, it will focus on the application of *tensor networks* (TNs), e.g. Refs. [9–13], in classical stochastic lattice models. Moreover, it will show how TNs can be used to determine the statistics and fluctuations of dynamical observables.

TNs have transformed the numerical study of quantum many-body models. The first works came from White in 1992 [14], introducing the *density matrix renormalization group* (DMRG) to estimate the ground state properties of finite and infinite *one dimensional* (1D) Hamiltonians. It was later discovered the DMRG could be naturally formulated as an optimization problem on the *matrix product state* (MPS) ansatz, e.g. Refs. [15, 16]. Following these advances, it was found that MPSs could also act as a platform to simulate *real-time evolution* on timescales where the entanglement remains sufficiently small [17]. Furthermore, the same methods could be applied to higher ranking TNs, such as *matrix product operators* (MPOs), to estimate properties of the Boltzmann distribution at non-zero temperatures [18–20]. Due to the immense success of MPSs, they have become the standard for studying quantum many-body systems, providing state-of-the-art results for various models. Since the invention of MPSs, the ideas have been extended to many other architectures, such as those able of capturing long-range correlations such as *tree tensor networks* (TTNs) [21], and those capable of describing entanglement beyond area laws, such as the *multi-scale entanglement renormalization ansatz* (MERA) [22]. Furthermore, MPSs have been generalized to higher dimensions: a particular important example is the *projected entangled-pair state* (PEPS) [23], an ansatz able to capture area-law states in *two dimensions* (2D).

Naturally, TNs have found their use in other areas of science, including classical systems [24–27] and machine learning [28–38]. An application of TNs which has not yet been extensively explored is in classical stochastic many-body problems, and, in particular, the study of their dynamical fluctuations. *Large deviation* (LD) theory, e.g. Refs. [39–42], is a useful framework which enables the study of the long-time statistics of dynamical observ-

ables within stochastic dynamics. For the case of stochastic lattice models, determining the long-time statistical properties is made possible through the leading eigenvectors of a deformed Markov generator which highly resembles quantum many-body Hamiltonians. By exploiting this connection, TN methods for estimating ground state properties of 1D quantum Hamiltonians can be used instead to find the leading eigenvectors of deformed Markov generators [43, 44], providing high-precision results for system sizes well beyond the usual sampling methods. However, there are still many unanswered questions. Can TNs be implemented to efficiently sample the dynamics responsible for the fluctuations of interest? How can the fluctuations at *finite times* be investigated? To what extent can TNs be used to study LDs in dimensions larger than one [45]? These are questions this thesis aims to address.

While the techniques discussed throughout this thesis are generally applicable to a broad class of stochastic lattice models, there will be a strong focus on *kinetically constrained models* (KCMs) [46–48]. KCMs were originally introduced for the dynamical modelling of structural glasses. When amorphous materials are supercooled below some glass transition temperature, there is a transition into the "glass state". While the structure of the supercooled liquids in the glass state look almost identical to that of a high-temperature liquid, there is a super-Arrhenius slowing of the dynamics [49]. Furthermore, there exist regions of space within the system with strong *dynamical* correlations. While there is no widely agreed consensus on the universal mechanisms responsible for the glass transition [50], one theoretical perspective for understanding this phenomenon is *dynamical facilitation* [51]. In particular, a region of space which is undergoing fast relaxation can facilitate the relaxation of neighbouring regions of space. This results in strong dynamical heterogeneity. That is, there is a coexistence of regions of space which are quickly relaxing and those which are not.

Perhaps one of the simplest ways to study such systems is to encode the properties of glasses into the dynamics of an Ising model with explicit kinetic constraints which capture the facilitated relaxation properties of glasses. This is at the heart of the prototypical *East* [48] and *Fredrickson-Andersen* (FA) [46] KCMs. While it is still an open question as to how one can arrive at such models from fundamental principles (e.g. through coarse graining), they are able to qualitatively reproduce some key properties of supercooled liquids, such as super-Arrhenius relaxation and dynamical heterogeneity [40]. Furthermore, these simple models allow for unique insights by means of LD theory. In particular, considering the fluctuations of the *dynamical activity* (a natural quantity to measure the level of "glassiness" in the system [40]) reveals the existence of a "dynamical phase transition" from an active-to-inactive dynamics [52]. The existence of the transition has dramatic consequences on the dynamics of interest. For KCMs — which undergo a first-order dynamical phase transition [52] — this is related to the existence of "space-time bubbles": regions of space and time which have strong dynamical correlations, demonstrating large amounts of dynamical heterogeneity in the system. Since these discoveries, ideas from LD theory and dynamical phase transitions have been used to better understand many interesting dynamical phenomena in the physical sciences, such as active matter [53, 54] and dissipative quantum many-body systems [55].

The first portion of this thesis provides a background to the physics and methods used throughout. First, an overview of stochastic dynamics is outlined in Chapter 2, with a focus on Markov Jump Processes. This section provides perspectives from both the evolution of a probability distribution under a master equation, and its unravelling onto random physical realizations of the dynamics (referred to as *trajectories*), explaining

details such as detailed balance, stationary states and efficient sampling algorithms. The object of interest will be the fluctuations of dynamical observables. In order to study these, it is necessary to consider the probability distribution over the observable. While this is typically difficult to calculate, it is made possible in the limit of large times where ideas and methodology from LD theory can be applied. It is explained how the LD statistics can be retrieved from spectral methods and from path sampling approaches. Subsequently, the two classes of models studied in this thesis are introduced: KCMs [46–48] and *simple exclussion processes* (SEPs) [56]. The former are used in the modelling of structural glasses (as described above). The latter describes the non-equilibrium transport of particles on a discrete lattice [57]. The dynamical observable typically of interest here is the particle current [44, 45, 58–63]. It is often the case that cooperative and collective behaviour can lead to dynamics which is not well described by the average. The explanation of this behaviour from the LD perspective is similar to that of KCMs. Also of interest is the dynamical activity [41, 60, 63, 64], which will be considered for the *symmetric simple exclusion process* (SSEP) in this thesis. As is the case for KCMs, evidence is provided for dynamical phase transitions at equilibrium, indicating large amounts of dynamical heterogeneity.

Chapter 3 then takes a detour to introduce the concept of TN states. It first introduces MPSs, explaining how they can be used to approximate vectors in many-body 1D system. It also explains how to *efficiently* contract MPSs, and optimize the MPS by variational means for approximating extremal eigenstates, and also simulating time evolution. Finally, a brief discussion is provided for extending these ideas to 2D with PEPS.

The next portion of the thesis is devoted to studying the dynamics and fluctuations of two KCMs. The first model introduced is the *XOR-Fredrickson-Andersen* (XORFA) model [65], motivated by Rydberg atoms in their "anti-blockade" regime [66–76], with a kinetic constraint shared by many other classical [77] and quantum [78–82] studies. As is the case for the closely related FA [46] and East [48] models, this model exhibits slow and glassy dynamics, which is shown through theoretical arguments and Monte Carlo simulations. It is also explained how moving to the *domain wall* (DW) picture allows a SEP representation of the model. The second model considered is the *Fredkin model* [83], inspired by quantum Fredkin spin chains [84–98]. Through comparisons to the ground state of the quantum model, it can be shown that the stationary state of the model displays an interesting transition between three different phases. It is again demonstrated that the model has glassy dynamics with slow relaxation properties. By means of LD theory and MPSs, the fluctuations of the dynamical activity in the long-time limit are calculated. The fluctuations in one of the stationary states of the Fredkin model displays evidence for a hierarchy of dynamical phase transitions close to equilibrium, providing a theoretical perspective on the observed metastability within the dynamics. Each of these works provide a further understanding to how kinetic constraints of varying strength can lead to the broadening of the probability distribution function of dynamical observables, and have links to other known systems of interest.

The final portion of the thesis aims to advance the algorithmic side of calculating dynamical fluctuations and rare event sampling through the use of TNs. The first of these works [99]—presented in Chapter 6—shows how the results of variational MPS can be used to implement a dynamics capable of generating the trajectories which are responsible for the fluctuations, but are not typical at equilibrium. In contrast to the MPS, which provides only the time-averaged properties of the dynamics, having access to the full dynamics associated with the fluctuations allows one to calculate arbitrary dynamical

properties of the fluctuations. Furthermore, by incorporating known path sampling techniques, this dynamics can be used to sample fluctuations at finite time events. This was further demonstrated with the works presented in Chapter 7, where time-evolved MPSs are used to calculate the finite-time canonical statistics of dynamical observables [100]. This approach is more optimal than the sampling approach due to the fact it evolves the entire ensemble of trajectories simultaneously, at the cost of systematic MPS error. Nevertheless, the results of this method can be used in tandem with the previous to correct on these errors, providing a reliable and efficient way to sample finite time events. Indeed, the investigation of finite time events, which were previously difficult to access, provides an interesting avenue for future exploration. Finally, the ideas and methods presented throughout this thesis are generalized to 2D with PEPS [101] in Chapter 8.

Chapter 2

Stochastic dynamics & fluctuations

Dynamical and complex processes in nature often have too many degrees of freedom to allow for an exact modelling. If just one of these degrees of freedom are slightly modified, the system can be in a vastly different state than expected at some later time. One approach to studying such systems is to model them as *stochastic* (random) dynamics [102]. This dynamics makes use of the known information in the system while assuming the incomprehensible degrees of freedom to be white noise. It is then possible to model the system by a probability distribution through which we can understand the average properties and their fluctuations. In the case of discrete systems, their dynamics are often modelled by a Markov Jump Processes, whereby the system will randomly and instantaneously transition between the discrete states of the system. The aim of this chapter is to give an introduction to Markov Jump Processes, providing perspectives from both the evolution of a probability vector, and random realizations of the dynamics.

2.1 Markov Jump Processes

Consider a finite and discrete space made up of the set of configurations, $\{x\}$. Now suppose that at some time t, the system is in the configuration (or state) x. Then the system is able to randomly and instantaneously transition into some other state $y \neq x$ with the transition rate $w_{x\to y}$. While the transitions are random, the transition rates define the average frequency at which transitions can occur, i.e. the times of the transitions are drawn from the exponential distribution $P_{x\to y}(t) = w_{x\to y}e^{-w_{x\to y}t}$. For simplicity, it is assumed these rates to be independent of time. The probability vector at some time t is then described by

$$|P_t\rangle = \sum_x P_t(x)\,|x\rangle\,, \tag{2.1}$$

where $P_t(x)$ is the probability that the system is in state x at time t, and $|x\rangle$ the Dirac bra-ket notation. Furthermore, the normalization $\sum_x P_t(x) = 1$ is enforced, which can be written as $\langle -|P_t\rangle = 1$ in Dirac notation, where $\langle -| = \sum_x \langle x|$ is the *flat state*. The evolution of the dynamics is determined by the *master equation*,

$$\frac{dP_t(x)}{dt} = \sum_{y\neq x} \left[w_{y\to x}P_t(y) - w_{x\to y}P_t(x)\right]. \tag{2.2}$$

The first term in the sum describes the increase in probability via some other configuration y transitioning to x, and the second term describes the loss in probability by the

transitioning away from state x to y. A convenient way of writing Eq. (2.2) is to define the *stochastic generator*

$$\mathbb{W} = \sum_x \sum_{y \neq x} w_{y \to x} \left|x\right\rangle \left\langle y\right| - \sum_x R_x \left|x\right\rangle \left\langle x\right|, \tag{2.3}$$

where $R_x = \sum_{y \neq x} w_{x \to y}$ is referred to as the *escape rate* from configuration x. This matrix representation then allows one to write the master equation as

$$\frac{d}{dt} \left|P_t\right\rangle = \mathbb{W} \left|P_t\right\rangle. \tag{2.4}$$

It then follows that the evolution of the probability vector is exponential with respect to the stochastic generator,

$$\left|P_{t+\tau}\right\rangle = e^{\tau \mathbb{W}} \left|P_t\right\rangle, \tag{2.5}$$

where $\tau > 0$. Furthermore, it can be shown that $\left\langle -\right| \mathbb{W} = 0$ and thus the normalization is preserved under time evolution,

$$\left\langle -|P_{t+\tau}\right\rangle = \left\langle -|e^{\tau \mathbb{W}}|P_t\right\rangle = \left\langle -|P_t\right\rangle. \tag{2.6}$$

2.1.1 Steady state solutions & detailed balance

Now further assume that the state space is irreducible. That is, for all configurations x_0 and x_k, there exists a path $x_0 \to x_1 \to \cdots \to x_k$ with $\omega_{x_i \to x_{i+1}} \neq 0$ for some $k > 0$. Then by the Perron-Frobenius theorem, there exists a unique steady state solution

$$\left|P_{\text{ss}}\right\rangle = \lim_{t \to \infty} e^{t \mathbb{W}} \left|P\right\rangle, \tag{2.7}$$

for all possible initial distributions, $\left|P\right\rangle$ [103]. [1] This steady state annihilates the stochastic generator $\mathbb{W} \left|P_{\text{ss}}\right\rangle = 0$. Furthermore, the dynamics described by $\mathbb{W}$ is said to obey *detailed balance* if and only if the condition

$$w_{x \to y} P_{\text{ss}}(x) = w_{y \to x} P_{\text{ss}}(y), \; \forall \, x, y \tag{2.8}$$

is satisfied [104], with $P_{\text{ss}}(x) = \left\langle x|P_{\text{ss}}\right\rangle$. Equivalently, for any path $x_1 \to x_2 \to \cdots \to x_n \to x_1$, detailed balance is obeyed if the product of transition rates is equal to the reverse path [103],

$$w_{x_1 \to x_2} w_{x_2 \to x_3} \cdots w_{x_n \to x_1} = w_{x_1 \to x_n} w_{x_n \to x_{n-1}} \cdots w_{x_2 \to x_1}. \tag{2.9}$$

If Eqs. (2.8-2.9) are satisfied, then $\left|P_{\text{ss}}\right\rangle$ is often referred to as an *equilibrium steady-state*, and can be related to some Boltzmann distribution through the expression

$$\frac{w_{x \to y}}{w_{y \to x}} = e^{-(\epsilon_y - \epsilon_x)/k_B T}, \tag{2.10}$$

where ϵ_x is an energy for state x, k_B is Boltzmann's constant and T is a temperature. If Eq. (2.8) is not satisfied, then $\left|P_{\text{ss}}\right\rangle$ is said to be a *non-equilibrium steady state*. This thesis will only consider models which obey detailed balance, and can thus be transformed onto a Hermitian matrix through the similarity transformation [52]

$$\mathbb{H} = \mathbb{P}^{-1/2} \mathbb{W} \mathbb{P}^{1/2}, \tag{2.11}$$

where $\mathbb{P}^{1/2} = \sum_x \sqrt{P_{\text{ss}}(x)} \left|x\right\rangle \left\langle x\right|$ and $\mathbb{P}^{-1/2}$ is the inverse. This will be a useful trick used throughout, as the expectation value of Eq. (2.11) is bounded by a Raleigh-Ritz principle, allowing for variational methods to be faithfully applied.

[1] This assumes that the initial probability distribution is defined *only* on the irreducible space.

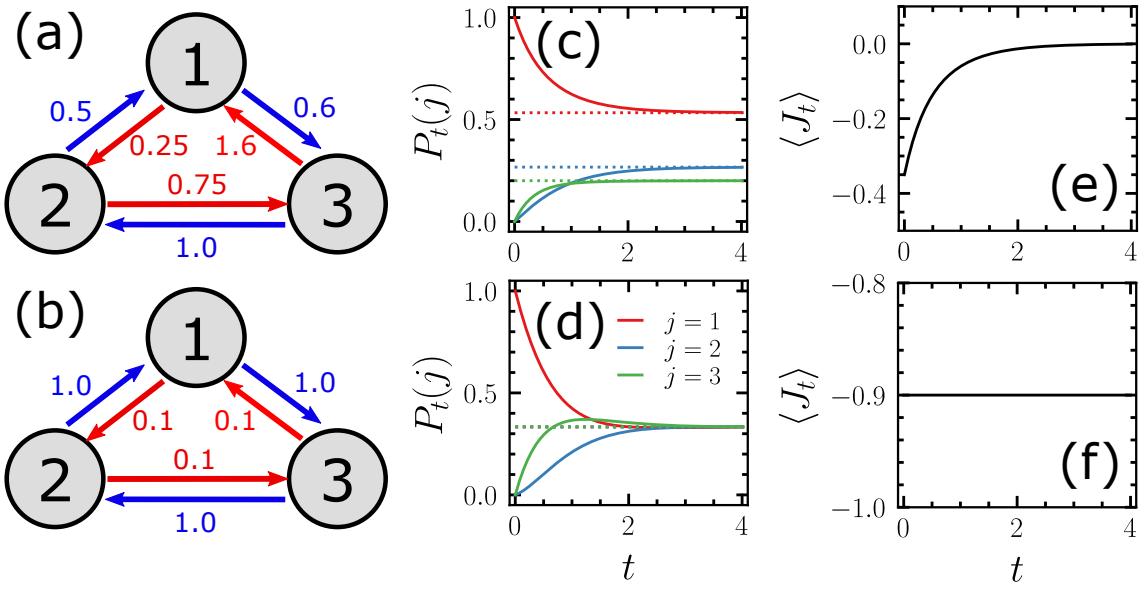


Figure 2.1: **Three state system.** A stochastic dynamics over a three state system with the transition rates chosen such that (a) the system obeys detailed balance and (b) the system does not obey detailed balance. The evolution of the probability vector elements, $P_t(j)$, are shown in (c, d) respectively. The dashed line shows their steady state values. The probability current, $\langle J_t \rangle$, is shown for each in (e, f). For the first system, the value eventually diminishes to zero as required by detailed balance. On the contrary, the second system maintains a constant current throughout.

Example: Three state system

As a simple example, consider a three state system with state space $\{|1\rangle, |2\rangle, |3\rangle\}$ and all-to-all transitions, as demonstrated in Figs. 2.1(a, b) for a system which obeys detailed balance and a system which does not, respectively. The stochastic generator for this system is

$$\mathbb{W} = \begin{bmatrix} -w_{1\to 2} - w_{1\to 3} & w_{2\to 1} & w_{3\to 1} \\ w_{1\to 2} & -w_{2\to 1} - w_{2\to 3} & w_{3\to 2} \\ w_{1\to 3} & w_{2\to 3} & -w_{3\to 1} - w_{3\to 2} \end{bmatrix}. \tag{2.12}$$

It is then easy to verify that $\mathbb{W}$ annihilates the steady-state $|P_{\text{ss}}\rangle$

$$|P_{\text{ss}}\rangle = \frac{1}{\mathcal{N}} \begin{bmatrix} w_{2\to 3}w_{3\to 1} + w_{2\to 1}w_{3\to 1} + w_{2\to 1}w_{3\to 2} \\ w_{1\to 2}w_{3\to 1} + w_{1\to 2}w_{3\to 2} + w_{1\to 3}w_{3\to 2} \\ w_{1\to 2}w_{2\to 3} + w_{1\to 3}w_{2\to 3} + w_{1\to 3}w_{2\to 1} \end{bmatrix}, \tag{2.13}$$

where $\mathcal{N}$ is the normalization factor to ensure $\langle -|P_{\text{ss}}\rangle = 1$. Now suppose for both cases, the system is initiated with $|P_0\rangle = |1\rangle$. Then the evolution of the probability vector, $|P_t\rangle$, is determined by the master equation Eq. (2.2). This can be determined numerically and is shown for both systems in Figs. 2.1(c, d). In both cases, the system relaxes to their steady state values shown by the dotted lines. The average current at some time t can be determined through $\langle J_t \rangle = \langle -|\mathbb{J}|P_t\rangle$, where

$$\mathbb{J} = \begin{bmatrix} 0 & -w_{2\to 1} & w_{3\to 1} \\ w_{1\to 2} & 0 & -w_{3\to 2} \\ -w_{1\to 3} & w_{2\to 3} & 0 \end{bmatrix} \tag{2.14}$$

is the operator which measures the net rate of transition in the direction $1 \to 2 \to 3 \to 1$. This is shown in Figs. 2.1(e, f) for both systems. The system which obeys detailed balance starts with a net current due to its out-of-equilibrium initial state, but eventually relaxes to a state with no net current. On the contrary, the system which does not obey detailed balance maintains a net current at all times.

2.1.2 Trajectories

The description for continuous-time Markov Jump Processes has so far been presented in terms of the probability vector and its evolution under the master equation. However, an alternative description can be provided in terms of *trajectory ensembles* of random realizations of the dynamics, where the properties are determined through the trajectory average. Consider some dynamics described by the stochastic generator $\mathbb{W}$ and the initial probability distribution $|P\rangle$. The partition function for the dynamics with some time t is then calculated by

$$Z_t = \langle -|e^{t\mathbb{W}}|P\rangle \,. \tag{2.15}$$

Notice that while the value of Eq. (2.15) is trivial ($Z_t = 1$), this is an important object to consider as it encodes all the possible realizations of the dynamics. At this point, it is convenient to express the stochastic generator as $\mathbb{W} = \mathbb{K} - \mathbb{R}$, where $\mathbb{K}$ is the off-diagonal components (transition rates), and $\mathbb{R}$ is the matrix of diagonal components (escape rates). Then one can perform a Dyson series expansion [105] of the exponential operator

$$Z_t = \sum_{k=0}^{\infty} \int dt_1 \cdots dt_k \, \langle -|e^{-(t-t_k)\mathbb{R}} \mathbb{K} e^{-(t_k - t_{k-1})\mathbb{R}} \mathbb{K} \cdots \mathbb{K} e^{-t_1 \mathbb{R}}|P\rangle \,, \tag{2.16}$$

where the integrals are done over the limits $t \geq t_k \geq \cdots \geq t_1 \geq 0$. One can then repeatedly insert the resolution of identity, $\mathbb{I} = \sum_x |x\rangle \langle x|$, resulting in

$$Z_t = \sum_{k=0}^{\infty} \sum_{\{x_i | i=0,\cdots,k\}} \int dt_1 \cdots dt_k \, \langle -|e^{-(t-t_k)\mathbb{R}}|x_k\rangle \langle x_k|\mathbb{K} e^{-(t_k - t_{k-1})\mathbb{R}}|x_{k-1}\rangle \cdots \\ \langle x_1|\mathbb{K} e^{-t_1 \mathbb{R}}|x_0\rangle \langle x_0|P\rangle \,. \tag{2.17}$$

Each $\langle \cdot \rangle$ can now be resolved, giving

$$Z_t = \sum_{k=0}^{\infty} \sum_{\{x_i | i=0,\cdots,k\}} \int dt_1 \cdots dt_k \; e^{-(t-t_k)R_{x_k}} \, w_{x_{k-1} \to x_k} \, e^{-(t_k - t_{k-1})R_{x_{k-1}}} \cdots \\ w_{x_0 \to x_1} \, e^{-t_1 R_{x_0}} P(x_0). \tag{2.18}$$

The paths within the path integral can be identified as *trajectories* of states, with the time records

$$\omega_t = \{(0, x_0), (t_1, x_1), \cdots, (t_k, x_k)\},$$

where the system will instantaneously transition into the state x_i at time t_i. The probability for each trajectory to occur is given by

$$\pi(\omega_t) = e^{-(t-t_k)R_{x_k}} \prod_{i=1}^{k} \left[w_{x_{i-1} \to x_i} e^{-(t_i - t_{i-1})R_{x_{i-1}}} \right] P(x_0). \tag{2.19}$$

In this description, an initial configuration x_0 is chosen from the distribution $P(x_0)$, followed by many *dwells* and *transitions*. The dwell times $\Delta t_i = t_i - t_{i-1}$ are drawn from the exponential distribution, $P(\Delta t_i) = R_{x_{i-1}} e^{-R_{x_{i-1}} \Delta t_i}$, and the transitions $x_{i-1} \to x_i$ are chosen with weightings $w_{x_{i-1} \to x_i} / R_{x_{i-1}}$. After the final transition, the configuration x_k survives with probability $P_S(t - t_k) = e^{-R_{x_k}(t - t_k)}$. It is often convenient to write the partition sum as a sum over all trajectories, $Z_t = \sum_{\omega_t} \pi(\omega_t)$, as a shorthand for the path integral.

2.1.3 Continuous-time Monte Carlo

The distribution Eq. (2.19) can be exactly and independently sampled through a *continuous-time Monte Carlo* (CTMC) algorithm. The algorithm often goes by the name Doob-Gillespie algorithm [106, 107], originally formulated for simulating models of dynamical chemical and biochemical systems. It was independently proposed by Bortz, Kalos and Lebowitz [108] as a rejection-free approach to sampling the Boltzmann distributions of Ising models (see Eq. (2.10)).

The algorithm begins by using a random number generator to randomly draw an initial configuration x from the distribution $P(x)$. The system is initialized at the time $\tau = 0$, with a total trajectory time t. Then the following steps are repeated until $\tau \geq t$:

1. Calculate the escape rate $R_x = \sum_{y \neq x} \omega_{x \to y}$.
2. Draw a uniformly distributed random number $r \in [0, 1]$. Use this to determine the transition time
$$\Delta\tau = -\frac{\log r}{R_x}, \tag{2.20}$$
and increment the trajectory time $\tau \leftarrow \tau + \Delta\tau$.
3. If $\tau > t$, set $\tau = t$ and end the simulation. Otherwise, update the system configuration to $x \leftarrow y$ with the weightings $\omega_{x \to y} / R_x$.

2.2 Dynamical fluctuations & large deviations

Consider the trajectory ensemble formulation of stochastic dynamics, which describes random realizations of the dynamics. To understand the properties of the system, one often considers some trajectory observables, $\hat{O}(\omega_t)$. These operators can take many forms, but two important ones are *static observables*, which are the time-integration of a quantity which depends only on the system configuration at some time, $\hat{O}(\omega_t) = \int_0^t d\tau\, \hat{O}(x(\tau))$, and *dynamical observables*, $\hat{K}(\omega_t)$, which depend on the transitions which occur in a trajectory. An important example for the latter is the *dynamical activity* [40, 52, 109, 110], a simple but rich observable which counts the number of transitions which occur. The dynamical activity quantifies the level of motion in a system and is thus the natural observable to quantify the "glassines" of a dynamics [40]. This will be considered extensively throughout this thesis.

Often, the important quantity considered is the trajectory average of an observable,

$$\langle \hat{O} \rangle = \sum_{\omega_t} \pi(\omega_t) \hat{O}(\omega_t). \tag{2.21}$$

This is typically easy to measure due to the accessibility of trajectories through CTMC sampling. However, it can frequently be the case that the average of some observable is not representative of the typical behaviour. In this case, one would have to investigate its *fluctuations*. The fluctuations of the dynamical observables are contained in its *probability distribution function* (PDF),

$$P_t(O) = \sum_{\omega_t} \pi(\omega_t)\delta[\hat{O}(\omega_t) - O], \tag{2.22}$$

where $\delta(x)$ is the delta function. For some given value, O, Eq. (2.22) is sometimes referred to as the *microcanonical* ensemble, as all trajectories without the desired value have zero contribution to the sum. Having access to the full PDF allows for the calculation of the moments of the observables, and thus allows one to learn about its fluctuations. On one hand, exactly calculating Eq. (2.22) is tricky due the path integral one has to carry out in Eq. (2.22). On the other hand, well estimating the PDF at its tails can require a large number of samples, which is increasingly costly in both the size of the configuration space, and in time.

As is the case in standard statistical mechanics [111], one can formulate an alternative *canonical* ensemble,

$$Z_t(s) = \sum_{\omega_t} \pi(\omega_t)e^{-s\hat{O}(\omega_t)}. \tag{2.23}$$

This ensemble, often called the *moment generating function* (MGF), allows all possible trajectories to contribute to the ensemble, but exponentially reweighs them by $-s\hat{O}(\omega_t)$, where s is a *conjugate variable* to the observable $\hat{O}$, which can be used to fix the ensemble average. The MGF contains essentially the same information as the PDF, and the two are related by a Laplace transform: while in theory this allows one to calculate one ensemble from the other, this can be difficult to do in practice. However, it becomes possible in the large time limit $t \to \infty$, where the Laplace transform becomes the *Legendre transform* [39].

The properties of the MGF and PDF in the long time limit $t \to \infty$ can be calculated through the framework of LDs [39]. In particular, the MGF and PDF go as exponentials in time,

$$P_t(O) \asymp e^{-t\varphi(o)}, \tag{2.24}$$

$$Z_t(s) \asymp e^{t\theta(s)}, \tag{2.25}$$

where $\varphi(o)$ is the *rate function* and $\theta(s)$ is the *scaled cumulant generating function* (SCGF), whose derivatives at $s = 0$ give the cumulants of the observable. Here the symbol $\asymp$ means that the functions on the left-hand side will asymptotically tend towards the right-hand side in the large time limit. The quantity $o = O/t$ is the time-averaged value of the observable. The two LD functions are connected by a Legendre transform,

$$\theta(s) = -\min_o \left[so + \varphi(o)\right]. \tag{2.26}$$

In practice, one typically aims to calculate the SCGF and then determine the rate function through Eq. (2.26). This is due to the convenience of the MGF formulation: if sampling approaches are used, all trajectories which contribute with non-zero probability in the original dynamics also contribute to the MGF, whereas in the PDF, contributions only come from the trajectories with the desired observable value. Furthermore, as explained below, the MGF admits a master equation calculation, which can be exploited to determine the SCGF.

2.2.1 Spectral approaches to the master equation

It is now demonstrated how spectral considerations allow for the calculation of the SCGF. This thesis will focus only on dynamical observables which are cumulative in the trajectory transitions, $\hat{K}(\omega_t) = \sum_{i=1}^{K} K(x_{i-1}, x_i)$, where $K(x_{i-1}, x_i)$ incorporates dependence on the types of transition. For the case of the dynamical activity, $K(x_{i-1}, x_i) = 1$ for all transitions. Generalizations to time-integrated static observables as defined earlier are obvious, but for brevity will not be discussed here. Notice that if one was to expand the MGF, Eq. (2.23), to retrieve the path integral formulation, then for every transition that occurs, there is also a factor of $e^{-sK(x_{i-1},x_i)}$. Thus it is possible to group this factor with the transition probability $\omega_{x_{i-1}\to x_i}$, and reverse the steps of the Dyson series expansion. It is at this point the *tilted* Markov generator (sometimes referred to as a biased or deformed generator) is defined,

$$\mathbb{W}_s = \sum_x \sum_{y\neq x} e^{-sK(x,y)} w_{x\to y} \left|y\right\rangle \left\langle x\right| - \sum_x R(x) \left|x\right\rangle \left\langle x\right|. \tag{2.27}$$

The MGF can then be calculated through the evolution of a probability vector under the tilted Markov generator,

$$Z_t(s) = \langle -| e^{t\mathbb{W}_s} |P\rangle . \tag{2.28}$$

While this looks equivalent to the previous considerations of the master equation, see Eq. (2.15), it is important to note that $\mathbb{W}_s$ is not a proper stochastic generator (except for at $s = 0$, where $\mathbb{W}_{s=0} = \mathbb{W}$). It does not conserve probability, $\langle -| \mathbb{W}_s \neq 0$, and its maximum eigenvalue is not zero. Nevertheless, by the Perron-Frobenius theorem, it is guaranteed that its leading eigenvalue Λ is real. Suppose that the leading left and right eigenvectors associated with Λ are $\langle l_s|$ and $|r_s\rangle$. Then it follows $\langle -| e^{t\mathbb{W}_s} \asymp e^{t\Lambda} \langle l_s|$ and $e^{t\mathbb{W}_s} |P\rangle \asymp e^{t\Lambda} |r_s\rangle$, and the leading eigenvalue of the tilted Markov generator is the SCGF, $\Lambda = \theta(s)$. Thus in the long time limit, it is not necessary to calculate Eq. (2.28), but only determine the leading eigenvalue of the tilted generator.

A number of methods have been developed to estimate the SCGF by exploiting the long-time limit. One class of methods are *variational* methods. These methods often have some form of functional approximation, $|\psi\rangle$, such as TNs [43–45, 60–62, 65, 83, 99–101, 112–114] or neural networks [30, 115–120], to approximate the left/right leading eigenvector. This is most optimal when the titled generator produces a dynamics which obeys detailed balance, and can be similarity transformed onto a Hermitian operator - this is not always possible, for example, if the original dynamics does not obey detailed balance, or if the biasing observable imposes net probability currents. When the resulting dynamics does obey detailed balance, however, one is able to define the Hermitian operator $\mathbb{H}_s = \mathbb{P}^{-1/2}\mathbb{W}_s\mathbb{P}^{1/2}$ where $\mathbb{P}^{1/2}$ is again the diagonal matrix of the square root of steady state probabilities of the *original dynamics*. This Hermitian operator then obeys the Raleigh-Ritz variationial principle, and the leading eigenvalue and its corresponding eigenvector can be retrieved through the optimization of

$$\tilde{\theta} = \frac{\langle\psi|\mathbb{H}_s|\psi\rangle}{\langle\psi|\psi\rangle} \leq \theta(s).$$

This approach will be used extensively throughout this thesis, with TNs as the functional approximation.

Another possibility for optimizing the functional approximation is to estimate the time-evolution of some initial probability vector

$$|r_s\rangle = \lim_{t\to\infty} e^{t\mathbb{W}_s} |P\rangle . \tag{2.29}$$

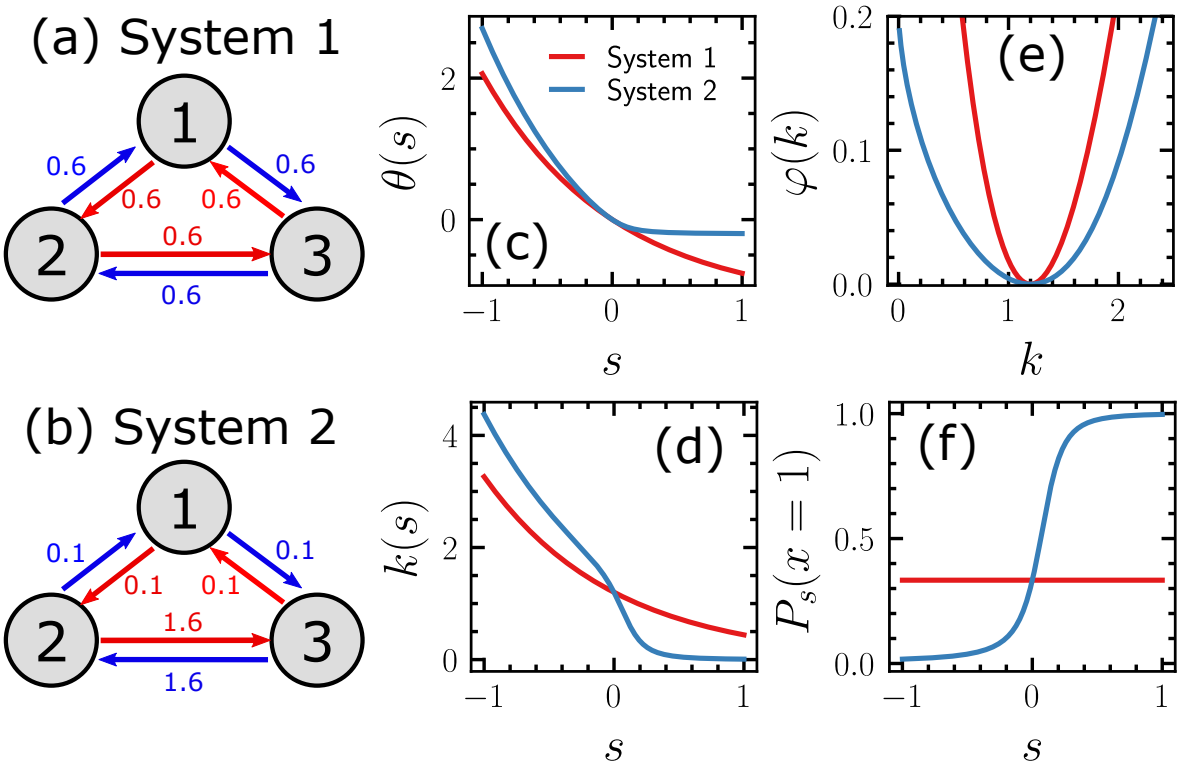


Figure 2.2: **Dynamical large deviations of the three state system.** (a, b) Three state systems with the mean activity per unit time $k(s=0) = 1.2$, and the transition rates as depicted. (c) The SCGF over a range of s for both systems. The first system has the Poisson statistics as explained in the main text, while the second system has a sharp change around $s = 0$. (d) The average dynamical activity of the trajectory ensembles $k(s) = -\theta'(s)$. (e) The rate function $\varphi(k)$ determine through the Legendre transform. (f) The probability of finding the system in the state $x = 1$ in the s-ensemble.

Indeed, this is a promising alternative for the case where detailed balance is not obeyed. For neural networks, this is made possible using natural gradient decent methods [121], and for TNs, can be done through a variety of approximate time evolution methods (e.g. Ref. [122]). This approach is adopted in Chapter 8 for PEPSs.

Example: Three state system

Again consider the three state systems shown in Figs. 2.2(a, b). In this instance, both systems have the property $w_{x\to y} = w_{y\to x}$ for all x and y, and thus the steady state solution is the equal superposition between all configurations, $|P_{\rm ss}\rangle = \frac{1}{3}(|1\rangle + |2\rangle + |3\rangle)$. Furthermore, the transition rates have been chosen such that the average dynamical activity per unit time $\bar{k} = \langle -|\mathbb{K}|P_{\rm ss}\rangle$ is equal for both systems. Despite these similarities, it should be clear that both systems will highly differ on the level of their dynamics.

Now consider the fluctuations in the dynamical activity, $K(x,y) = 1\,\forall\, x, y$. The tilted generator then takes the form

$$\mathbb{W}_s = \begin{bmatrix} -w_{1\to 2} - w_{1\to 3} & e^{-s}w_{2\to 1} & e^{-s}w_{3\to 1} \\ e^{-s}w_{1\to 2} & -w_{2\to 1} - w_{2\to 3} & e^{-s}w_{3\to 2} \\ e^{-s}w_{1\to 3} & e^{-s}w_{2\to 3} & -w_{3\to 1} - w_{3\to 2} \end{bmatrix}, \tag{2.30}$$

which can be exactly diagonalized to determine the leading left and right eigenvectors $\langle l_s|$ and $|r_s\rangle$, along with their associated eigenvalue (SCGF) $\theta(s)$. The SCGF is shown in Fig. 2.2(c) for both systems. For the first system, which is a simple Poisson process, the probability distribution goes as

$$P_t(K) = \frac{(\bar{k}t)^K e^{-\bar{k}t}}{K!}. \tag{2.31}$$

The MGF can then be calculated as

$$Z_t(s) = \sum_{K=0}^{\infty} P_t(K) e^{-sK} = e^{\bar{k}t(e^{-s}-1)}. \tag{2.32}$$

Thus the SCGF takes the exponential form $\theta(s) = \bar{k}(e^{-s} - 1)$. The second process, however, is more complicated due to its spatial heterogeneity. The results of numerical exact diagonalization is shown in Fig. 2.2(c). Notice the kink in the SCGF around $s = 0$: there is a sharp change in the gradient. Indeed, the gradient of the SCGF gives the average activity per unit time for the s-ensemble, $k(s) = -\theta'(s)$, as shown in Fig. 2.2(d). For the first system, the activity is clearly exponential in s. For the second, however, we see a sharp drop in activity around $s = 0$. Despite the fact that this sharp drop occurs over a range of s, and the true dynamics of the system only occurs at $s = 0$, this sharp change has drastic consequences on the dynamics of the system. It is a hallmark of *dynamical heterogeneity* in stochastic dynamics - in this case, a coexistence of high and low dynamical activity. This becomes more obvious when one considers the rate function. In the case of the first system, one can approximate the rate function using Stirling's formula to be $\varphi(k) \simeq k\ln(k) - k\ln(\bar{k}) - k + \bar{k}$. For the second system, this can be retrieved numerically through the Legendre transformation. This is shown for both systems in Fig. 2.2(e). Notice that the rate function for the second system is more broad than the first, showing a larger variance in activity than the simple Poisson distribution.

As was alluded to earlier, the broadening of the probability distribution is a consequence of the spatial dependence of the transition rates; the most probable trajectories for a low activity spend a significant amount of time in the state $x = 1$. This can be probed through the s-ensemble by measuring the probability of finding the system in the state x,

$$P_s(x) = \frac{\langle l_s|x\rangle \langle x|r_s\rangle}{\langle l_s|r_s\rangle}. \tag{2.33}$$

This is shown in Fig. 2.2(f) for the state $x = 1$: notice that the sharp change in probability corresponds to the sharp change in activity.

2.2.2 Auxiliary & Doob dynamics

The methods mentioned previously make use of optimization methods to estimate properties of the leading eigenvector(s) of a tilted generator. An alternative approach is to directly sample the path integral Eq. (2.23). Naively, this could be done by brute force sampling trajectories via CTMC, with the addition of importance sampling to account for the weighting $e^{-s\hat{K}(\omega_t)}$. In particular, the mean of some trajectory observable $\hat{O}(\omega_t)$ in the biased trajectory ensemble would be

$$\langle \hat{O} \rangle_s = \frac{\sum_{\omega_t} \hat{O}(\omega_t)\pi(\omega_t)e^{-s\hat{K}(\omega_t)}}{\sum_{\omega_t} \pi(\omega_t)e^{-s\hat{K}(\omega_t)}} \approx \frac{\sum_{i=1}^{N_{\rm sp}} \hat{O}(\omega_t^i)e^{-s\hat{K}(\omega_t^i)}}{\sum_{i=1}^{N_{\rm sp}} e^{-s\hat{K}(\omega_t^i)}}, \tag{2.34}$$

where N_{sp} is the number of trajectories sampled, and ω_t^i are the trajectories sampled from the original dynamics. While some success could come from this approach, it is ultimately doomed due to the exponential cost in time (and increasing cost in the size of the state space). The objective of this section is to explore alternative sampling dynamics which improve the could the sampling convergence.

Auxiliary dynamics

As is often the case in statistical mechanics, the sampling can be improved on by considering a different PDF to sample from. For continuous time stochastic dynamics, one can propose a new *auxiliary dynamics*, or *reference dynamics*. This dynamics must respect the transitions of the original dynamics, meaning all (non-)zero transition rates must also be (non-)zero in this new dynamics. Other than this restriction, there is complete freedom in the choosing of transition rates,

$$w^{\text{ref}}_{x\to y} = f_s(x,y)\, w_{x\to y}, \tag{2.35}$$

where $f_s(x,y) > 0$. The escape rate for this reference dynamics then goes as

$$R^{\text{ref}}_x = \sum_{y\neq x} f_s(x,y)\, w_{x\to y}. \tag{2.36}$$

As was the case for sampling with the original dynamics, the reference dynamics might have inconsistencies with the target dynamics, which must be accounted for with importance sampling. [2] The estimation of the mean of the observable $\hat{O}$ now goes as

$$\langle \hat{O} \rangle_s = \frac{\sum_{\omega_t} \pi^{\text{ref}}(\omega_t) \frac{\pi(\omega_t) e^{-s\hat{K}(\omega_t)}}{\pi^{\text{ref}}(\omega_t)} \hat{O}(\omega_t)}{\sum_{\omega_t} \pi^{\text{ref}}(\omega_t) \frac{\pi(\omega_t) e^{-s\hat{K}(\omega_t)}}{\pi^{\text{ref}}(\omega_t)}} \equiv \frac{\sum_{\omega_t} \pi^{\text{ref}}(\omega_t) g(\omega_t) \hat{O}(\omega_t)}{\sum_{\omega_t} \pi^{\text{ref}}(\omega_t) g(\omega_t)}, \tag{2.37}$$

where $\pi^{\text{ref}}(\omega_t)$ is the probability that the trajectory ω_t was sampled from the reference dynamics. The factor $g(\omega_t)$ accounts for the ratio of probabilities,

$$g(\omega_t) = \frac{\pi(\omega_t) e^{-s\hat{K}(\omega_t)}}{\pi^{\text{ref}}(\omega_t)} = \frac{e^{-s\hat{K}(\omega_t)} e^{-\int_0^t dt' R(x(t')) - R^{\text{ref}}(x(t'))}}{\prod_{i=1}^k f(x_{i-1}, x_i)} \frac{P(x_0)}{P^{\text{ref}}(x_0)}, \tag{2.38}$$

where $R(x(t'))$ and $R^{\text{ref}}(x(t'))$ return the escape rates R_x and R^{ref}_x for the configuration x of the system at some time t'. Notice that the reference dynamics also allows for the possibility of an alternative initial distribution, $P^{\text{ref}}(x_0)$, to sample from. Then the observable can be estimated from a finite number of trajectories sampled from the reference dynamics,

$$\langle \hat{O} \rangle_s \approx \frac{\sum_{i=1}^{N_{\text{sp}}} g(\omega_t) \hat{O}(\omega_t)}{\sum_{i=1}^{N_{\text{sp}}} g(\omega_t)}. \tag{2.39}$$

To gain any benefit from the auxiliary dynamics, a smart choice of $f(x,y)$ is required. [3] Finding this choice is difficult, and is usually inspired by some physical knowledge of the system [123–126] or the optimization of "control forces" [115, 127, 128].

[2] More precisely, the reference dynamics might not generate trajectories with the same probability as the target dynamics.

[3] One requirement for a good choice of $f(x,y)$ would be for the variance of $g(\omega_t)$ to be small.

Doob dynamics

While the tilted generator is the object required to generate the desired dynamics, it cannot be used to directly sample trajectories due to the fact that the dynamics is unnormalized (except for the case $s = 0$). Nevertheless, it is is possible to define a proper stochastic dynamics which generates trajectories with the same probabilities as the tilted generator, $\pi_s(\omega_t) = \pi(\omega_t)e^{-s\hat{K}(\omega_t)}/Z_t(s)$. In this sense, this dynamics is the most optimal auxiliary dynamics. The dynamics is defined by its transition rates, which can be calculated using Eq. (2.27). Suppose at some time $0 \leq \tau \leq t$, the system is in some configuration x. Given this fact, the instantaneous transition rate onto some configuration y is calculated as the probability that the system transitions from $x \to y$, divided by the probability that the system is in configuration x (to account for the conditional probability),

$$\begin{aligned}\tilde{w}_{x\to y}(\tau) &= \frac{\langle -|e^{(t-\tau)\mathbb{W}_s}|y\rangle \langle x|e^{\tau\mathbb{W}_s}|P\rangle}{\langle -|e^{(t-\tau)\mathbb{W}_s}|x\rangle \langle x|e^{\tau\mathbb{W}_s}|P\rangle} e^{-sK(x,y)}\, w_{x\to y} \\ &= \frac{\langle -|e^{(t-\tau)\mathbb{W}_s}|y\rangle}{\langle -|e^{(t-\tau)\mathbb{W}_s}|x\rangle} e^{-sK(x,y)}\, w_{x\to y},\end{aligned} \tag{2.40}$$

where the factor $e^{-sK(x,y)}\, w_{x\to y}$ comes from the off-diagonal transition rate in the tilted generator. This dynamics is sometimes referred to as the *Doob dynamics* [55, 126, 129–135]. While Eq. (2.40) gives the optimal transition rate for all times, it is not always useful for a variety of reasons, the predominant one being it requires one to carry out the explicit time-evolution of the master equation (this is made clear if the numerator/denominator in Eq. (2.40) is compared to Eq. (2.28) with the initial distribution $|y\rangle$). These time-dependant dynamics are dealt with in Chapter 7 by simulating the evolution of the master equation with TNs.

It is at this point the results of the previous sections can be applied by considering the large time limits $t \to \infty$ and $t - \tau \to \infty$ (only considering times sufficiently far from the time-edges of the trajectories). The transition rates, Eq. (2.40), then go as

$$\tilde{w}_{x\to y} \asymp \frac{\langle l_s|y\rangle}{\langle l_s|x\rangle} e^{-sK(x,y)}\, w_{x\to y} \equiv \frac{l_s(y)}{l_s(x)} e^{-sK(x,y)}\, w_{x\to y}. \tag{2.41}$$

This long-time limit Doob dynamics is much simpler than the true Doob dynamics due to the fact one must only calculate the leading left eigenvector of the tilted generator. Furthermore, it is time independent, meaning trajectories can be sampled from CTMC in its standard form. It is easy to show that the escape rate for the Doob dynamics is related to the escape rate of the original dynamics by the expression

$$\tilde{R}_x \equiv \sum_{y\neq x} \tilde{w}_{x\to y} = R_x + \theta(s). \tag{2.42}$$

Together, the transition rates and escape rates of the true biased dynamics can be compactly written as the *generalized Doob transformation*,

$$\tilde{\mathbb{W}}_s = \mathbb{L}\left[\mathbb{W}_s - \theta(s)\right]\mathbb{L}^{-1}, \tag{2.43}$$

where $\mathbb{L}$ is the maximal left eigenvector, $\langle l_s|$, as a diagonal matrix. It is easy to verify that the leading eigenvalue of $\tilde{\mathbb{W}}_s$ is zero with the eigenvectors $\langle -|$ and $|\tilde{P}\rangle = \sum_x \sqrt{\langle l_s|x\rangle \langle x|r_s\rangle}\, |x\rangle$. Thus it follows that Eq. (2.43) is a proper stochastic matrix.

The Doob transition rates Eq. (2.41) motivate the auxiliary dynamics with

$$w^{\text{ref}}_{x\to y} = e^{-sK(x,y)}\frac{l^{\text{ref}}_s(y)}{l^{\text{ref}}_s(x)}w_{x\to y}. \tag{2.44}$$

While this restricts the possibility of allowed dynamics, it naturally incorporates the bias $e^{-sK(x,y)}$. Indeed, the better $l^{\text{ref}}_s(x)$ estimates the left eigenvector, $l_s(x)$, the better the dynamics will be (for large times). The probability that some trajectory ω_t is generated under Eq. (2.44) us

$$\begin{aligned}\pi^{\text{ref}}(\omega_t) &= P^{\text{ref}}(x_0)e^{-\int_0^t d\tau\, R^{\text{ref}}(x(\tau))}\prod_{i=1}^{K} w^{\text{ref}}_{x_{i-1}\to x_i} \\ &= P^{\text{ref}}(x_0)e^{-\int_0^t d\tau\, R^{\text{ref}}(x(\tau))}\frac{l^{\text{ref}}_s(x_K)}{l^{\text{ref}}_s(x_0)}\prod_{i=1}^{K}\left[e^{-sK(x_{i-1},x_i)}w_{x\to y}\right],\end{aligned} \tag{2.45}$$

where $R^{\text{ref}}(x(\tau))$ denotes the escape rate of the reference dynamics when the system is in configuration x at the time τ, and the intermediate $l^{\text{ref}}_s(x)$ terms cancel through telescoping. Equation (2.45) can then be used to calculate the reweighting factor used in importance sampling,

$$g(\omega_t) = \frac{\pi(\omega_t)}{\pi^{\text{ref}}(\omega_t)} = \frac{l^{\text{ref}}_s(x_0)P(x_0)}{l^{\text{ref}}_s(x_K)P^{\text{ref}}(x_0)}e^{-\int_0^t d\tau\left[R(x(\tau))-R^{\text{ref}}(x(\tau))\right]}. \tag{2.46}$$

Notice that Eq. (2.46) has two contributions. The first is from the time-edges of the trajectories. These account for the fact the dynamics is (an approximation to) the infinite-time Doob dynamics, whereas the true dynamics is a finite time one. The exponential factor comes from the fact that the true Doob dynamics is only being approximated. Indeed, in the case the approximation is exact, the exponent goes as $\int_0^t d\tau\,\left[R_s(x(\tau)) - R^{\text{ref}}_s(x(\tau))\right] = -t\theta(s)$, and thus only serves as a constant.

2.2.3 Transition path sampling

Auxiliary dynamics with importance sampling in the manner described previously can often fall short if the auxiliary dynamics does not closely resemble the true dynamics. The sampling can be improved on, however, by introducing more advanced methods of trajectory sampling. The approach described here, known as *transition path sampling* (TPS), e.g. Refs. [115, 123, 124, 136–141], can be considered the *Markov chain Monte Carlo* (MCMC) for trajectory sampling. Suppose there is some trajectory generated from some dynamics, ω_t. The objective is to perturb the trajectory, and then accept/reject it using the Metropolis rule. As is the case with standard MCMC, the choice of update is not unique, but has the requirements that it obeys detailed balance and is ergodic in trajectory space. Once an update to the trajectory has been proposed, the new trajectory is accepted with the probability

$$P(\omega_t, \omega^{\text{prop}}_t) = \min\left(\frac{g(\omega^{\text{prop}}_t)}{g(\omega_t)}, 1\right). \tag{2.47}$$

This procedure is then repeated until the required termination conditions have been met.

There are many trajectory updates one could propose. One effective and popular method of proposing trajectories for dynamics which obey detailed balance is the *shifting*

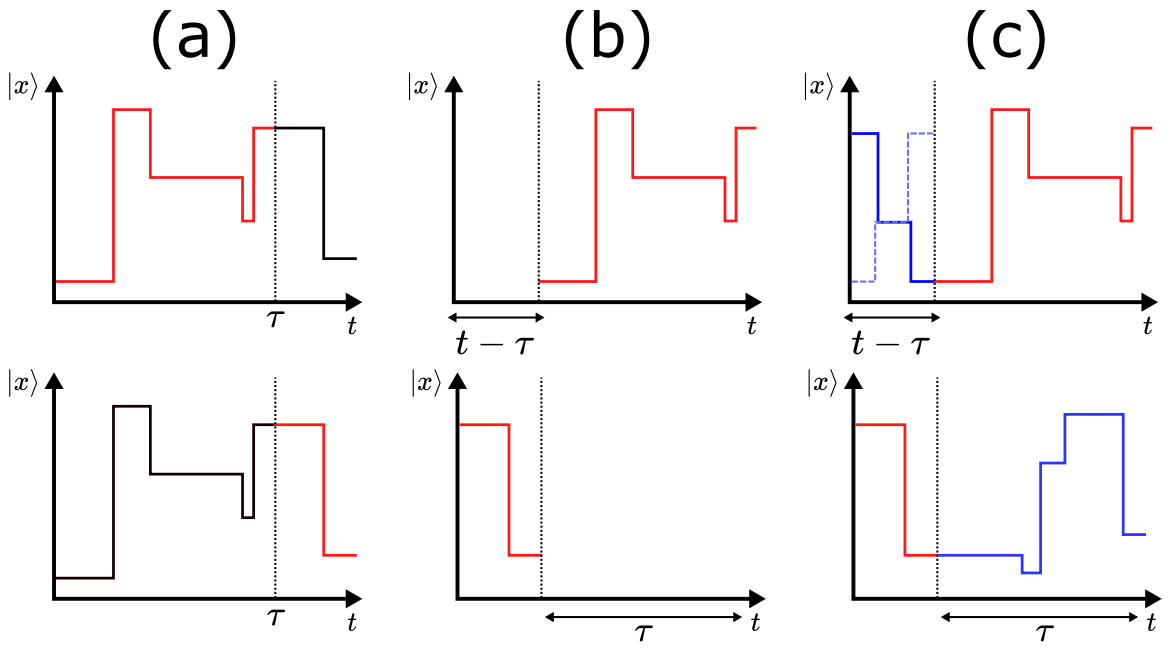


Figure 2.3: **Transition path sampling.** An illustration of the shifting method for proposing new trajectories. (a) A trajectory is split at some random time $0 \leq \tau \leq t$, and only the (top) first portion or (bottom) second portion is kept, highlighted in red. (b) The other portion of the trajectory is removed and the remaining trajectory is shifted to the (top) end or the (bottom) beginning. (c) The empty part of the trajectory is rejuvenated. For the first case (top), the trajectory is initiated with the configuration at time $t-\tau$, and is generated over the time $t-\tau$. It is then reversed. For the second case (bottom), the trajectory is initiated with the configuration at time $t-\tau$ and is generated over the time τ.

method [115, 123, 124, 138]. The first step is to uniformly sample some time to split the trajectory, $\tau \in [0, t]$, see Fig. 2.3(a). It is then randomly chosen to either keep the first part of the trajectory, $[0, \tau_{\text{split}}]$, or the second part, $[\tau_{\text{split}}, t]$, discarding the part which is not chosen. If the first part is kept, the trajectory is then shifted by time $t-\tau$. If the second part is kept, the trajectory is shifted by $-\tau$. These two options are illustrated in Fig. 2.3(b). The remaining part of the trajectory is then rejuvenated using the sampling dynamics. If the first part was kept, then the trajectory must be regenerated such that configurations match at time $t-\tau$. Since detailed balance is obeyed, the configuration at time $t-\tau$ can be used as the initial state for the rejuvenated trajectory. The dynamics are then run for a time of $t-\tau$. The trajectory is then reversed, satisfying this requirement they meet at time $t-\tau$. This process is illustrated in Fig. 2.3(c). If the second part was kept, then the trajectory can be regenerated without the need to reverse it.

The shifting method provides a way to propose trajectory updates which can be accepted or rejected, keeping the parts of the trajectory which might be desirable to the biased ensemble. The updates can be proposed over only a small portion of the trajectory, reducing the exponential cost in time. Notice it does little to hinder the cost of in the configuration space. However, for systems which can be partitioned into various subsystems (such as the lattice models considered in this thesis), it is sometimes possible to also localize the updates in space too [124, 142, 143].

TPS with an approximate Doob dynamics

The reference dynamics Eq. (2.41) can be incorporated with TPS to provide an efficient route to sampling the biased trajectories. One can easily show the reference dynamics has the stationary dynamics $P^{\text{ref}}(x) = l_s^{\text{ref}}(x)^2 P_{\text{ss}}(x)$, where $P_{\text{ss}}(x)$ is the steady state of the *original dynamics*. Assuming the initial probability distribution of the original dynamics is its stationary state, $P(x) = P_{\text{ss}}(x)$, then the reweighting factor goes as

$$g(\omega_t) = \frac{\pi(\omega_t)}{\pi^{\text{ref}}(\omega_t)} = \frac{1}{l_s^{\text{ref}}(x_0) l_s^{\text{ref}}(x_K)} e^{-\int_0^t d\tau \left[R_s(x(\tau)) - R_s^{\text{ref}}(x(\tau))\right]}. \tag{2.48}$$

This thesis will only deal with dynamics at equilibrium, but it is worth noting that it is easy to modify this approach for dynamics with different initial distributions.

2.2.4 Hybrid approaches

Three different approaches to sampling a biased dynamics have been presented, each with their own advantages and flaws. Spectral approaches can be used to directly determine the time-averaged properties of biased dynamics, but exact calculations are often limited to systems with small configuration spaces. Importance sampling methods can be used with trajectory sampling to estimate properties from the ensemble of biased trajectories. These are exponentially costly in time, but can be improved by sampling from a different auxiliary dynamics. Finally, path sampling methods, such as TPS, can be used to propose new trajectories with acceptance and rejection. While these methods slightly hinder the exponential cost of trajectory sampling, they can require many iterations to converge, and have no promise of converging onto the true dynamics. In practice, it is often the case that a hybrid of methods is used. Examples of these have already been demonstrated in the previous sections. For example, one can propose an auxiliary dynamics and use TPS to better sample it. An appropriate dynamics is often difficult to find: sometimes known physical properties of the system are exploited to construct a dynamics [123–126]. Alternatively, feedback strategies can be employed to find an optimal dynamics, where the results of path sampling methods are used to directly improve the sampling dynamics [115]. The approach used in this thesis is to use the approximate spectral results to propose an auxiliary dynamics which approximates the Doob dynamics, and then use TPS to account for errors in the approximation.

2.3 Kinetically constrained models

A significant class of lattice models considered throughout this thesis is KCMs [46–48]. Consider a lattice of N sites (or spins), which can take the binary values $n_j = 0, 1$ for $j = 1, \ldots, N$, which are labelled down/up or unexcited/excited. Excited sites have an energy cost of J, and thus the total lattice energy is $E = J \sum_i n_i$. Through standard thermodynamics, the average excitation density can be calculated to be $c = e^{J/T}/(1 + e^{J/T})$, where T is the temperature and the Boltzmann constant is set to $k_B = 1$. It is possible to define a set of stochastic lattice models which obey detailed balance, with a stationary state with the same properties as the Boltzmann distribution,

$$\mathbb{W} = \sum_{i=1}^{N} f_i \left[c\sigma_i^+ + (1-c)\sigma_i^- - c(1-n_i) - (1-c)n_i\right], \tag{2.49}$$

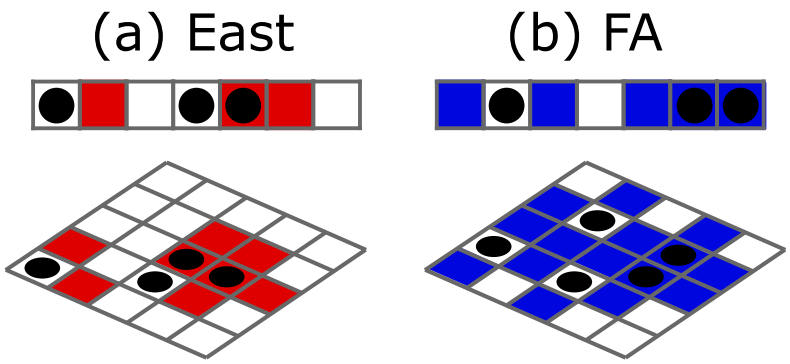


Figure 2.4: **Kinetically constrained models.** (a) The East model and (b) the FA model in 1D (top) and 2D (bottom). The circles represent excitations at sites, and the coloured tiles indicate the site is dynamically active.

where f_j is the *kinetic constraint* for the site j. This constraint has no explicit dependence on the site j. It instead depends on the lattice sites neighbouring j. Indeed, one can easily verify that the stationary state goes as

$$|\mathrm{ss}\rangle = \bigotimes_{i=1}^{N} \left[(1-c)\,|0\rangle_i + c\,|1\rangle_i\right]. \tag{2.50}$$

The averaged properties of this stationary state correspond with the thermodynamic properties of the system, irregardless of the *kinetic constraint*, f_j. [4] That is, the long-time averaged properties of the model have the required thermodynamics. However, the choice of f_j allows for vastly varying behaviour on the level of the dynamics.

Two popular choices of kinetic constraint used in the modelling of structural glasses are the one-dimensional *East* [48] and *Fredrickson-Andersen* (FA) [46] models, with the kinetic constraints

$$f_i^{\mathrm{East}} = n_{i-1}, \tag{2.51}$$
$$f_i^{\mathrm{FA}} = n_{i-1} + n_{i+1}, \tag{2.52}$$

which describes the "activation" of a local transition. For the East, the site is only allowed to flip if its neighbouring left site is excited, while the FA allows for excitations in either direction. The generalization for both models to higher dimensions is obvious - the East model becomes the "North-or-East" model in 2D [117, 144–146], where an excitation can facilitate a jump to the East or to the North, and the "North-or-East-or-Front" in 3D [145, 147]. For the FA model, a neighbouring excitation can facilitate a jump in any direction [145, 148, 149]. These constraints are illustrated in Fig. 2.4. This thesis will deal mostly with 1D lattices, but there will be some consideration of 2D lattices in Chapter 8.

One theoretical perspective of the formation of glasses is *dynamical facilitation* [51, 150]. It is easy to understand why KCMs are prototypical models of glasses if one considers an excitation to be a region of space which is undergoing fast relaxation. This

[4] This is only true if the kinetic constraint does not restrict the dynamical space the system can explore: in the case of the East and FA models, the constraint restricts the configuration $n_j = 0$ for all j. For the XORFA, which is introduced later in the thesis, there is a U(1) conservation of DWs, and the stationary properties thus look vastly different from the thermodynamics.

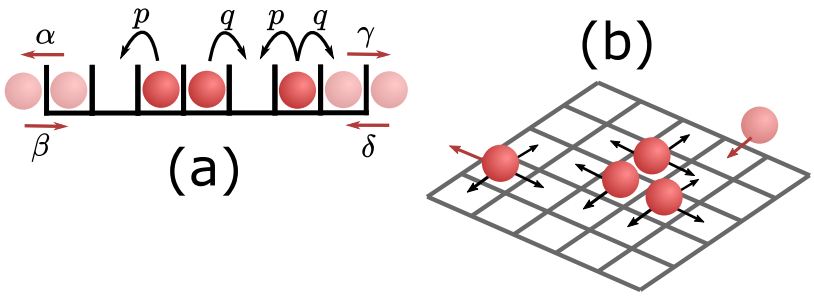


Figure 2.5: **Simple exclusion processes.** The (a) ASEP in 1D and (b) the SSEP in 2D. The balls represent particles, and the opaque balls at the boundaries illustrated particles can be injected or removed at the boundaries. The black arrows show a particle hopping to neighbouring sites, and the red arrows show the insertion or removal of particles at the boundaries. For the ASEP, the hopping rates are directional dependant, in contrast to the SSEP, which is isomorphic in space.

region is able to facilitate the relaxation of its neighbouring regions through the activation of the kinetic constraints, Eqs. (2.51, 2.52). There are many consequences of having such a constrained dynamics. One consequence is when the system is quenched from a high-to-low temperature, the relaxation time scales with the inverse temperature. Indeed, the KCMs exhibit large amounts of metastability on their road to relaxation, with increasing timescales between each metastable regime [151]. These relaxation times are super-Arrhenius with respect to the inverse temperature of the system [152]. A closely related property, relevant to structural glasses, is *dynamical heterogeneity* [150, 153, 154]: highly varying dynamics which are highly correlated in space and time.

The study of dynamical LDs have proved a useful tool for probing the properties of KCMs. On one hand, they have allowed for a thorough investigation of the probability distribution over dynamical observables through a variety of theoretical [155–159] and numerical [43, 52, 145, 157, 158, 160] techniques. This has revealed the existence of "dynamical phase transitions": the coexistence of highly active and inactive dynamics at equilibrium, indicating large amounts of dynamical heterogeneity [43, 52, 117, 145]. While this transition in the dynamics occurs in the parameter space defined by s, its very existence has severe consequences on the true dynamics (the dynamics defined by $s = 0$). For the East and FA models, the transition point between active and inactive dynamics occurs at $s \to 0^+$ for $N \to \infty$ [43]. At the point of transition, there is a coexistence between a globally active and inactive phase. However, at $s = 0$, this manifests as "space-time bubbles" [51]: the coexistence of local regions of space-time which are active or inactive. On the other hand, LD theory has been used to investigate the metastable regimes of KCMs. In the case of the East model, LD methods reveal a hierarchy of metastability close to equilibrium [43, 117, 157], related to the hierarchical ageing of the dynamics [151]. The typical timescales to reach these metastable states increase exponentially between metastable regimes; LD theory provides a way to more easily study these timescales.

2.4 Exclusion processes

Exclusion processes, which describe the movement of volume excluded particles on discrete lattices, have become a prototypical model of transport in non-equilibrium physics [57]. Like KCMs, SEPs are defined on a discrete lattice of N sites, where each site can contain a particle, $n_j = 1$, or a hole, $n_j = 0$. Each particle is able to hop to a neighbouring site *if* the target site is not already occupied. Consider the case of a 1D lattice, where particles hop to the right with the rate q, and to the left with the rate p. Also consider that the particles can be removed (inserted) at the lattice site $j = 1$ with rate α (β). Similarly, particles can also be removed (inserted) at the lattice site $j = N$ with rate γ (δ). The stochastic generator for such a process is written as

$$\mathbb{W} = \sum_{i=1}^{N-1} \Big[p\big(\sigma_i^+ \sigma_{i+1}^- - (1-n_i)n_{i+1}\big) + q\big(\sigma_i^- \sigma_{i+1}^+ - n_i(1-n_{i+1})\big) \Big] + \alpha\big(\sigma_1^- - n_1\big) + \beta\big(\sigma_1^+ - (1-n_1)\big) + \gamma\big(\sigma_N^- - n_N\big) + \delta\big(\sigma_N^+ - (1-n_N)\big). \quad (2.53)$$

The local particle currents are controlled by the parameters p and q. When $p = q$, the model is named the SSEP, with no net difference in the current between two neighbouring sites. On the other hand, when $p \neq q$, the model is named the *asymmetric simple exclusion process* (ASEP) and maintains a net current. Furthermore, the rates of insertion and removal of the particles at the boundaries control the average particle density within the lattice, and can also be used to drive currents through the boundaries [161, 162]. The ASEP is illustrated in Fig. 2.5(a).

Despite only being a simple model, the SEP allows for very rich dynamical behaviour. Regions of space can become compact in particles, restricting the transport of particles and halting any local currents. This phenomena is often referred to as *dynamical jamming* [86], and leads to rich phase diagrams in the steady steady [163], and a broadening in the probability distribution of currents [44, 45, 60–62, 164–167] and the dynamical activity [60, 63, 64, 168]. Furthermore, SEPs have already served as a famous example where analytical tensor network techniques can be applied to exactly determine the stationary state (e.g. Refs. [163, 169]).

Throughout this thesis, SEPs will be used as a toy model for non-equilibrium behaviour. In Chapters 6, 7 and 8, the activity fluctuations of the SSEP will be considered, including for the SSEP on a 2D lattice, shown in Fig. 2.5(b). Furthermore, in Chapter 4, where the XORFA model will be introduced. It will be shown that the XORFA has a mapping onto a SEP, where neighbouring particles have alternative and asymmetric hopping rates. In Chapter 5, the "Fredkin" model will be studied, which can be understood to be an ASEP with additional kinetic constraints.

Chapter 3

Tensor networks

The study of quantum many-body dynamics is an important area of research in many of the physical sciences, including condensed matter [170], quantum chemistry [171], and quantum computation [172]. However, numerical studies through standard linear algebra approaches are often limited to just a few spins due to the *curse of dimensionality*; the computational space on which calculations are performed grows exponentially with system size. However, in many cases of interest, the wavefunction (or the density matrix) can be well approximated by an ansatz with a lower dimensionality. This is the fundamental idea behind tensor networks: to decompose a large tensorial object, such as the wavefunction of a many-body system, into a product over many smaller tensors (for reviews, see Refs. [9–13]).

Typically, each degree of freedom, such as a subsystem or a spin, is allocated its own tensor which contains the essential local information. Correlations with other spins on the lattice can be obtained through the contraction of *virtual dimensions* between the tensors. While in theory one is free to choose any tensor decomposition to model their system, in practice it is convenient to choose one which matches the geometry of the system. For example, the appropriate and popular choice of TN for a 1D lattice is the MPS, shown in Fig. 3.1(a), where the shapes correspond to the tensors, the red open lines to their local physical dimensions, and the closed black lines to the virtual dimensions which are contracted over. The decomposition of this TN connects each tensor to the tensors of its neighbouring sites, allowing it to efficiently account for *local* interactions. Furthermore, the MPS is also a natural ansatz [173] to perform variational methods, such as DMRG [14], which is used in the estimation of the ground states properties of local 1D quantum many-body Hamiltonians. Alternative ansätze capable of capturing long range correlations or larger amounts of entanglement are the TTN [21] and the MERA [22, 174, 175]. These decompositions have a hierarchical structural to allow for information to be easily propagated across the chain, as illustrated for the TTN in Fig. 3.1(c). For 2D lattices, the direct generalization of the MPS is the PEPS [23], shown in Fig. 3.1(b).

The first step of any TN approach is to pick a relevant ansatz for the problem. The second step is to define a contraction scheme: a way to contract over tensors within a TN to measure some quantity. For the case of TNs which are absent of loops, such as MPSs and TTNs, there exist straightforward and efficient ways to exactly contract them. PEPSs on the other hand are harder to contract [176], and rely on approximate contraction schemes [177, 178]. Once the contraction scheme has been chosen, one is then in a position to optimize the TN. The approach taken highly depends on the objective at hand. If one wishes to study the ground state properties of a 1D lattice, for example,

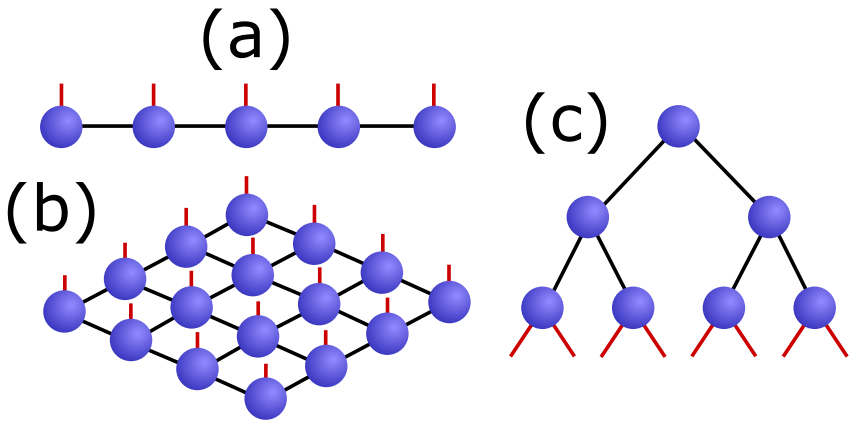


Figure 3.1: **Tensor Networks.** A schematic drawing of (a) MPS, (b) PEPS and (c) TTNs. The shapes represent tensors, with the lines the dimensions of the tensors. Black lines represent virtual dimensions, and red (open) lines represent physical dimensions.

DMRG (or variational MPS) can be applied to optimize the tensors within an MPS. This approach has proven to be highly efficient for gapped 1D local Hamiltonians [179, 180], and can also be effective for gapless systems [181]. On the other hand, if one wanted to estimate the unitary time evolution of a wavefunction, they could employ a time evolution approach [122] such as the *time-evolving block decimation* [17] method or the *time-dependant variational principle* [182]. This can also be applied to calculating the statistical canonical properties of a quantum Hamiltonian [20, 183], with some application also to the microcanonical properties [5]. Finally, once the TN is optimized, one can measure observables. This entire process is demonstrated in this chapter for MPSs, with a brief discussion on how the approach is extended to PEPSs.

While the growing popularity of TN ansätze is a consequence of their success in studying many-body quantum problems, they have also proven to be highly effective for the study of classical statistical systems [7, 27], including some application in stochastic dynamics [24–26, 169], with recent attention on studying their dynamical fluctuations [43–45, 60–62, 112–114]. This chapter will focus on providing an overview of TN methods with the intention of applying them to dynamical LDs and rare event sampling in the later chapters.

3.1 Matrix product states

Matrix product states (MPS), e.g. Refs. [9, 12, 184, 185], are a 1D TN ansatz used to represent a vector over a 1D space of N subsystems, each with a local physical dimension d. It is composed of N rank-3 tensors (except for the edge tensors, which are rank-2), with the j-th tensor denoted as $A^j_{d_j,\mu_{j-1},\mu_j}$. Some general vector on the system can be written as

$$|\psi\rangle = \sum_{\{d_j\}} \alpha_{d_1\dots d_N} |d_1 \cdots d_N\rangle \,, \tag{3.1}$$

where the coefficients $\alpha_{d_1\dots d_N}$ are retrieved by contracting over the virtual indices $\{\mu_j\}$ of the tensors,

$$\alpha_{d_1\dots d_N} = A^1_{d_1,\mu_1} A^2_{d_2,\mu_1,\mu_2} \cdots A^{N-1}_{d_{N-1},\mu_{N-2},\mu_{N-1}} A^N_{d_N,\mu_{N-1}}. \tag{3.2}$$

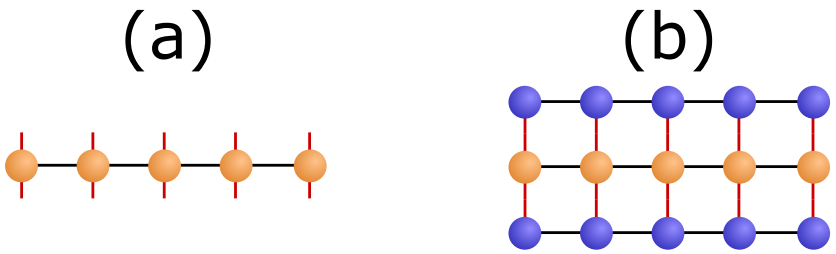


Figure 3.2: **1D Tensor Networks.** A schematic drawing of some (a) MPO, $\hat{O}$, and (b) the inner product of the MPO with some wavefunction, $\langle\psi|\hat{O}|\psi\rangle$.

This thesis will adopt the Einstein summation convention, where an index appearing twice in an equation implies a sum over the index. The virtual dimensions, μ_j — which are often referred to as the *bond dimensions* — are of dimension D, and control the complexity of the MPS. In particular, the number of parameters needed to describe the MPS scales as $O(ND^2)$, and the entanglement entropy (and thus the mutual information) between two partitions of the system is bound by $S \leq O(\log D)$ [9, 180]. To describe any arbitrary vector across the full space, one would require $D = d^{\lfloor N/2 \rfloor}$, and would gain no computational benefit from using an MPS. However, vectors with low amounts of entanglement between subsystems can be well estimated with a finite bond dimension [9]. This is the case for ground states of 1D local Hamiltonians with a non-zero gap [179, 180], which are said to obey an *area law* [186], whereby the bipartite entanglement entropy between any two neighbouring subsystems grows as $O(1)$. Vectors which satisfy this property are often said to live in a small "corner" of the exponentially big Hilbert space, and can be well estimated by a low entanglement ansatz.

The decomposition of a vector into an MPS can be easily extended to higher ranking tensorial objects. The most common extension is to a rank-2 tensor (a *matrix)*, such as a 1D Hamiltonian. In general, some matrix over a 1D system can be written as

$$\hat{O} = \sum_{d_1=1}^{d}\sum_{d_1'=1}^{d}\cdots\sum_{d_N=1}^{d}\sum_{d_N'=1}^{d} O_{d_1\ldots d_N}^{d_1'\ldots d_N'} |d_1\ldots d_N\rangle\langle d_1'\ldots d_N'|\,, \tag{3.3}$$

where the coefficients $O_{d_1\ldots d_N}^{d_1'\ldots d_N'}$ once again take the tensor decomposition

$$O_{d_1\ldots d_N}^{d_1'\ldots d_N'} = B^1_{d_1,d_1',\mu_1} B^2_{d_2,d_2',\mu_1,\mu_2}\cdots B^{N-1}_{d_{N-1},d_{N-1}',\mu_{N-2},\mu_{N-1}} B^N_{d_N,d_N',\mu_{N-1}}. \tag{3.4}$$

We call this tensor decomposition a matrix product operator (MPO) [9, 187]. An operator, $\hat{O}$, is said to be *local* if it can be exactly described by some MPO with some constant finite bond dimension D_O for any system size: this is the case for many 1D Hamiltonians of interest, and many stochastic generators [40]. It will not be explained here how to choose the tensors $B^j_{d_j,d_j',\mu_{j-1},\mu_j}$ – instructions on how to systematically construct the MPO for a 1D Hamiltonian can be found at in the literature, e.g. Refs. [9, 188]. A schematic drawing of the MPO can is shown in Fig. 3.2(a).

To perform a calculation such as calculating the expectation value $\langle\hat{O}\rangle = \langle\psi|\hat{O}|\psi\rangle$ of an operator $\hat{O}$ with respect to the vector $|\psi\rangle$, one simply needs to contract over all tensors of each object,

$$\langle\hat{O}\rangle = (\alpha_{d_1\ldots d_N})^* O_{d_1\ldots d_N}^{d_1'\ldots d_N'} \alpha_{d_1'\ldots d_N'}. \tag{3.5}$$

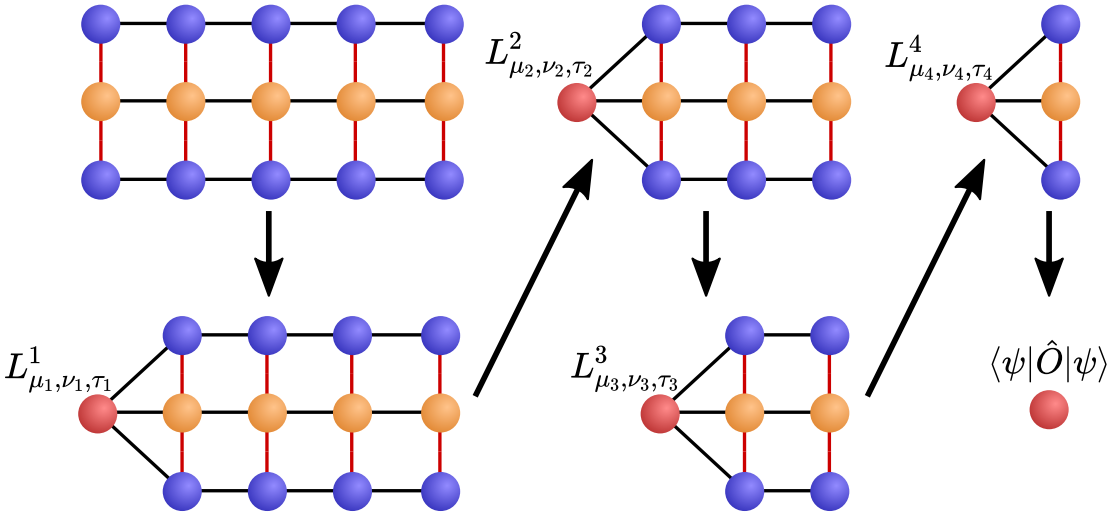


Figure 3.3: **Contracting the expectation value.** The tensor network is contracted from the left. The three leftmost tensors are contracted to give $L^1_{\mu_1,\nu_1,\tau_1}$. The next three tensors are then contracted to give $L^2_{\mu_2,\nu_2,\tau_2}$. This is repeated until all tensors have been contracted yielding the expectation value $\langle\psi|\hat{O}|\psi\rangle$.

Indeed, inserting the tensor decomposition for each of these terms makes the expression rather cumbersome. Instead, it is convenient to use the pictorial description to represent the contracted network, as shown in Fig. 3.2(b). Nevertheless, it is not immediately obvious what the optimal way to contract the network is.

3.1.1 Contracting matrix product states

With every TN ansatz, it is important to devise a scheme to efficiently contract it. For the case of 1D TNs which are absent of loops, calculating quantities such as expectation values can be done efficiently and exactly with a computational cost which scales only as a polynomial of the system size and the bond dimensions. Consider the expectation value $\hat{O} = \langle\psi|\hat{O}|\psi\rangle$. The most optimal way to contract this is from the left-most (or right-most) edge of the TN, as is demonstrated in Fig. 3.3. The three left-most tensors are contracted to give

$$L^1_{\mu_1,\nu_1,\tau_1} = (A^1_{d_1,\mu_1})^* O^1_{d_1,d'_1,\nu_1} A^1_{d'_1,\tau_1}, \tag{3.6}$$

where $\{\mu_j\}$, $\{\nu_j\}$ and $\{\tau_j\}$ are the index sets for the virtual dimensions of $\langle\psi|$, $\hat{O}$ and $|\psi\rangle$, respectively. The TN is then further contracted one column at a time,

$$L^{j+1}_{\mu_{j+1},\nu_{j+1},\tau_{j+1}} = L^j_{\mu_j,\nu_j,\tau_j} (A^{j+1}_{d_{j+1},\mu_j,\mu_{j+1}})^* O^{j+1}_{d_{j+1},d'_{j+1},\nu_j,\nu_{j+1}} A^{j+1}_{d'_{j+1},\tau_j,\tau_{j+1}}, \tag{3.7}$$

until the TN is fully contracted. This process is depicted in Fig. 3.3. Notice that at each stage of the contraction, the intermediate tensors $L^j_{\mu_j,\nu_j,\tau_j}$ have $O(D^2 D_O)$ number of parameters. It is also convenient to define the contraction of the TN from the right,

$$R^N_{\mu_{N-1},\nu_{N-1},\tau_{N-1}} = (A^N_{d_N,\mu_{N-1}})^* O^N_{d_N,d'_N,\nu_{N-1}} A^N_{d'_N,\tau_{N-1}}, \tag{3.8}$$

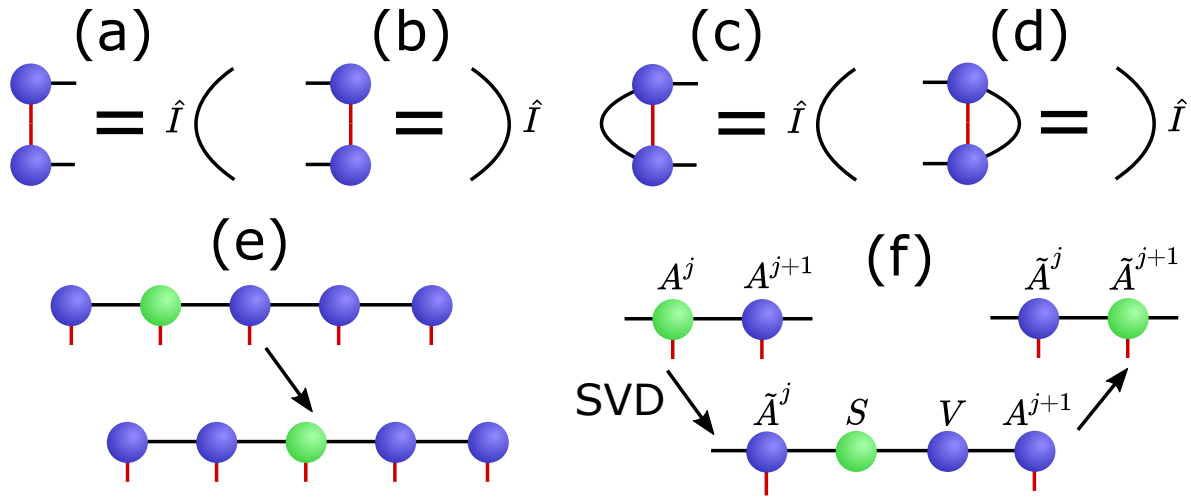


Figure 3.4: **Canonical representations of MPS.** (a, b) The tensors A^1 and A^N are orthonormal if they give the identity matrix when contracted with their conjugates. (c, d) The same for tensors A^i for $i < j$ and $i > j$ respectively, where j is the center of orthogonality of the MPS. (e) The center of orthogonality can be moved across the MPS. (f) The center is moved from A^j to A^{j+1} through a SVD. The singular matrix S and the orthonormal matrix V are contracted with the matrix A^{j+1} to give the updated matrices.

with subsequent blocks defined as

$$R^{j-1}_{\mu_{j-2},\nu_{j-2},\tau_{j-2}} = R^{j}_{\mu_{j-1},\nu_{j-1},\tau_{j-1}} (A^{j-1}_{d_N,\mu_{j-2}}, \mu_{j-1})^* O^N_{d_{j-2},d'_{j-2},\nu_{j-2},\nu_{j-1}} A^N_{d'_{j-2},\tau_{j-2},\tau_{j-1}}. \tag{3.9}$$

For contracting an expectation value it does not matter whether it is contracted from the left or the right, but it will later be convenient to calculate both when variationally optimizing some MPS.

3.1.2 Canonical representations

Consider the MPS decomposition Eq. (3.2). By defining the identity matrix as $\hat{I} = \sum_{\tau,\tau'} \delta_{\tau,\tau'}$, it is possible to write Eq. (3.2) as

$$\alpha_{d_1 \ldots d_N} = A^1_{\mu_1,d_1} \cdots A^j_{\mu_{j-1},\tau,d_j} \delta_{\tau,\mu_j} A^{j+1}_{\mu_j,\mu_{j+1},d_j} \cdots A^N_{\mu_{N-1},d_N}, \tag{3.10}$$

where the identity has been inserted between the tensors at j and $j+1$. This equivalence can be exploited by noting that the identity matrix can be expressed as $\hat{I} = \hat{U}^{-1}\hat{U}$ for some invertible matrix $\hat{U}$. In particular, one can define the transformed tensors

$$\tilde{A}^j_{d_j,\mu_{j-1},\mu_j} = A^j_{d_j,\mu_{j-1},\tau} U^{-1}_{\tau,\mu_j}, \tag{3.11}$$

$$\tilde{A}^{j+1}_{d_{j+1},\mu_j,\mu_{j+1}} = U_{\mu_j,\tau} A^{j+1}_{d_{j+1},\tau,\mu_{j+1}}, \tag{3.12}$$

as the new tensors at positions j and $j+1$. While the tensors within the MPS have changed, the coefficients in Eq. (3.1) have not. This gauge transformation allows one to move the *center of orthogonality* of the MPS [9, 184]. The first (last) tensors within an MPS are said to be in *left-canonical* (*right-canonical*) form if they respectively satisfy the

following conditions,

$$(A^1_{d_1,\mu_1})^* A^1_{d_1,\nu_1} = \delta_{\mu_1,\nu_1}, \tag{3.13}$$

$$(A^N_{d_N,\mu_{N-1}})^* A^N_{d_N,\nu_{N-1}} = \delta_{\mu_{N-1},\nu_{N-1}}, \tag{3.14}$$

or put more simply, if the contraction between the tensors and their conjugates give the identity matrix. This is shown in Figs. 3.4(a, b). This analogy can be extended to the remaining tensors in the MPS, where each can take a left- or right-canonical form if they respectively satisfy the following conditions,

$$\delta_{\mu_{i-1},\nu_{i-1}} (A^i_{d_i,\mu_{i-1},\mu_i})^* A^i_{d_i,\nu_{i-1},\nu_i} = \delta_{\mu_i,\nu_i} \tag{3.15}$$

$$\delta_{\mu_i,\nu_i} (A^i_{d_i,\mu_{i-1},\mu_i})^* A^i_{d_i,\nu_{i-1},\nu_i} = \delta_{\mu_{i-1},\nu_{i-1}}, \tag{3.16}$$

as shown in Figs. 3.4(c, d). Then it can be said that the center of orthogonality of an MPS is at site j (or the MPS is in mixed-canonical form around site j) if all tensors for $i < j$ are in left-canonical form and all tensors for $i > j$ are in right-canonical form. Having the center of orthogonality around j is useful in updating and optimizing MPSs in the later algorithms. For example, for the variational methods discussed later, having the center of orthogonality at site j allows for a huge simplification for updating the tensor A^j. On the other hand, for time evolution methods, local "gates" can be applied to tensors without the need to consider the other tensors of the MPS. Fortunately, moving the center of orthogonality in an MPS can be easily achieved through the *singular value decomposition* (SVD). For some matrix M, the SVD allows one to write $M = USV$, where U and V are orthonormal matrices, and S is a diagonal matrix with non-negative entries (the "singular values"). By remapping the tensors of the MPS onto a rectangular matrix with one dimension being the virtual bond in the direction we wish to move the center, and the other index containing all other dimensions, it is possible to use the SVD to put the tensors in left- or right-canonical form. Consider Fig. 3.4(e), which describes an MPS with the center of orthogonality at the second tensor. The center of orthogonality of the MPS can be moved to the third site by forcing the second tensor into left-canonical form. This process is illustrated in Fig. 3.4(f), where a SVD is applied at site j. The matrices S and V are then absorbed into the tensor at site $j+1$. Similarly, the center of orthogonality can be moved left by applying the SVD in the opposite direction.

3.1.3 Variational matrix product states

Now that a suitable ansatz with the appropriate contraction scheme has been specified, a suitable optimization algorithm is necessary to calculate the MPS. The first class of optimization algorithms discussed here will be *variational* algorithms, with the objective of minimizing the expectation value of an observable with respect to a vector [9, 14, 189–191]. For quantum problems, this vector is the wavefunction, and the observable is an MPO representation of a Hamiltonian. However, for the stochastic problems considered in this thesis, the object of interest is determining the SCGF through the maximal eigenvalue of tilted stochastic generators $\mathbb{W}_s$, where here the vector is the probability distribution over configurations. As described previously, for the fluctuations of interest here, it is possible to transform the tilted generator onto a Hermitian operator (which we will refer to as the Hamiltonian) through the similarity transformation $\mathbb{H}_s = -\mathbb{P}^{-1/2}\,\mathbb{W}_s\,\mathbb{P}^{1/2}$, where a minus sign has been introduced to change the problem from a maximization to a minimization.

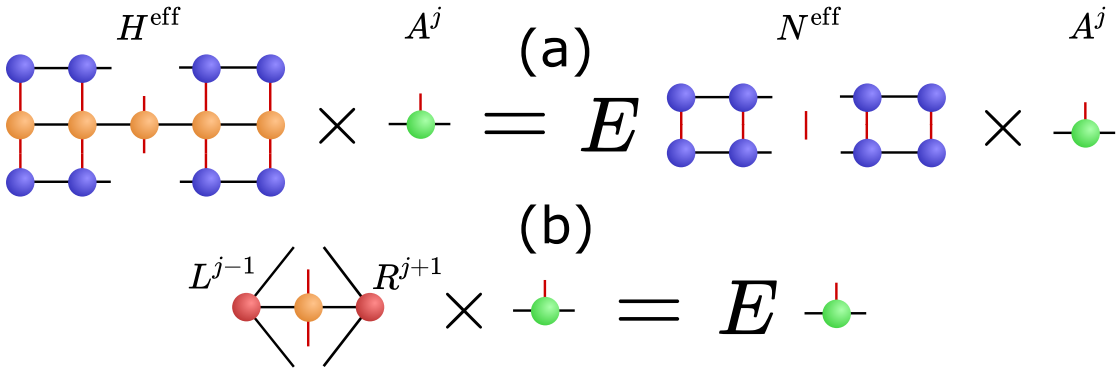


Figure 3.5: **Variational matrix product states.** (a) The eigenproblem to be solved to optimize the tensor A^j in pictorial form. (b) The effective Hamiltonian, H^{eff}, can be contracted from the relevant left L^{j-1} and right R^{j+1} respectively, where the subscript indices have been omitted. When the center of orthogonality is at site j, the effective norm becomes the identity operator $N^{\text{eff}} = \hat{I}$.

Then by the Raleigh-Ritz variational principle, it follows that the expectation of the vector with respect to the Hamiltonian is

$$E = \frac{\langle\psi|\mathbb{H}_s|\psi\rangle}{\langle\psi|\psi\rangle} \geq -\theta(s). \tag{3.17}$$

Thus by minimizing the expectation of some vector $|\psi\rangle$, it is possible to have a bound on the SCGF. For the MPS representation of $|\psi\rangle$, a strategy can be devised to sequentially update each tensor within the MPS, with each update decreasing the energy.

Suppose the tensor to be updated is the tensor at site j, A^j, where the subscripts have been dropped for brevity. To find the optimal update for the tensor while keeping all other tensors fixed, one would need to minimize the Eq. (3.17). By differentiating with respect to the tensor A^j, one finds the eigenproblem

$$H^{\text{eff}} A^j = E N^{\text{eff}} A^j, \tag{3.18}$$

where the effective Hamiltonian, H^{eff}, and the effective norm, N^{eff}, are obtained by contracting over every tensor but A^j in $\langle\psi|\mathbb{H}|\psi\rangle$ and $\langle\psi|\psi\rangle$ respectively, where again the subscripts have been omitted. These partially contracted networks are shown in Fig. 3.5(a). It is obvious Eq. (3.18) is a generalized eigenproblem if one considers A^j to be a vector, and H^{eff} and N^{eff} to be matrices. Then the optimal choice for A^j is the solution to Eq. (3.18) with the smallest eigenvalue.

The problem can be made far simpler if the center of orthogonality of the MPS is placed at j. In this instance, the effective norm N^{eff} reduces to the identity matrix, and Eq. (3.18) reduces to a standard eigenproblem. Furthermore, determining the solution to Eq. (3.18) requires an iterative eigensolver, such as the Lanczos method [192]. [1] Each iteration requires contracting H^{eff} with the current guess of A^j. Therefore it is optimal to calculate the blocks L^{j-1} and R^{j+1} beforehand, allowing them to be reused in the update procedure. The reduced equation is illustrated in Fig. 3.5(b). Once the minimal

[1] While in practice one can use a standard eigensolver, it is important to note that only the eigenvector with the smallest eigenvalue is required. Using an iterative eigensolver is necessary for an implementation with minimal cost.

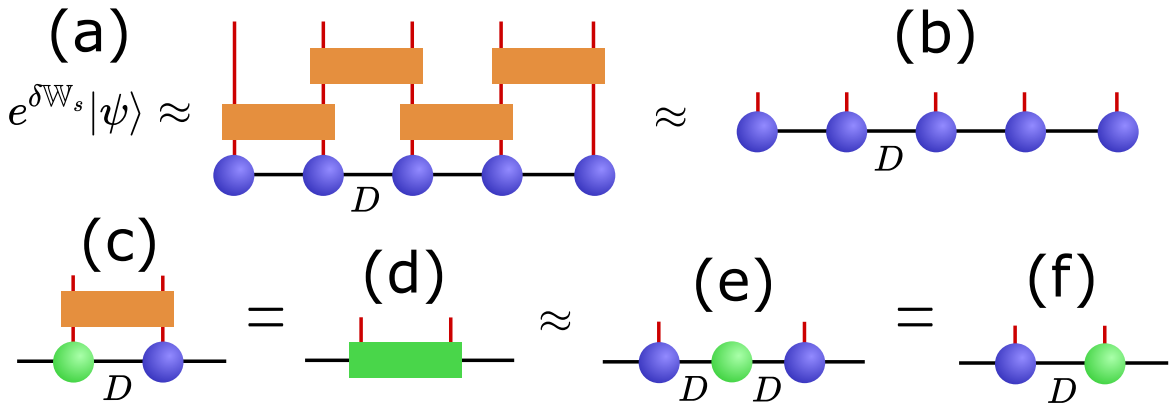


Figure 3.6: **Time-evolved matrix product states.** (a) The time evolution of an MPS for a small time step δ can be approximated by the application of local Trotter gates, shown by the orange rectangles. (b) The application of the Trotter gates can be approximately described by an MPS with the original bond dimension, D, using truncation. (c) The application of a Trotter gate on two sites, with the center of orthogonality at the left site. (d) The gate is then contracted with the two MPS tensors. (e) A truncated SVD is performed to restore the MPS representation. (f) The singular values are contracted with the rightmost tensor to move the MPS into canonical form on the right tensor.

eigenvector has been found (or well estimated up to some tolerance), it is substituted in the MPS. This procedure is conducted over each tensor in the MPS, sweeping from left-to-right and then right-to-left, until convergence in the energy is achieved. This algorithm is dubbed *variational matrix product states* (vMPS), but is often referred to as DMRG [14] due to their similarities. Indeed, vMPS can just be considered DMRG on the MPS ansatz [189].

3.1.4 Trotterized time-evolution

An alternative approach to optimizing an MPS is to use approximate time evolution methods. These days, there is a whole ensemble of methods which can be used to estimate the time evolution of an MPS [17, 122, 182, 183, 193–196]. A simple but versatile approach is the so-called Trotter approach [17, 183, 195]. At the heart of this method is the Trotter-Suzuki decomposition [197] of the time-evolution operator. In particular, suppose the objective is to time evolve an MPS by some time, t. The first step is to split the time evolution operator into a sequence of smaller operators,

$$U(t) = e^{t\mathbb{W}_s} = \prod_{i=1}^{T} e^{\delta \mathbb{W}_s} = \prod_{i=1}^{T} U(\delta), \tag{3.19}$$

where $T = t/\delta$ is the number of time steps. Now suppose the operator $\mathbb{W}_s$ is local, meaning it can be decomposed into the sum of operators over local sites. In this instance we will consider the operator to be a sum over nearest neighbour interactions, $\mathbb{W}_s = \sum_i W_{i,i+1}$, although generalizations to interactions with a larger number of sites is possible. Then, for small δ, the evolution operator can be estimated by

$$U(\delta) = e^{\delta \sum_i W_{i,i+1}} \approx e^{\delta \sum_{i\,\text{even}} W_{i,i+1}} e^{\delta \sum_{i\,\text{odd}} W_{i,i+1}} = \prod_{i\,\text{even}} e^{\delta W_{i,i+1}} \prod_{i\,\text{odd}} e^{\delta W_{i,i+1}}, \tag{3.20}$$

where the expression is inexact due to the fact that terms with even i do not commute with terms with odd i. The last equality follows from the fact that all terms in both sums commute. Notice that since each matrix exponential in Eq. (3.20) is defined over just two sites, it can be calculated exactly. These matrices are given the name "Trotter gates", and can be contracted with the MPS $|\psi\rangle$ to estimate time evolution over the small-time step δ. This is shown in Fig. 3.6(a).

The Trotter gates allow a way to estimate the time evolution operator as a TN made of local tensors, but their application to the MPS is not obvious. Indeed, contracting one of the gates with the relevant tensors within the MPS will destroy the structure of the MPS. Through an SVD, the MPS structure can be retrieved, but with a larger bond dimension between the two tensors (which grows exponentially with the number of gates applied). To maintain a finite bond dimension, it is crucial to *truncate* the tensors. This means when an SVD is applied, only the largest D singular values are kept. For this update to be optimal with respect to D, it is important that the centre of orthogonality of the MPS is at one of the tensors contracted with the gate beforehand. The process of applying the gate to the MPS is depicted in Figs. 3.6(c-f). The result of an application of all the gates is the approximately time-evolved MPS, Fig. 3.6(b).

Notice that time evolution through Trotter gates introduces a few systematic errors. The first error, referred to as the "Trotter error", is introduced through the Trotter-Suzuki decomposition of the time evolution operator. For a first-order decomposition, this error goes as $O(\delta^2)$ per time step, entailing a total error of $O(\delta t)$. Indeed, one can use higher order approximations which introduce more Trotter gates, but with a reduced error. For example, the commonly used second order Trotter decomposition has an error $O(\delta^3)$ per time step, and a total error $O(\delta^2 t)$. In practice, these errors can often be made negligible on timescales typically investigated by MPS by using a small-time step, with an increase in the computational cost, where one has to apply $O(t\delta^{-1})$ gates. The second, more dominant error, is acquired through the SVD truncation of the MPS. Each application of a Trotter gate will increase the local bond dimensions of the tensors, leading to an exponential increase in bond dimension with time. The truncation of the MPS keeps this bond dimension bounded, at the cost of loosing some information. While there is no easy way to avoid this error, it highly depends on the amount of entanglement in the vector: some instances may have manageable amounts of entanglement, whereas others might quickly spiral out of control. Approaches to deal with this for stochastic fluctuations are discussed in Chapter 7.

3.2 Projected entangled-pair states

The natural generalization of the MPS is the projected entangled-pair state (PEPS), e.g. Ref. [23]. As was the case for the MPS, each subsystem is assigned its own tensor, with virtual dimensions which connect it to the tensors of each neighbouring subsystem. However, unlike MPSs, PEPSs cannot be contracted exactly in the large system size limit, and require approximate contraction schemes. The most popular choice for finite systems is the "MPS boundary" scheme [178]. The fundamental idea is to equate the boundary of the PEPS to a 1D system, and approximate it by an MPS. The subsequent rows (or columns) in the network can then be interpreted as MPOs which act on the MPS: the MPO-MPS network can then be approximately represented by another MPS through SVDs or by variational minimization [178].

While in principle one can formulate a method to optimize the tensors in a PEPS with

respect to the minimization of some Hamiltonian, in practice this is often too expensive to do efficiently, with the approximate contraction schemes sometimes making this approach unstable [178]. Instead, the standard approach is to use Trotterized time-evolution with large enough times to converge to the steady state (or ground state for quantum problems). There are a variety of methods to achieve this task, which range from low-complexity and low-accuracy [198] to high-complexity and high-accuracy [199, 200]. While a detailed explanation of the methods used to optimize PEPS is omitted from this thesis, some basic detailed can be found in Appendix B. More detailed explanations can be found in the literature, e.g. Ref. [178], which details the many strategies to optimize PEPS, and gives insights onto improving the stability of the algorithms. These approaches are used in Chapter 8 to deal with LDs of 2D problems.

Chapter 4

Dynamics and large deviation transitions of the XOR-Fredrickson-Andersen kinetically constrained model

The following work is from the publication *"Dynamics and large deviation transitions of the XOR-Fredrickson-Andersen kinetically constrained model"* by Luke Causer, Mari Carmen Bañuls, Igor Lesanovsky and Juan P. Garrahan, published in Physical Review E 102 (5), 052132 (2020).

This work introduces a 1D kinetically constrained model in stochastic classical dynamics. The constraint is inspired by Rydberg atoms in their "anti-blockade" regime, only allowing spins to flip if their neighbouring spins are in different states. The model can be expressed by its DW representation, allowing for it to be studied as an exclusion process, with each neighbouring particle having alternating asymmetries.

The steady state properties and non-equilibrium dynamics of the model are determined by means of Monte Carlo simulations and analytical MPSs, which demonstrate the model undergoes highly non-trivial and glassy relaxation. The slow relaxation properties of the model motivates the investigation of its dynamical fluctuations through LD theory and numerical MPSs. As is the case for many other KCMs, the model exhibits a first order phase transition in the dynamical activity, indicating large amounts of dynamical heterogeneity and the existence of "space-time bubbles" in the dynamics: the coexistence of regions in space-time which display small and large amounts of activity, confirming the glassy behaviour of the model. An accurate finite-size scaling analysis is done, revealing a scaling exponent which differs from other KCMs of interest (such as the East of FA), placing this model in a different universality class. This difference in the exponent is explained by the diffusive nature of the dynamics.

Dynamics and large deviation transitions of the XOR-Fredrickson-Andersen kinetically constrained model

Luke Causer,[1,2] Igor Lesanovsky,[3,1,2] Mari Carmen Bañuls,[4,5] and Juan P. Garrahan[1,2]

[1]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

[2]*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

[3]*Institut für Theoretische Physik, Universität Tübingen, Auf der Morgenstelle 14, 72076 Tübingen, Germany*

[4]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

[5]*Munich Center for Quantum Science and Technology (MCQST), Schellingstrasse 4, D-80799 München, Germany*

(Received 24 August 2020; revised 5 November 2020; accepted 9 November 2020; published 25 November 2020)

We study a one-dimensional classical stochastic kinetically constrained model (KCM) inspired by Rydberg atoms in their "facilitated" regime, where sites can flip only if a single of their nearest neighbors is excited. We call this model "XOR-FA" to distinguish it from the standard Fredrickson-Andersen (FA) model. We describe the dynamics of the XOR-FA model, including its relation to simple exclusion processes in its domain wall representation. The interesting relaxation dynamics of the XOR-FA is related to the prominence of large dynamical fluctuations that lead to phase transitions between active and inactive dynamical phases as in other KCMs. By means of numerical tensor network methods we study in detail such transitions in the dynamical large deviation regime.

DOI: 10.1103/PhysRevE.102.052132

I. INTRODUCTION

Systems with constraints often display interesting cooperative dynamics [1–4]. This is true both in classical and quantum settings. Broadly speaking there are three classes of constrained systems. One is that of problems where state space is constrained. The canonical example is lattice coverings, for example dimers on a square lattice [5–11]. In such systems, the constrained nature of configuration space implies constraints in the allowed transitions, making both their classical and quantum dynamics very rich. A second class encompasses systems where constraints in the dynamics are emergent, such as in classical and quantum "fracton" models where the motion of certain effective excitations is severely impeded [12–16]. A third class comprises systems known as kinetically constrained models (KCMs) with explicit constraints in the allowed dynamical transitions. Here we focus on KCMs.

KCMs were first introduced [1,2] in the 1980s as models of classical glasses. The ones studied most throroughly, such as the Fredrickson-Andersen (FA) [1] and East models [3], are stochastic lattice spin systems with the interesting combination of a trivial thermodynamics and a strongly fluctuating cooperative dynamics (under appropriate conditions—typically low temperatures and/or high densities) due to the constraints. For reviews on classical KCMs, see, e.g., Refs. [17–19]. Like their classical counterparts, quantum KCMs also display complex nonequilibrium dynamics, both under closed unitary [20–22] or open dissipative [23] evolution.

Modeling dynamics via KCMs can be motivated in many different areas. For example, in classical soft matter, specifically for glasses [24,25], kinetic constraints are meant to encode the local steric interactions of dense fluids. Another application is in the context of ensembles of Rydberg atoms in optical lattices, modeled as a collection of local two-level systems (representing for each atom their ground and some high-lying Rydberg state). When driven on resonance, due to "Rydberg blockade" [26], their dynamics is subject to a kinetic constraint where an atom can change state only if all their nearest neighbors are in their ground state. In a one-dimensional lattice such constraint gives rise to the much studied PXP model [27–30], the quantum counterpart of the classical "two-spin facilitated" FA model [17].

Here we study a one-dimensional classical KCM which to our knowledge has not been considered in the past. We call it the XOR-FA model to distinguish it from the standard FA model (i.e., the "one-spin facilitated" FA model). The kinetic constraint in the XOR-FA is such that a spin can flip only if one of its nearest neighbors is in the excited state, but not if both are (which is allowed in the FA). Such condition makes the XOR-FA more constrained than the standard FA model. Conversely, the XOR-FA is less constrained than the PXP, whose transitions require the two neighboring sites to be simultaneously down. The constraint in the XOR-FA model can be motivated by Rydberg atoms in their "facilitated" (or "antiblockade") regime [31–41]: When driven out of resonance, specifically when blue-detuned, conditions can be such that an atom may change state only if a single neighbor is in the excited state, but not both.

The paper is organized as follows. In Sec. II we introduce the XOR-FA model and discuss its connection to simple exclusion processes. In Sec. III we consider the relaxation

dynamics of the model. In Sec. IV we study the dynamical large deviations by means of numerical tensor networks, and show the existence of a phase transition between active and inactive dynamical phases. In Sec. V we draw comparisons between the FA, XOR-FA, and PXP models. In Sec. VI we give our conclusions.

II. MODEL

We consider a system of binary variables $n_j = 0, 1$ (we call these states down/up or unexcited/excited) on the sites $i = 1, \ldots, N$ of a one-dimensional lattice (with boundary conditions to be specified below). Similar to other KCMs [17,42] the dynamics will be that of singe-spin flips subject to a constraint. Specifically, the allowed transitions are

$$\begin{array}{ll} 001 \to 011 & \text{rate} = c, \\ 011 \to 001 & \text{rate} = 1-c, \\ 100 \to 110 & \text{rate} = c, \\ 110 \to 100 & \text{rate} = 1-c, \end{array} \tag{1}$$

where $c \in (0, 1)$. That is, a site can flip only if both nearest neighboring sites are in different states. This means that the constraint is a boolean XOR operation on the nearest neighbors of the site that is attempting to flip. We therefore call this model the XOR-FA (short for XOR-Fredrickson-Andersen) to distinguish it from the standard Fredrickson-Andersen (FA) model, where a site can flip if either of its nearest neighbors is up, which in this nomenclature would correspond to the OR-FA (while the PXP would be the AND-FA).

The generator of the continuous-time Markov dynamics is the operator

$$\mathbb{W} = \sum_{j=1}^{N} \mathbb{P}_j [c\sigma_j^+ + (1-c)\sigma_j^- - c(1-n_j) - (1-c)n_j], \tag{2}$$

where $\sigma_j^{\pm}$ are Pauli operators acting on site j, $n_j = \sigma_j^+\sigma_j^-$, and the kinetic constraint $\mathbb{P}_j$ on site j reads

$$\mathbb{P}_j = (n_{j-1} + n_{j+1} - 2n_{j-1}n_{j+1}) = \tfrac{1}{2}\left(1 - \sigma_{j-1}^z \sigma_{j+1}^z\right), \tag{3}$$

where $\sigma_j^z = 2n_j - 1$. The operator Eq. (3) enforces the impossibility of the transitions ruled out in Eq. (1). Note that Eq. (2) has an explicit symmetry between up/down spins and is unchanged under the transformation $c \to 1-c$ and $n_j \to 1 - n_j$.

Dynamics with the kinetic constraint Eq. (3) is naturally motivated [39] in quantum many-body systems, specifically in the context of Rydberg atoms in their facilitated/antiblockade regime [31–41], whereby an up (down) spin represents an atom in its excited Rydberg (ground) state, and the drive is such that an atom can get excited resonantly only when one of its nearest neighbors is also excited, but not both (as that would make the transition off resonant). The constraint Eq. (3) has also been studied in certain quantum spin chains [43,44] in particular in relation to "quantum scars" (special nonthermal states in constrained quantum systems [30,45,46]) [47–49], and additionally in the context of quantum cellular automata [50]. Our aim here is to consider the classical stochastic dynamics of a system with such a constraint, thus extending the set of known KCMs.

Conservation of the number of domain walls and relation to simple exclusion processes

The dynamical rules Eq. (1) impose a conservation law in the dynamics, that of the total number of domain walls (DWs) [39,43]. Consider two neighboring domains of, say, up and down spins

$$\cdots 11110000 \cdots .$$

Due to the constraint Eq. (3) the only allowed changes are to the spins next to the DW, since inside the domains both neighbors to every spin are the same. This means that the possible moves are

$$\cdots 11110000 \cdots \begin{array}{c} \nearrow \cdots 111\underline{0}0000 \cdots \\ \searrow \cdots 1111\underline{1}000 \cdots \end{array},$$

where we have underlined the sites that changed in each allowed transition.

We can perform a duality transformation to have an explicit DW representation of the problem. We write

$$\sigma_j^x = X_j X_{j+1}, \tag{4}$$

$$\sigma_j^y = (-1)^{j+1} \prod_{k=1}^{j-1} Z_k Y_j X_{j+1}, \tag{5}$$

$$\sigma_j^z = (-1)^{j+1} \prod_{k=1}^{j} Z_k, \tag{6}$$

where X_j, Y_j, Z_j are Pauli operators for the DW between sites $j-1$ and j. Notice that this is a canonical (rather than unitary) transformation that preserves the commutation relations between the Pauli operators. The generator in this representation is then

$$\begin{aligned} \mathbb{W}^{\mathrm{DW}} = \sum_{j=1}^{N} \mathbb{P}_j^{\mathrm{DW}} \Bigg[& \frac{1}{2} X_j X_{j+1} \\ & - i\left(\frac{1}{2} - c\right)(-1)^{j+1} \prod_{k=1}^{j-1} Z_k Y_j X_{j+1} \\ & - \left(\frac{1}{2} - c\right)(-1)^{j+1} \prod_{k=1}^{j} Z_k - \frac{1}{2} \Bigg], \end{aligned} \tag{7}$$

where the constraint is

$$\mathbb{P}_j^{\mathrm{DW}} = \tfrac{1}{2}(1 - Z_j Z_{j+1}), \tag{8}$$

and we have used the superscript "DW" to indicate operators in the domain wall representation. Combining the factors we can simplify the generator to

$$\begin{aligned} \mathbb{W}^{\mathrm{DW}} = \sum_{j=1}^{N} \frac{1}{2} \Bigg[& \frac{1}{2}(X_j X_{j+1} + Y_j Y_{j+1}) \\ & - i\left(\frac{1}{2} - c\right)(-1)^{j+1} \prod_{k=1}^{j-1} Z_k (Y_j X_{j+1} - X_j Y_{j+1}) \end{aligned}$$

$$+\frac{1}{2}(Z_jZ_{j+1}-1)$$
$$-\left(\frac{1}{2}-c\right)(-1)^{j+1}\prod_{k=1}^{j-1}Z_k(Z_j-Z_{j+1})\Big]. \quad (9)$$

The conservation law is now explicit, as the kinetic term simply corresponds to DW hopping. That is, we conserve the quantity $\mathbb{N}_{\rm DW}=\frac{1}{2}\sum_j \mathbb{I}+Z_j$. For the special case of $c=1/2$, the generator simplifies to that of the symmetric simple exclusion process (SEP) [51,52],

$$\mathbb{W}^{\rm DW}_{c=1/2}=\sum_{j=1}^{N}\frac{1}{4}(X_jX_{j+1}+Y_jY_{j+1}+Z_jZ_{j+1}-1). \quad (10)$$

In the XOR-FA language this is the "infinite temperature limit," where the cost of creating and destroying an excitation is the same. Away from $c=1/2$, the dependence of c in the DW-representation generator Eq. (9) encodes the fact that moving left or right a DW depends on whether it is energetically favorable to extend or contract the corresponding spin domain.

For $c\neq 1/2$ the generator Eq. (9) corresponds to a SEP with particles with alternating asymmetries in their hopping rates. That is, we have a model where particles (DWs) can hop to neighboring sites if the sites are not already occupied: the odd particles (DWs) hop left with rate c and right with rate $1-c$, while the even particles hop left with rate $1-c$ and right with rate c. Since particles (DWs) cannot cross due to the exclusion, these rates are maintained. This is a special case of the general model introduced in Ref. [53], where each particle is given an individual hopping rate which is maintained under the dynamics. Independently from Ref. [53], this exclusion process was studied in Ref. [54]. In that paper the authors use a transformation onto a representation which coincides to our spin model, allowing them to find a hydrodynamic limit with a nontrivial diffusion rate for the exclusion process with alternating hopping rates.

III. EQUILIBRIUM AND RELAXATION

A. Equilibrium properties

We consider the XOR-FA with N sites, $N_{\rm DW}$ domain walls and periodic boundary conditions (PBC), which formally corresponds to setting $n_0=n_N$. The dynamics generated by Eq. (2) obeys detailed balance and therefore any initial condition eventually relaxes to an equilibrium state. Since the dynamics conserves the number of DWs, there is one such equilibrium probability for each DW sector. For PBC the number of DWs is even, and the sectors can be classified by the number p of up/down (one/zero) domains, $p=N_{\rm DW}/2$. One can then construct the equilibrium state within each sector in the following way.

Consider a configuration for fixed p where the zero (or down) domains and the one (or up) domains have lengths d_m and u_m, respectively, for $m=1,\ldots,p$, with the first domain being a down one,

$$\left|0..0_{d_1}1..1_{u_1}\ldots 0..0_{d_p}1..1_{u_p}\right\rangle.$$

Note that the total length of the domains must be equal to the system size, so in the state above we have

$$\sum_{m=1}^{p}(d_m+u_m)=N, \quad (11)$$

and each domain must have at least one site, so that

$$d_m,u_m\geqslant 1\ \ \forall m. \quad (12)$$

We now define a state which is the translationally invariant superposition of all possible translations of the state above,

$$|d_1,\,u_1,\,\ldots,\,d_p,\,u_p\rangle=\sum_{m=1}^{N}\mathbb{T}^m\left|0..0_{d_1}1..1_{u_1}\ldots 0..0_{d_p}1..1_{u_p}\right\rangle,$$

where the operator $\mathbb{T}$ shifts the chain by a single site.

The equilibrium probability vector for the sector with $2p$ DWs is given by

$$|{\rm eq}_p\rangle=\mathcal{N}\sum_{d_1=1}^{\gamma}\cdots\sum_{d_p=1}^{\gamma}\sum_{u_1=1}^{\gamma}\cdots\sum_{u_p=1}^{\gamma}\delta(d_1+\cdots+u_p-N)$$
$$(1-c)^{\sum_m d_m}c^{\sum_m u_m}\,|d_1,\,u_1,\,\ldots,\,d_p,\,u_p\rangle\,, \quad (13)$$

where $\gamma=N-2p+1$ and $\mathcal{N}$ is a normalization constant. One can check that the state Eq. (13) is annihilated by all terms of the generator Eq. (2). This state corresponds to the equilibrium state with noninteracting energy $E=\sum_j n_j$ (i.e., each up spin costs a unit of energy) at temperature T such that $c=e^{-1/T}/(1+e^{-1/T})$, and subject to the conditions Eqs. (11) and (12).

We now study the basic properties of the equilibrium state Eq. (13). In Fig. 1 we show two average observables in equilibrium. The first one is the average excitation density, $\langle n\rangle=N^{-1}\sum_i\langle -|n_i|{\rm eq}_p\rangle$, where $\langle -|=\sum_{\boldsymbol{n}}\langle \boldsymbol{n}|$ is the flat state and $\langle \boldsymbol{n}|=\langle n_1,\ldots,n_N|$, see Fig. 1(a). We show $\langle n\rangle$ for several values of the filling fraction defined as $n_{\rm DW}=2p/N$ (note that the mean domain length is $1/n_{\rm DW}$). The symbols are numerical results from standard Monte Carlo simulations. Note that in contrast to the FA or East models [17,42], $\langle n\rangle$ is not just equal to c, due to the conservation of the number of DWs. Figure 1(a) shows the agreement of the numerics with an analytical prediction in the $N\to\infty$ limit described in the Appendix.

The second observable coincides with the average dynamical activity (per site) in equilibrium, $\langle k\rangle$. While the activity is an observable at the level of trajectories (see Sec. IV for further details), its average in equilibrium is given by the average escape rate per site, which is a static observable [55]. The escape rate operator $\mathbb{R}$ is (minus) the diagonal part of the generator Eq. (2). Since $\mathbb{R}$ is a local operator we can also obtain analytically its equilibrium average in the large size limit, see the Appendix. In Fig. 1(b) we show the agreement between $\langle k\rangle$ from simulations and the analytic result. Note that $\langle n\rangle$ and $\langle k\rangle$ are symmetric around $c=1/2$ as functions of c as a consequence of the up/down symmetry of the model, cf. Eq. (2) (while there is no corresponding symmetry in terms of DW filling $n_{\rm DW}$).

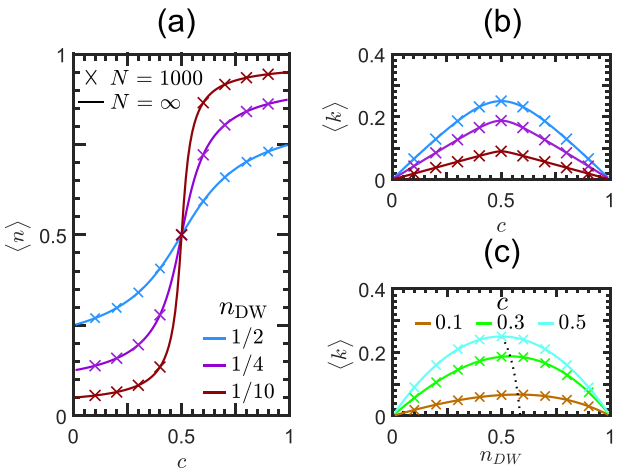


FIG. 1. Equilibrium properties of the XOR-FA model. We show various properties of the XOR-FA at equilibrium for both finite-size systems ($N = 1000$, symbols) obtained via Monte Carlo simulations, and $N = \infty$ (solid lines) obtained through the analytical considerations from the Appendix. (a) The average excitation density for various DW fillings as a function of c. (b) The average dynamical activity $\langle k \rangle$ as a function of c for the same DW fillings of panel (a). (c) The average dynamical activity $\langle k \rangle$ as a function of DW filling $n_{\rm DW}$ for various c. Note that the peak in the dynamical activity with respect to $n_{\rm DW}$ varies with c (dotted line).

B. Relaxation dynamics

The dynamics of the XOR-FA model is entirely determined by expansion and contraction of the domains (or the movement of domain walls which cannot cross). The system behaves like an "accordion." Depending on the value of c there may be an energetic preference to expand or contract domains of one orientation or the other. Figure 2 shows typical trajectories sampled from the XOR-FA model by running continuous time Monte Carlo at various conditions. The trajectories of the figure are at a quarter filling for three values of c. The top row of Fig. 2 is for $c = 0.4$, the middle one for $c = 0.5$, and the bottom one, $c = 0.6$. The columns correspond to different initial conditions. Column (a) shows equilibrium trajectories, i.e., those that start from an initial condition sampled from Eq. (13). They show pronounced space-time fluctuations in the dynamics associated to the breathing of domains. Column (b) corresponds to the most unfavorable initial state, where DWs are maximally clustered. Relaxation to equilibrium in this case is slow, as DWs within the bulk of the cluster cannot move until the DWs on the outside of cluster diffuse away. Column (c) shows an opposite nonequilibrium initial condition, where DWs are maximally spread out. In this case relaxation to equilibrium is faster, cf. Fig. 2(c). The large space-time fluctuations that are evident in these example trajectories anticipate the large deviation phase transitions that we uncover in the next section.

The different timescales involved in the relaxation of the XOR-FA model can be quantified using time-correlation functions. In particular we focus on two different correlators in the equilibrium dynamics. The first one is the autocorrelation function, $C(t)$, which measures how many sites that were in

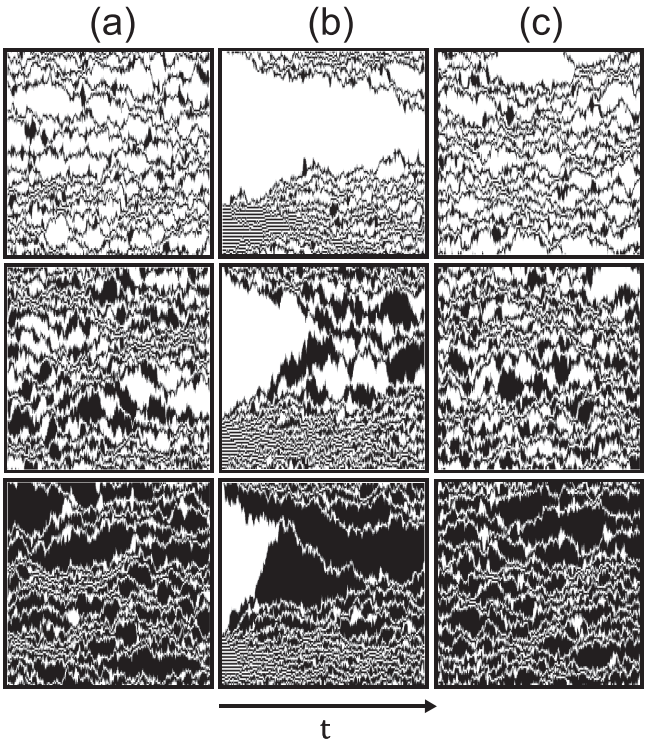


FIG. 2. Trajectories of the XOR-FA model. Representative trajectories from continuous-time Monte Carlo simulations using generator Eq. (2) with a time of $t = 10^3$. The rows are for different values of c, with $c = 0.4$ (top), $c = 0.5$ (middle), $c = 0.6$ (bottom). All trajectories are at quarter filling, $n_{\rm DW} = 1/4$. The columns correspond to different initial conditions: (a) typical equilibrium configuration, (b) DWs maximally clustered, (c) DWs maximally spread out. Column (b) shows the slowest approach to equilibrium.

the excited state at time 0 are also in an excited state at a later time t. Subtracting the disconnected part, and normalizing so that it takes values between 1 and 0, it reads

$$C(t) = \frac{1}{N}\sum_{j=1}^{N} \frac{\langle n_j(t) n_j(0)\rangle - \langle n\rangle^2}{\langle n\rangle - \langle n\rangle^2}, \tag{14}$$

where the average is over realisations of the dynamics in equilibrium, i.e., starting from a configuration sampled from the equilibrium state Eq. (13) and evolved according to Eq. (2).

The second correlator we study is the persistence function, $P(t)$, which quantifies the average probability for a randomly selected site to have not changed state up to time t. We can define it in terms of a local dynamical variable $p_j(t)$ at each site j, where $p_j(t) = 1$ if the site has never changed from its initial state at time t, and $p_j(t) = 0$ as soon as it changes for the first time. The resulting aggregate function is then

$$P(t) = \frac{1}{N}\sum_{j=1}^{N} \langle p_j(t)\rangle \,. \tag{15}$$

This function is automatically normalized between 1 at the initial time and 0 eventually when all sites flip at least once.

In Fig. 3 we show results for time-correlators. We focus mostly on the persistence function as it better captures

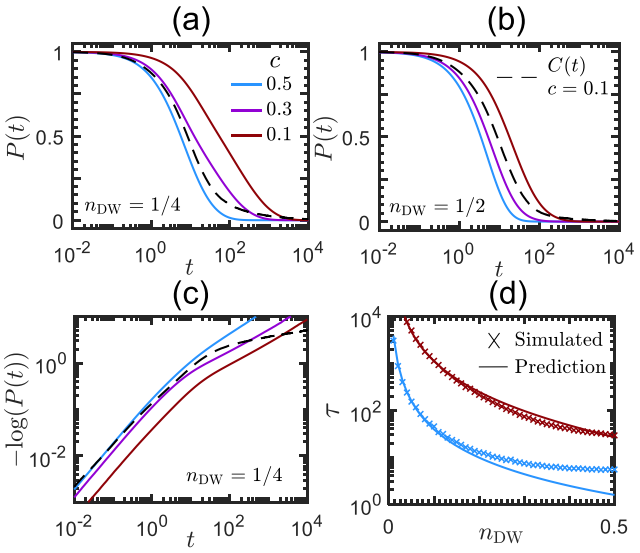


FIG. 3. Time correlations of the XOR-FA model. All results are done for $N = 400$. We show the persistence function $P(t)$ plotted for (a) $n_{\rm DW} = 1/4$ and (b) $n_{\rm DW} = 1/2$. In both cases we show for various c. The black dashed line also shows the autocorrelator $C(t)$ for $c = 0.1$ to compare. (c) We plot the same functions as shown for (a) but with a double-logarithmic scale on the ordinate. (d) We show the time taken τ for the persistence function to drop to $P(\tau) = e^{-1}$ (crosses) for $c = 0.5$ (blue, bottom), $c = 0.1$ (red, top) and various $n_{\rm DW}$. We also show our estimate τ (solid lines) given in Eq. (17).

overall relaxation. Figure 3 shows $P(t)$ for various c and two filling fractions of DWs, $n_{\rm DW} = 1/4$ (a) and $n_{\rm DW} = 1/2$ (b). For comparison we also show the autocorrelator for $c = 0.1$ (dashed). We see that decreasing c away from $c = 0.5$ leads to slower relaxation times. The same can be said for decreasing the density of the DWs. Figure 3(c) shows the same functions as in Fig. 3(a) but in a double-logarithmic scale on the ordinate. The change in slope in this representation emphasises the change from exponential decay at short times, to stretched exponential decay at long times [56].

From the persistence function we can extract a characteristic relaxation time, τ, customarily from the time the function decays to e^{-1}, that is, $P(\tau) = e^{-1}$. These times are shown in Fig. 3(d) for two values of c and as a function of the DW filling. Their behavior can be understood approximately with simple heuristic arguments.

We first note that for $c \ll 1$, we can treat the dynamics of the XOR-FA model as small up domains diffusing around a "vacuum" of down domains. To move, the up domain must first expand by exciting a neighboring spin. This happens slowly at rate c. Following this, the domain then shrinks at rate $1 - c \approx 1$. It can either shrink back to its original position, or shrink such that it shifts by one site across, each happening with equal probability. Thus, we say it diffuses around the lattice with diffusion constant $D_c \approx c/2$. The time taken for the system to fully relax can then be estimated as the time it takes for the DWs to diffuse from their original positions around the available space surrounding them, until they hit another DW. On average, the length of each zero domain is given by the average number of down spins split among the number of zero domains. Namely,

$$l_0 = \frac{2}{n_{\rm DW}}(1 - \langle n \rangle). \tag{16}$$

It then follows that the timescale for the system to relax goes as

$$\tau \sim \frac{(l_0/2)^2}{D_c} = \frac{2}{c}\frac{(1 - \langle n \rangle)^2}{{n_{\rm DW}}^2} \tag{17}$$

for c small. As Fig. 3(d) shows, this prediction works well for c small in the whole $n_{\rm DW}$ range, while for $c \approx 0.5$ it qualitatively accounts for τ for small DW density [57].

Thinking of the dynamics in this way can also explain the two timescales in Fig. 3(c). At some small time after $t \gtrsim 0$, the first successful shift of domain(s) will occur. When this happens for $c \ll 1/2$, the original site is no longer excited, but the site next to it is. In the language of the persistence, this means two sites have flipped from their initial state. For both the persistence and the autocorrelator, this gives a fast initial relaxation, and as these are random uncorrelated events, the initial decay is exponential. Further successive moves of the domain only change at most one more site from its initial state (or in the case of the autocorrelator, will only slightly reduce the probability that the domain may end up in its original position). Thus, the rate at which relaxation occurs is reduced, the time is longer, and the decay of the correlators is stretched as the relaxation becomes more collective.

IV. DYNAMICAL LARGE DEVIATIONS AND MATRIX PRODUCT STATES

In this section we study the statistics of trajectories of the XOR-FA model in the long-time regime where we can apply large deviation (LD) methods [19,58,59]. Recent work [60–62] has shown the effectiveness of numerical tensor network methods for studying the LDs of KCMs. Here, by means of numerical matrix product states (MPS) we are able to study the LDs of the XOR-FA for large systems to high accuracy. As we show below, the XOR-FA has a trajectory-space phase transition between between dynamical phases with very distinct characteristics, similar to what occurs in several other KCMs.

A. LDs and tilted generators

The dynamical activity [19,63–65] is a trajectory observable which counts the number of configuration changes (in our case the number of spin flips) in some given time. It is the natural trajectory observable to quantify the amount of motion in the dynamics. A question one can ask is what is the probability of observing the activity K for trajectories ω_t which run for a total time t. The probability distribution for K is given by

$$P_t(K) = \sum_{\omega_t} \pi(\omega_t)\, \delta[K(\omega_t) - K], \tag{18}$$

where $\pi(\omega_t)$ is the probability of observing trajectory ω_t. For long times this obeys a large deviation (LD) principle $P_t(K) \approx e^{-t\varphi(K/t)}$ where $\varphi(K/t)$ is the LD rate function [58]. One can

also consider the moment generating function

$$Z_t(s) = \sum_K P_t(K)\, e^{-sK} = \sum_{\omega_t} \pi(\omega_t)\, e^{-sK(\omega_t)}, \tag{19}$$

which also obeys a LD principle, $Z_t(s) \approx e^{t\theta(s)}$ where $\theta(s)$ is the scaled cumulant generating function (SCGF) whose derivatives at $s = 0$ give the cumulants of K, scaled by time [58]. The SCGF plays the role of the thermodynamical free energy and is related to the LD rate function by a Legendre transform $\theta(s) = -\min_k[sk + \varphi(k)]$ [58].

We can deform the generator given in Eq. (2) by multiplying the off-diagonals by a factor of e^{-s} to give the tilted generator,

$$\begin{aligned}\mathbb{W}_s = \sum_{j=1}^{N} \mathbb{P}_j\{e^{-s}[c\sigma_j^+ + (1-c)\sigma_j^-] \\ - c(1-n_j) - (1-c)n_j\},\end{aligned} \tag{20}$$

whose largest eigenvalue is the SCGF $\theta(s)$ [58]. It has the associated left and right eigenvectors, $\langle l_s|\, \mathbb{W}_s = \theta(s)\, \langle l_s|$ and $\mathbb{W}_s\, |r_s\rangle = \theta(s)\, |r_s\rangle$, respectively. As the dynamics obeys detailed balance, we can transform the generator into a Hermitian one by using a similarity transformation independent of s [55]. We first define the diagonal matrix Q with matrix elements $\langle \mathbf{n}|Q|\mathbf{n}\rangle = (1-c)^{N/2}[c/(1-c)]^{\sum_i n_i/2}$. The tilted Hamiltonian $\mathbb{H}_s = -Q^{-1}\,\mathbb{W}_s Q$ is then given by

$$\mathbb{H}_s = -\sum_{j=1}^{N} \mathbb{P}_j\big[e^{-s}\sqrt{c(1-c)}\sigma_j^x - c(1-n_j) - (1-c)n_j\big], \tag{21}$$

which has the ground state $\mathbb{H}_s\, |\psi_s\rangle = -\theta(s)\, |\psi_s\rangle$. As was done for the generator, one can write the tilted Hamiltonian in the DW representation

$$\begin{aligned}\mathbb{H}_s^{\mathrm{DW}} = -\sum_{j=1}^{N} \frac{1}{2}\Big[& e^{-s}\sqrt{c(1-c)}(X_jX_{j+1} + Y_jY_{j+1}) \\ & + \frac{1}{2}(Z_jZ_{j+1} - 1) \\ & - \left(\frac{1}{2} - c\right)(-1)^{j+1}\prod_{k=1}^{j-1} Z_k(Z_j - Z_{j+1})\Big].\end{aligned} \tag{22}$$

The eigenvector $|\psi_s\rangle$ of $\mathbb{H}_s$ is related to the left and right eigenvectors of $\mathbb{W}_s$ by

$$|\psi_s\rangle = \sum_{\mathbf{n}} \sqrt{l_s(\mathbf{n})r_s(\mathbf{n})}\, |\mathbf{n}\rangle\,, \tag{23}$$

where $l_s(\mathbf{n}) = \langle l_s|\mathbf{n}\rangle$ and $r_s(\mathbf{n}) = \langle \mathbf{n}|r_s\rangle$. Thus, studying the LDs reduces to diagonilizing Eq. (21) to find $\theta(s)$ and $|\psi_s\rangle$.

B. Matrix product states

A matrix product state (MPS) is an ansatz for the vector state of a many-body system [66–68]. For a chain of N finite dimensional subsystems (of dimension d), it corresponds to states of the form

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N}^{d} \mathrm{Tr}\big(A_1^{i_1}A_2^{i_2}\ldots A_N^{i_N}\big)\, |i_1\, i_2\, \ldots\, i_N\rangle\,, \tag{24}$$

where i_k labels the states of the physical basis for the kth subsystem and each A_k is a rank-3 tensor with dimensions $d \times D \times D$, with D the so-called bond dimension. Thus, the MPS is described by $O(NdD^2)$ parameters. Notice that by increasing D, any arbitrary state can be exactly written in the form of Eq. (24), although this may require up to $D = d^{\lfloor N/2\rfloor}$.

The bond dimension D limits the entanglement within the state. More precisely, in a MPS with bond dimension D, the entanglement entropy for a subchain L (defined as $S_E = -\mathrm{Tr}\rho_L \log \rho_L$ where $\rho_L = \mathrm{Tr}_{N\setminus L}\, |\Psi\rangle\, \langle\Psi|$ is the subchain reduced density matrix) is upper-bounded by $S_E \leqslant 2\log D$, independent of the subchain length. This implies that the MPS satisfies an entanglement area law, which is intimately related to their success at approximating relevant physical states [69]. In particular, ground states of local gapped Hamiltonians, which in one spatial dimension are known to satisfy an area law [70], but also of critical models, can be efficiently approximated by MPS [70,71]. Furthermore, MPS constitute the basis of efficient numerical methods, including the celebrated density matrix renormalization group (DMRG) algorithm [72,73] which we use to approximate the ground state of $\mathbb{H}_s$.

The DMRG, originally formulated in Ref. [72], can be understood as a variational minimization of energy over the set of MPS. By writing the operator Eq. (21) as a matrix product operator (MPO) [74,75], one can perform a local optimization on a single tensor within the MPS to minimize the energy. We iterate through each tensor, applying local updates until convergence. This variational MPS search (vMPS) is well detailed in many reviews (e.g., Refs. [68,76]). For completeness, we give a brief description in the Appendix.

When applying the vMPS to study the LD statistics of the XOR-FA model, we use open boundary conditions (OBC) which formally corresponds to setting $n_0 = n_{N+1} = 0$, as this allows for the most efficient MPS calculations, with computational cost $O(D^3)$. In our problem, the number of DWs defines a global conserved quantity, and we need to find the ground state in a particular sector. Although it is possible to encode this symmetry in the tensors [77–79], as other local constraints have [80], we opt here for a simpler approach. Namely, we add an energy penalty to the Hamiltonian to favor the desired sector. Specifically, the penalty is $\lambda(\mathbb{N}_{\mathrm{DW}} - N_{\mathrm{DW}})^2$ where $\lambda > 0$ is some Lagrange multiplier and $\mathbb{N}_{\mathrm{DW}} = \sum_{i=0}^{N} n_i(1 - n_{i+1}) + (1-n_i)n_{i+1}$ is the operator which counts the number of DWs.

C. Results

As we now show, the MPS ansatz combined with the variational search proves to be very efficient for studying the LDs of the XOR-FA model, just like for other KCMs and exclusion processes [60–62]. In this way we are able to achieve results for system sizes superior to traditional methods such as exact diagonalization or importance sampling.

1. First-order phase transition in the SCGF

A key property of other KCMs such as the FA or the East model is the presence of a first-order dynamical phase transition [64] in the thermodynamical limit $N \to \infty$, manifested as a singularity in the SCGF $\theta(s)$ at $s = 0$. Consequentially

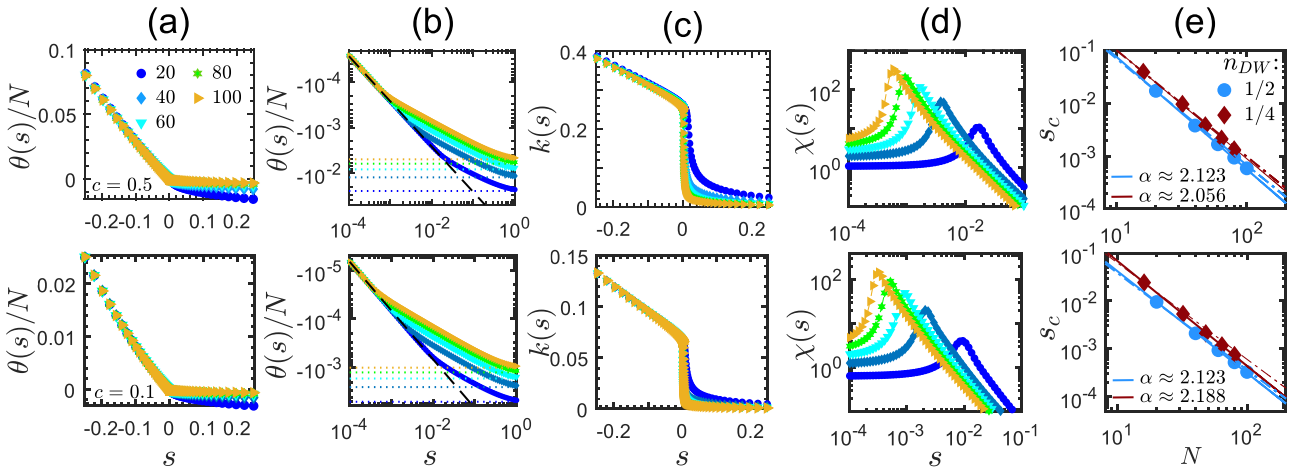


FIG. 4. First-order phase transition in the SCGF. We consider the finite-size scaling of the phase transition for $N \in [20, 100]$ and $n_{\mathrm{DW}} = 1/2$. (a) A linear-linear plot of the SCGF $\theta(s)/N$. There is an apparent change in behavior at some critical point $s_c(N) > 0$. (b) The SCGF $\theta(s)/N$ for $s > 0$ on a log-log plot. On the first branch, $\theta(s)/N$ follows a LR (dashed line) up to $s_c(N)$ after which it follows another branch, which is no longer linear or scales with system size. The dotted lines show the value of the SCGF at $s \to \infty$. (c) The activity $k(s) = -\theta'(s)/N$ has a drop around $s = s_c(N)$ which becomes sharper with system size. (d) The dynamical susceptibility $\chi(s) = \theta''(s)$ which peaks at $s_c(s)$. The peaks become narrower and larger as N becomes larger. (e) The critical point $s_c(N)$ extracted from the peak of the susceptibility plotted against system size, for the DW densities $n_{\mathrm{DW}} = 1/2$ (blue circle) and $n_{\mathrm{DW}} = 1/4$ (red diamonds). We fit the data the the power law $s_c(N) \propto N^{-\alpha}$ (solid line) and to the polynomial $s_c(N) = aN^{-2} + bN^{-3} + cN^{-4}$ (dashed lines) which are the subleading corrections to N^{-2}.

there are two dynamical phases—the active phase for $s < 0$ and the inactive phase for $s > 0$. We look for evidence for this transition in the XOR-FA model.

Figures 4(a) and 4(b) show the SCGFs obtained numerically for system sizes $N \in [20, 100]$, in linear and log-log scales, respectively. The upper row of Fig. 4 is for $c = 0.5$ while the lower row corresponds to $c = 0.1$. For finite size, near enough $s = 0$ the SCGF should obey the linear response (LR), $\theta(s) \sim -sk_{\mathrm{eq}}$, where k_{eq} is the average activity per unit time in the equilibrium state. For the FA and East models, the equilibrium activity is straightforward to calculate exactly (see, e.g., Ref. [81]). For the XOR-FA, it is more difficult due to the conservation of DWs. In the Appendix we give a way to compute it to a good approximation. As we see from Fig. 4(b), the SCGF we obtain numerically does obey LR close to $s = 0$.

Still for $s \gtrsim 0$, beyond the LR regime the SCGF changes behavior, notably stops scaling with system size, see Figs. 4(a) and 4(b). This change in behavior becomes even more apparent when one considers the activity per unit time as a function of s, $k(s) = -\theta'(s)/N$, Fig. 4(c). We see a sudden drop close to $s = 0$ that becomes more pronounced with system size, a hallmark of a first-order phase transition. From the point of the numerics, this occurs when where there two smallest energy levels of Eq. (21) cross.

The transition point s_c can be estimated from the peak of the susceptibility $\chi(s) = \theta''(s)$; see Fig. 4(d). The peaks become higher and sharper with system size. From the numerics we can make a finite-size scaling analysis of the critical point. We find that $s_c(N)$ seems to obey $s_c(N) \propto N^{-\alpha}$ as shown in Fig. 4(e). For $n_{\mathrm{DW}} = 1/2$, we find that the best fit exponent is $\alpha \approx 2.123$ for both $c = 0.5$ and $c = 0.1$. Furthermore, for $n_{\mathrm{DW}} = 1/4$ we find that $\alpha \approx 2.056$ and $\alpha \approx 2.188$ for $c = 0.5$ and $c = 0.1$, respectively. In each case α is close to the value 2 expected from a diffusive behavior of the gap in the spectrum of Eq. (21). It could be that these are subleading corrections to N^{-2}; see Fig. 4(e).

2. *Spatial structure of active and inactive phases*

The singularity in the SCGF represents a phase transition in the trajectories of the dynamics at the level of fluctuations: The behavior at $s < 0$ corresponding to dynamics with activity that is larger than the typical (equilibrium) one is fundamentally different from that at $s > 0$ corresponding to dynamics which is less active than typical. This difference is also manifested in the configurations that are predominantly visited by the different trajectories. That is, active and inactive dynamical phases are associated with very different spatial structures.

We mean the following. At $s = 0$ there is no tilting in the ensemble of trajectories which is the one given by the original dynamics. It corresponds to typical behavior. Dynamics is ergodic over configuration space, and over long-times the distribution of configurations that are visited—for some fixed value of the number of DWs—is given by the equilibrium probability Eq. (13). The state Eq. (13) is the leading right eigenstate of generator Eq. (20) at $s = 0$.

At $s \neq 0$ the probability of a trajectory is reweighted by the exponential of its activity, cf. Eq. (19). How often configurations are visited in such reweighted ensembles depends on s, and for long times the distribution over configurations is given by the leading eigenstate of Eq. (20), or equivalently Eq. (21) for the detailed balance-obeying XOR-FA. We have access to these states, $|\psi_s\rangle$, from our MPS numerics.

The easiest way to characterise the spatial structure of the characteristic configurations associated with dynamics tilted by s is to study the average local occupations $\langle n_i \rangle_s = \langle \psi_s | n_i | \psi_s \rangle$ [82]. In Figs. 5(a) and 5(b) we show these local

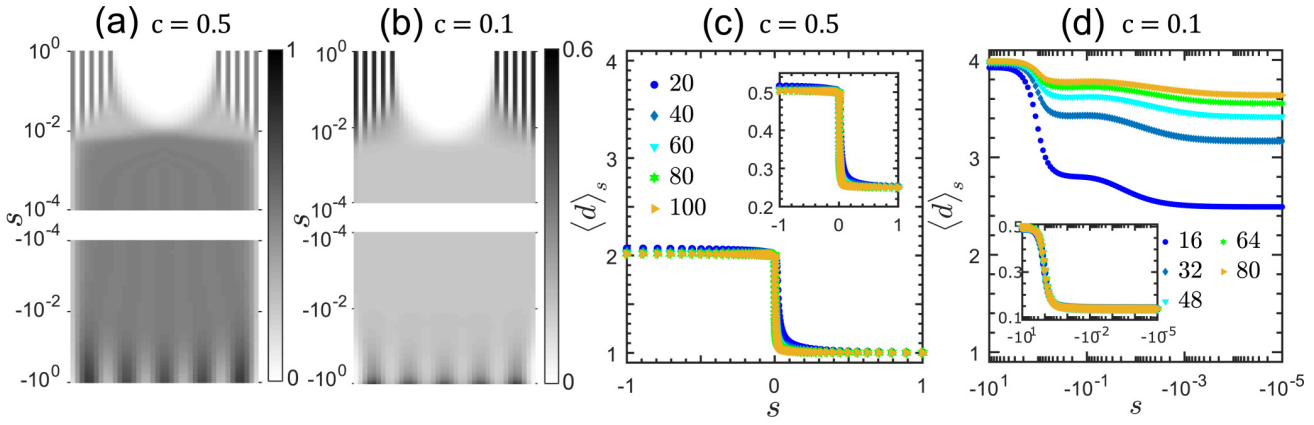


FIG. 5. Spatial structure of the active and inactive phases. (a) The average occupation of each site for a system with $c = 0.5$, $N = 40$ and $N_{\mathrm{DW}} = 10$. For $s < 0$ there is a clear localization of domains. Each domain becomes (on average) equally sized and hence the DWs are equally spread. For $s > 0$, we see the DWs gather at the edge(s). (b) The same as (a) but for $c = 0.1$. (c) The average distance between neighboring DWs for a systems with $c = 0.5$ and $n_{\mathrm{DW}} = 1/2$. This shows there is a maximal separation between DWs in the active phase, and only one site separating DWs for the inactive phase. Inset: The average excitation density as a function of s. (d) The same as (c) but with $c = 0.1$ and $n_{\mathrm{DW}} = 1/4$. Here we observe multiple plateaus in the growth. Inset: The average excitation density as a function of s.

densities as a function of s for two values of c and $n_{\mathrm{DW}} = 1/4$. For s large and negative we see that that domains becomes maximally sized, that is, DWs become maximally spaced apart, maximizing the activity, as expected. In contrast, for s large and positive DWs become localized at either edge of the system. When DWs become minimally separated and clustered together, only the sites next to the last DW are allowed to move and activity becomes subextensive and thus minimal.

We can further quantify the average distance between neighboring DWs by considering the operator

$$\mathbb{D}_d = \sum_{i=1}^{N+1} n_{i-1}(1-n_i)(1-n_{i+1})\dots(1-n_{i+d-1})n_{i+d} + (1-n_{i-1})n_i n_{i+1}\dots n_{i+d-1}(1-n_{i+d}), \tag{25}$$

which measures the likelihood of observing two neighboring DWs at distance d apart. The average distance between neighboring DWs is then given by

$$\langle d\rangle_s = (N_{\mathrm{DW}}-1)^{-1}\sum_d \langle\psi_s|\mathbb{D}_d|\psi_s\rangle \tag{26}$$

(as we have $N_{\mathrm{DW}} - 1$ pairs of neighboring DWs). In Fig. 5(c) we show $\langle d\rangle_s$ as a function of s for $c = 0.5$ and $n_{\mathrm{DW}} = 1/2$. It is evident from the plot that the dynamical transition at $s_c \approx 0$ is also one where there is a singular change in the distance between DWs in the corresponding characteristic configurations, from maximal distance between DWs at s negative, to minimal at s positive.

Figure 5(d) shows the same for $c = 0.1$ and $n_{\mathrm{DW}} = 1/4$. We see that away from the SEP limit of the XOR-FA, there is even richer spatial structure due to the energetic cost associated with domains. At small $s < 0$ there is an initial plateau in the growth of the average distance between DWs. This is an extension of the equilibrium behavior, where domains are randomly sized according to the behavior described in Sec. III. As we move further into the active phase, we observe another plateau. At this point, the excitation density $\langle n\rangle_s = N^{-1}\sum_i \langle n_i\rangle_s$ (as shown in the inset) has not varied much from the equilibrium value. This leads us to believe that this change in structure is due to the excited domains spreading apart and becoming localized as shown in Fig. 5(b). This maximizes the activity without having to grow the excited domains as is energetically favorable for $c < 1/2$. Following this plateau, there is a steady growth to the maximum $\langle d\rangle_s$ corresponding to the growth of the one domains, such that every DW is on average equally spaced. This is further supported by the fact that following this plateau, the excitation density rapidly grows.

3. *Maximally and minimally active limits*

While we cannot calculate the SCGF analytically in general, there are limits where the calculation becomes tractable [apart from the obvious case of $\theta(0) = 0$]. These are the $s \to \pm\infty$ limits corresponding to the tilting of the dynamics that maximizes and minimizes the activity. For the former we can easily calculate the ground state of Eq. (21) via vMPS to obtain the rescaled SCGF $\tilde{\theta}/N = \lim_{s\to-\infty} e^s\theta(s)/[N\sqrt{c(1-c)}]$ using only a small bond dimension of $O(10)$. The numerical data is shown in Fig. 6(a) for various filling fractions. Note that for $s \to -\infty$ the dependence on c is irrelevant and can be scaled out as in our definition of $\tilde{\theta}$.

We can extrapolate from the numerical results for finite size to obtain an estimate of $\lim_{N\to\infty}\tilde{\theta}/N$ as a function of the density of domain walls n_{DW}. This large-size limit of the SCGF can be obtained analytically. For $s \to -\infty$, after scaling out the e^{-s} and $\sqrt{c(1-c)}$ factors, the Hamiltonian Eq. (22) becomes that of the XX model. Being noninteracting, the ground state can be obtained exactly by standard means [83], to give: $\lim_{N\to\infty}\tilde{\theta}/N = \frac{\pi}{2}\sin(\pi\, n_{\mathrm{DW}})$. Figure 6(a) shows the agreement between the numerical extrapolation and the exact result. The structure of the state in the maximally active limit for a system with $N = 40$ sites and $N_{\mathrm{DW}} = 20$ domain walls is

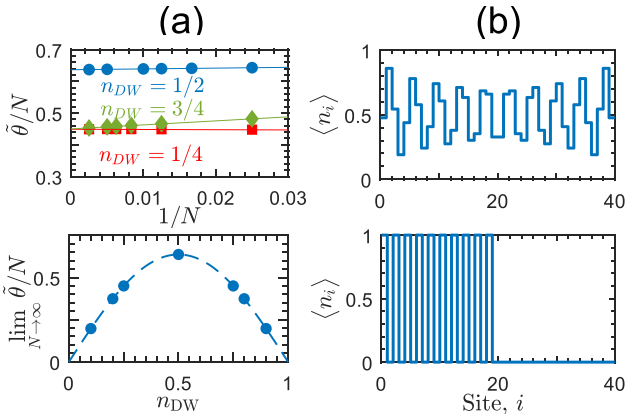


FIG. 6. Maximally and minimally active limits. (a) Top: The rescaled SCGF $\tilde{\theta}/N = \lim_{s\to-\infty} e^s\theta(s)/\sqrt{c(1-c)}$ for various number of DWs. We fit the data with the curves $\tilde{\theta}/N = a + bN^{-1}$. Bottom: $\tilde{\theta}/N$ in the limit $N \to \infty$ obtained by extrapolating the fitted curves in top figure, plotted against the density of domain walls $n_{\rm DW}$. We fit the data using $\lim_{N\to\infty} \tilde{\theta}/N = \frac{\pi}{2}\sin(\pi n_{\rm DW})$. (b) The average occupation at each site for systems a system with $N = 40$ and $N_{\rm DW} = 20$ in the limits $s \to -\infty$ (top) and $s \to \infty$ (bottom). For the latter, we show only one of the degenerate ground states. In this case, DWs localize at the left.

shown in Fig. 6(b). In the limit $N \to \infty$ we would expect that $\lim_{s\to-\infty} \langle n \rangle_s = 1/2$, and $\lim_{s\to-\infty} \langle d \rangle_s = {n_{\rm DW}}^{-1}$ which are both in excellent agreement with the numerical vMPS data.

For the minimally active limit $s \to \infty$, the Hamiltonian given in Eq. (21) becomes a diagonal one. Each configuration is an eigenstate and one can easily show that the configuration with the least energy is the one where all the DWs are clustered at the edge of the system (which is doubly degenerate). The SCGF at this limit is given by $\theta(s \to \infty) = -c$ and the exact structure for $N = 40$ and $N_{\rm DW} = 20$ is given in Fig. 6(b) for just one of the degenerate states (in practice the vMPS prefers to converge onto just one to keep the bond dimension minimal). Additionally, the excitation density and the distance between DWs are given by $\lim_{s\to\infty} \langle n \rangle_s$, $\langle d \rangle_s = n_{\rm DW}/2,\ 1$, respectively.

V. COMPARISON BETWEEN THE FA, XOR-FA, AND PXP MODELS

As discussed above, the kinetic constraint of the XOR-FA model is stronger than that of the FA model (which is a binary OR operation on the spins neighboring the one attempting to flip), but weaker than that of the PXP model (which is a binary AND operation on the neighboring spins). This has significant consequences on the dynamics.

In the case the FA model [17], configuration space is all connected by the dynamics, except for the configuration with $n_i = 0$ for all i. This means that in practice dynamics is irreducible and there is one giant ergodic component (as the probability of starting from the unique all-zero site is suppressed exponentially with size). Despite the relative weakness of the constraint, the dynamics of the FA is strongly fluctuating [84]. Figure 7 (top left) shows an example trajectory in equilibrium, displaying the characteristic

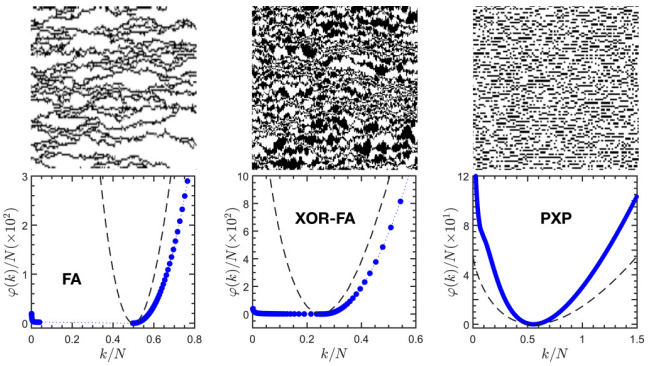


FIG. 7. Comparison between FA, XOR-FA and PXP models. The top panels show sample equilibrium trajectories for the FA model (left, taken from Ref. [84], $c \approx 0.27$, $N = 100$), the XOR-FA (middle, 1/4 filling, $c = 0.5$, $N = 200$), and the PXP model (right, no pairs of contiguous up spins, $c = 0.5$, $N = 100$). The bottom panels show the corresponding LD rate functions for the activity in the three models ($c = 0.5$ for each) obtained via numerical MPS. The data for the FA model is from Ref. [60] ($N = 200$), and that for the PXP is from Ref. [80] ($N = 400$). Both the FA and XOR-FA (1/2 filling, $N = 100$) rate functions show first-order transitions close to the typical dynamics (flatness of the bottom of the curves), while the PXP has a continuous transition for highly atypical low values of the activity (kink at left of curve, see Ref. [80]). The dashed curves are Poisson rate functions for comparison.

"space-time bubbles" responsible for dynamic heterogeneity [24,85]. This preponderance of fluctuations is manifest in the form of the LD rate function for the dynamical activity, see Fig. 7 (bottom left), corresponding to a (dynamical) first-order transition [64].

On the other extreme of this comparison is the PXP model [27–30]. In the stochastic terminology this corresponds to the two-spin facilitated FA model in one-dimension [17]. Here the constraint is such that configuration space is broken into exponentially many dynamically disconnected components. Specifically, pairs of sites with up spins, $n_i = n_{i+1} = 1$ cannot be flipped and are conserved by the dynamics. This means that dynamics is reducible as configuration space is partitioned into an exponential in size number of irreducible components, classified by local conservation laws (i.e., the location of the unmovable contiguous clusters of up spins). The largest ergodic component is that where no two up spins are adjacent. But despite the strength of the constraint, the resultant dynamics is much less interesting than for the FA or the XOR-FA models, see for example the sample trajectory of Fig. 7 (top right). Correspondingly, a detailed quantification of the statistics of the dynamics shows no significant fluctuations, see the LD rate function of Fig. 7 (bottom right).

The middle panels of Fig. 7 show the XOR-FA for comparison. Given that its constraint is somewhat in between that of the FA and PXP model, we see that trajectories display less pronounced "bubbles" than the FA but are more fluctuating that the PXP. Specifically, the constraint does break configuration space, but the number of disconnected ergodic components is only linear in the system size. These components are classified by the globally conserved number of domain walls. This allows for rich dynamics within the

components, leading to singular LD behavior as demonstrated in this paper; see LD rate function Fig. 7 (middle bottom).

The key observation is that both FA and XOR-FA models have trajectory transitions which manifest in fluctuating dynamics, while the PXP does not. In both the FA and XOR-FA models there are configurations which are completely void of dynamics, the all-zero state in the FA, and the state with maximal number of domain walls in the XOR-FA. But while these configurations are disconnected dynamically, many other configurations have finite regions that resemble them locally. Such spatial rare regions can only be relaxed from the outside, and thus give rise to the dynamical bubbles of the trajectories, see Fig. 7, and are at the source of the large fluctuations in the dynamics. In contrast, the PXP constraint makes inactive configurations local rather than global, and they can either be relaxed locally or not. In the PXP there are no space-time bubbles and no LD transition close to $s = 0$ (close at the edge of typical dynamics, see discussion of previous sections).

VI. CONCLUSIONS

We have studied a new classical stochastic KCM, the one-dimensional XOR-FA model, which could be experimentally realized with a laser-driven dissipative Rydberg lattice gas under facilitation (antiblockade) conditions [33–35]. The kinetic constraint in this model is stronger than that of the standard FA model, as spins can flip if only one nearest neighbor is in the excited state, in contrast to the FA where spins can also flip if more than one neighbor is excited. It is also less constrained than the PXP, or two-spin facilitated FA, model which requires both neighbors to be simultaneously in the same state. As such it shares features with both these models. The constraint imposes a conservation law, that of the total number of domain walls, breaking configuration space into a number of disconnected components that scales with system size. This contrasts to the FA model where all but one configuration are dynamically connected, and is closer to the PXP where configuration space is also disconnected. The distinction with the PXP is that in the XOR-FA the conserved quantity is global, while in the PXP there are many local conserved quantities and configuration space is broken into exponentially many disconnected components. This less severe disconnection of state space makes the dynamics within connected components in the XOR-FA still interesting as there is scope for large dynamical fluctuations (in contrast to the PXP). An interesting question would be to study the analogous problem in constrained lattice gases, considering variants of the Kob-Andersen (KA) model [4] or the triangular lattice gas (TLG) [86] but where constraints are "selective" (as in Ref. [87]). The KA and the TLG are models where particles can hop as long as a minimum number of neighbors are unoccupied (cf. the FA model where at least one neighbor has to be up for the spin to flip). A "selective" [87] version of such models where the number of required unoccupied neighbors is fixed would be slightly more constrained, just like the XOR-FA is slightly more constrained than the FA model. Studying such models in the manner of the current paper would require, however, to extend the tensor network methods to dimensions higher than one.

The XOR-FA is a "thermal" model, in the sense that it obeys detailed balance with respect to the Boltzmann distribution of a noninteracting binary gas, where each excited spin costs a unit of energy, and subject to the conservation of the number of DWs. In analogy to other KCMs like the FA and East models [17], the XOR-FA has a trivial (up to the conservation law) thermodynamics, but complex relaxation dynamics due to the constraint. The conservation law allows the XOR-FA to be represented in terms of the dynamics of its DWs. This establishes a connection to exclusion processes. At infinite temperature the XOR-FA can be mapped via a duality transformation to the SEP [39,43,44], while at finite temperature this mapping leads to a dynamics of hopping DWs with long-range interactions (in contrast to the original spin representation which is always local).

Like in other KCMs [19,55,64,81], we have shown here that the XOR-FA has a trajectory level phase transition between active and inactive dynamical phases. We have proved this to high accuracy via numerical MPS. This adds to recent work [60–62] demonstrating the effectiveness of numerical tensor network methods for studying the full counting statistics of stochastic systems. One of the many advantages of this approach is the direct availability of spatial structural information in the various dynamical phases. For the XOR-FA we find spatial structure as expected from its connection to the SEP [88,89]: a change from repulsion of DWs in the active phase, maximizing the possibility of motion, through structureless configurations in the typical equilibrium dynamics, to DW clustering in the inactive phase. Away from the strict SEP limit, these structures remain overall, with further richness due to energetics.

Here we have studied the XOR-FA in one-dimension. It is easy to generalise the model to higher dimensions, once again motivated for example by the problem of atoms interacting strongly in Rydberg states. For example, in the blockade regime, when going from one dimension to a, say, two dimensional square lattice, the PXP model becomes the hard square model [90]. Similarly, in the antiblockade regime of Rydbergs, the XOR-FA would generalise to a model that is less constrained than that of hard squares, but more constrained than the two-spin facilitated FA model in two dimensions [17]. From the results here it is safe to speculate that such higher dimensional generalisations of the XOR-FA will also display very rich dynamics. Additionally, the structure of the ground state observed at $s > 0$ is reminiscent of the localized ground states found in quantum KCMs [22], which has dramatic consequences for the quantum dynamics of the model. It may be interesting to investigate whether an analogous localization can be characterized in the ground state of the XOR-FA model in the inactive regime.

ACKNOWLEDGMENTS

We thank Oriane Blondel, Martin Evans, and Sarang Gopalakrishnan for useful comments on the manuscript. We acknowledge financial support from EPSRC Grant No. EP/R04421X/1 and the Leverhulme Trust Grant No. RPG-2018-181. I.L. acknowledges support from the DFG through SPP 1929 (GiRyd) and by the Wissenschaftler-Rückkehrprogramm GSO/CZS of the Carl-Zeiss-Stiftung and

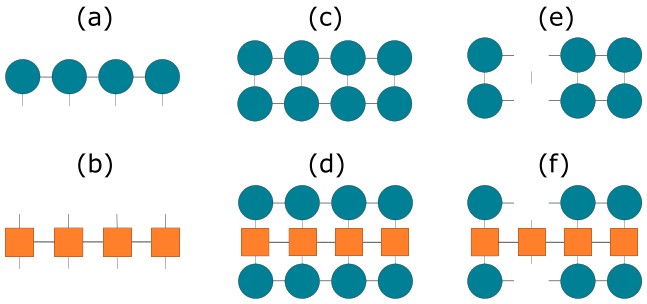


FIG. 8. Tensor networks in the vMPS. The diagrammatic representations of tensor networks in the variational MPS search. (a) The state $|\Psi\rangle$ as an MPS. (b) The Hamiltonian $\mathbb{H}_s$ as an exact MPO. (c) The inner product $\langle\Psi|\Psi\rangle$. (d) The expectation value $\langle\Psi|\mathbb{H}_s|\Psi\rangle$. (e) The effective norm $\mathcal{N}_k$. (e) The effective Hamiltonian $\mathcal{H}_k$.

the German Scholars Organization e.V. M.C.B. acknowledges support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC-2111–390814868. We are grateful for access to the University of Nottingham Augusta HPC service.

APPENDIX

1. Variational MPS

The vMPS algorithm used in Sec. IV goes as follows. We have some MPS, $|\Psi_{\text{guess}}\rangle$ as defined in Eq. (24), which is our guess to the true ground state. See Fig. 8(a) for the diagrammatic representation, where the shapes represent the local tensors and the legs represent contractions over tensors. One can then write the Hamiltonian in Eq. (21) as a matrix product operator (MPO) [74,75],

$$\mathbb{H}_s = \sum_{i_1,\ldots,i_N} \sum_{j_1,\ldots,j_N} \mathrm{Tr}(M_1^{i_1 j_1} M_2^{i_2 j_2} \ldots M_N^{i_N j_N}) \\ |i_1\, i_2 \ldots i_N\rangle \langle j_1\, j_2 \ldots j_N|\,, \tag{A1}$$

where M_k is a rank-4 tensor with dimensions $d\times d\times D_H\times D_H$. The locality of $\mathbb{H}_s$ allows us to exactly represent it in MPO form with only a small bond dimension $D_H = 4$, where each tensor is identical. As with the MPS, this can be represented in the diagrammatic form Fig. 8(b). The energy of the guess with respect to Eq. (A1) is then given by

$$E_{\text{guess}} = \frac{\langle\Psi_{\text{guess}}|\mathbb{H}_s|\Psi_{\text{guess}}\rangle}{\langle\Psi_{\text{guess}}|\Psi_{\text{guess}}\rangle} \geqslant E_s, \tag{A2}$$

where $E_s = -\theta(s)$ is the true ground state energy. The expectation value and inner product can be expressed as tensor network contractions, as illustrated in Figs. 8(c) and 8(d). This allows for an efficient calculation that exploits the MPS structure.

At each step, a single tensor is optimized by minimizing equation Eq. (A2) with respect to A_k, which gives

$$\mathcal{H}_k A_k = E_{\text{guess}}\, \mathcal{N}_k A_k, \tag{A3}$$

where $\mathcal{N}_k$ and $\mathcal{H}_k$ are the effective norm and effective Hamiltonian computed by contracting over all tensors except for A_k within $\langle\Psi_{\text{guess}}|\Psi_{\text{guess}}\rangle$ and $\langle\Psi_{\text{guess}}|\mathbb{H}_s|\Psi_{\text{guess}}\rangle$, respectively. Both effective operators can be expressed also as tensor networks, as shown in Figs. 8(e) and 8(f). If we treat A_k as a $(D^2 d)$-vector and $\mathcal{N}_k$, $\mathcal{H}_k$ as $(D^2 d)$-matrices, then Eq. (A3) is simply a generalized eigenvalue problem which should be solved using a sparse eigensolver to keep the computational scaling to $O(D^3)$. The solution to Eq. (A3) with the smallest E_{guess} is our new choice for A_k.

We sweep back and forth through each tensor in the MPS, applying local updates in the way detailed above. Since each local minimization can be solved exactly, the energy can only decrease at each step, and the algorithm is guaranteed to converge. However, it may do so to a local minimum. As a quality criterion, we require that the (efficiently computable) variance of the Hamiltonian in the guess state falls below some specified value $\langle\mathbb{H}_s^2\rangle - \langle\mathbb{H}_s\rangle^2 < \epsilon$, where here $\langle\cdot\rangle$ denotes an expectation with respect to the $|\Psi_{\text{guess}}\rangle$. If $|\Psi_{\text{guess}}\rangle$ does not satisfy this criterion for a run of the algorithm at some bond dimension D, then we run it again with an MPS with a higher bond dimension, typically using the previous run as our initial guess.

2. MPS steady-state solutions

Here we follow the workings of Ref. [54] to present an MPS steady-state solution to Eq. (13) in the thermodynamic limit, $N\to\infty$, with which we can determine the equilibrium properties of the model. We consider the XOR-FA model with N sites and PBC, and describe the probability vector $|\nu\rangle$ as an MPS, cf. Eq. (24), where of course we have translational invariance and each tensor is identical, $A_k = A$ for all k. Let us now guess the solution

$$A^0 = \begin{bmatrix} 0 & 0 \\ 1-p_0 & p_0 \end{bmatrix}, \quad A^1 = \begin{bmatrix} p_1 & 1-p_1 \\ 0 & 0 \end{bmatrix}, \tag{A4}$$

where $0 < p_0,\ p_1 < 1$. We first require that our solution annihilates the generator, $\mathbb{W}\,|\nu\rangle = 0$. It is easy to verify that this is the case if we have

$$\frac{p^1}{p^0} = \frac{c}{1-c}. \tag{A5}$$

Additionally, we require that $|\nu\rangle$ is normalized. The partition function is calculated by taking the inner product with the flat state,

$$Z = \langle -|\nu\rangle = \mathrm{Tr}(M^N), \tag{A6}$$

where $M^N = A^0 + A^1$. It is easy to show via induction that

$$M^t = (2-p_0-p_1)^{-1}\left(\begin{bmatrix} 1-p_0 & 1-p_1 \\ 1-p_0 & 1-p_1 \end{bmatrix} \right. \\ \left. + (p_0+p_1-1)^t \begin{bmatrix} 1-p_1 & -(1-p_1) \\ -(1-p_0) & 1-p_0 \end{bmatrix}\right). \tag{A7}$$

It follows that the partition function is already normalized in the infinite limit

$$\lim_{N\to\infty} Z = 1. \tag{A8}$$

The average DW density $\langle n_{\text{DW}}\rangle$ can be calculated as the DW occupation between any two neighboring sites in the

lattice,

$$\langle n_{\mathrm{DW}}\rangle = \frac{1}{Z}\langle -|[n_i(1-n_{i+1})+(1-n_i)n_{i+1}]|\nu\rangle = \frac{1}{Z}\mathrm{Tr}([A^1A^0+A^0A^1]M^{N-2}). \tag{A9}$$

Taking the infinite limit, one finds that

$$\lim_{N\to\infty}\langle n_{\mathrm{DW}}\rangle = \frac{2(1-p_0)(1-p_1)}{2-p_0-p_1}. \tag{A10}$$

We can now determine the p_0, p_1 necessary to have the required DW density by substituting Eq. (A5) into Eq. (A10) and solving as a quadratic equation. Thus, we have found an MPS steady-state solution with bond dimension 2, which also has the required DW density in the thermodynamic limit.

To calculate other local observables, we can again simply use the procedure described above. The average excitation density can be calculated using a local MPO on just one site,

$$\lim_{N\to\infty}\langle n\rangle = \lim_{N\to\infty}\frac{1}{Z}\langle -|n_i|\nu\rangle = \lim_{N\to\infty}\frac{1}{Z}\mathrm{Tr}(A^1M^{N-1}) = \frac{1-p_0}{2-p_0-p_1}. \tag{A11}$$

Likewise, the average dynamical activity can be calculated as the escape rate of just a single site, which can be calculated using the three-body operator

$$r_i = c[n_{i-1}(1-n_i)(1-n_{i+1})+(1-n_{i-1})(1-n_i)n_{i+1}] + (1-c)[n_{i-1}n_i(1-n_{i+1})+(1-n_{i-1})n_in_{i+1}]. \tag{A12}$$

After a lengthy calculation, we find

$$\lim_{N\to\infty}\langle k\rangle = \lim_{N\to\infty}\frac{1}{Z}\langle -|r_i|\nu\rangle = \frac{2(1-p_0)(1-p_1)}{2-p_0-p_1}[(1-c)p_1+cp_0]. \tag{A13}$$

We compare our analytical results to numerical data obtained for large, but finite system sizes in Fig. 1. Both results show excellent agreement.

[1] G. H. Fredrickson and H. C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984).
[2] R. G. Palmer, D. L. Stein, E. Abrahams, and P. W. Anderson, Phys. Rev. Lett. **53**, 958 (1984).
[3] J. Jäckle and S. Eisinger, Z. Phys. B **84**, 115 (1991).
[4] W. Kob and H. C. Andersen, Phys. Rev. E **48**, 4364 (1993).
[5] D. S. Rokhsar and S. A. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988).
[6] F. Alet, J. L. Jacobsen, G. Misguich, V. Pasquier, F. Mila, and M. Troyer, Phys. Rev. Lett. **94**, 235702 (2005).
[7] F. Alet, Y. Ikhlef, J. L. Jacobsen, G. Misguich, and V. Pasquier, Phys. Rev. E **74**, 041124 (2006).
[8] O. F. Syljuåsen, Phys. Rev. B **73**, 245105 (2006).
[9] C. L. Henley, Annu. Rev. Condens. Matter Phys. **1**, 179 (2010).
[10] T. Oakes, J. P. Garrahan, and S. Powell, Phys. Rev. E **93**, 032129 (2016).
[11] T. Oakes, S. Powell, C. Castelnovo, A. Lamacraft, and J. P. Garrahan, Phys. Rev. B **98**, 064302 (2018).
[12] S. Vijay, J. Haah, and L. Fu, Phys. Rev. B **92**, 235136 (2015).
[13] S. Vijay, J. Haah, and L. Fu, Phys. Rev. B **94**, 235157 (2016).
[14] A. Prem, J. Haah, and R. Nandkishore, Phys. Rev. B **95**, 155133 (2017).
[15] H. Song, A. Prem, S.-J. Huang, and M. A. Martin-Delgado, Phys. Rev. B **99**, 155118 (2019).
[16] R. M. Nandkishore and M. Hermele, Annu. Rev. Condens. Matter Phys. **10**, 295 (2019).
[17] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[18] J. P. Garrahan, P. Sollich, and C. Toninelli, in *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media*, International Series of Monographs on Physics, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos (Oxford University Press, Oxford, UK, 2011).
[19] J. P. Garrahan, Physica A **504**, 130 (2018).
[20] M. van Horssen, E. Levi, and J. P. Garrahan, Phys. Rev. B **92**, 100305(R) (2015).
[21] Z. Lan, M. van Horssen, S. Powell, and J. P. Garrahan, Phys. Rev. Lett. **121**, 040603 (2018).
[22] N. Pancotti, G. Giudice, J. I. Cirac, J. P. Garrahan, and M. C. Bañuls, Phys. Rev. X **10**, 021051 (2020).
[23] B. Olmos, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. Lett. **109**, 020403 (2012).
[24] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[25] G. Biroli and J. P. Garrahan, J. Chem. Phys. **138**, 12A301 (2013).
[26] A. Browaeys and T. Lahaye, Nat. Phys. **16**, 132 (2020).
[27] P. Fendley, K. Sengupta, and S. Sachdev, Phys. Rev. B **69**, 075106 (2004).
[28] I. Lesanovsky, Phys. Rev. Lett. **106**, 025301 (2011).
[29] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Nature **551**, 579 (2017).
[30] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Nat. Phys. **14**, 745 (2018).
[31] C. Ates, T. Pohl, T. Pattard, and J. M. Rost, Phys. Rev. Lett. **98**, 023002 (2007).
[32] T. Amthor, C. Giese, C. S. Hofmann, and M. Weidemüller, Phys. Rev. Lett. **104**, 013001 (2010).
[33] I. Lesanovsky and J. P. Garrahan, Phys. Rev. Lett. **111**, 215305 (2013).
[34] I. Lesanovsky and J. P. Garrahan, Phys. Rev. A **90**, 011603(R) (2014).
[35] M. Hoening, W. Abdussalam, M. Fleischhauer, and T. Pohl, Phys. Rev. A **90**, 021603(R) (2014).
[36] A. Urvoy, F. Ripka, I. Lesanovsky, D. Booth, J. P. Shaffer, T. Pfau, and R. Löw, Phys. Rev. Lett. **114**, 203002 (2015).
[37] M. M. Valado, C. Simonelli, M. D. Hoogerland, I. Lesanovsky, J. P. Garrahan, E. Arimondo, D. Ciampini, and O. Morsch, Phys. Rev. A **93**, 040701(R) (2016).

[38] M. Marcuzzi, J. Minář, D. Barredo, S. de Léséleuc, H. Labuhn, T. Lahaye, A. Browaeys, E. Levi, and I. Lesanovsky, Phys. Rev. Lett. **118**, 63606 (2017).

[39] M. Ostmann, M. Marcuzzi, J. P. Garrahan, and I. Lesanovsky, Phys. Rev. A **99**, 060101(R) (2019).

[40] T. M. Wintermantel, Y. Wang, G. Lochead, S. Shevate, G. K. Brennen, and S. Whitlock, Phys. Rev. Lett. **124**, 070503 (2020).

[41] P. P. Mazza, R. Schmidt, and I. Lesanovsky, Phys. Rev. Lett. **125**, 033602 (2020).

[42] J. P. Garrahan and I. Lesanovsky, Phys. Rev. Lett. **104**, 160601 (2010).

[43] M. Suzuki, Prog. Theor. Phys. **46**, 1337 (1971).

[44] U. Borla, R. Verresen, F. Grusdt, and S. Moroz, Phys. Rev. Lett. **124**, 120503 (2020).

[45] N. Shiraishi and T. Mori, Phys. Rev. Lett. **119**, 030601 (2017).

[46] S. Moudgalya, S. Rachel, B. A. Bernevig, and N. Regnault, Phys. Rev. B **98**, 235155 (2018).

[47] T. Iadecola and M. Schecter, Phys. Rev. B **101**, 024306 (2020).

[48] Z.-C. Yang, F. Liu, A. V. Gorshkov, and T. Iadecola, Phys. Rev. Lett. **124**, 207602 (2020).

[49] D. K. Mark, C.-J. Lin, and O. I. Motrunich, Phys. Rev. B **101**, 195131 (2020).

[50] S. Gopalakrishnan and B. Zakirov, Quantum Sci. Technol. **3**, 044004 (2018).

[51] R. A. Blythe and M. R. Evans, J. Phys. A **40**, R333 (2007).

[52] K. Mallick, Physica A **418**, 17 (2015).

[53] M. R. Evans, Europhys. Lett. **36**, 13 (1996).

[54] V. Sidoravicius, D. Surgailis, and M. Vares, Markov Process Relat. **4**, 131 (1998).

[55] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 75007 (2009).

[56] A stretched exponential has the form $P(t) \sim \exp[-(t/t_0)^b]$. For $c = 0.1$ and $n_{\mathrm{DW}} = 1/4$, we estimate the "stretching" parameter to be $b = 0.488$.

[57] The diffusion constant used in Eq. (17) is only true for small c. For $c = 0.5$, the diffusion constant can be estimated to be $D_c \approx c$. In this way, the estimate Eq. (17) only accounts for the results qualitatively where we have to re-scale by some constant.

[58] H. Touchette, Phys. Rep. **478**, 1 (2009).

[59] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).

[60] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).

[61] P. Helms, U. Ray, and G. Kin-Lic Chan, Phys. Rev. E **100**, 022101 (2019).

[62] P. Helms and G. Kin-Lic Chan, Phys. Rev. Lett. **125**, 140601 (2020).

[63] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).

[64] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).

[65] C. Maes, Phys. Rep. **850**, 1 (2020).

[66] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).

[67] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004).

[68] U. Schollwöck, Ann. Phys. **326**, 96 (2011).

[69] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 030504 (2008).

[70] M. B. Hastings, J. Stat. Mech. (2007) P08024.

[71] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).

[72] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[73] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).

[74] I. P. McCulloch, J. Stat. Mech. (2007) P10014.

[75] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, New J. Phys. **12**, 025012 (2010).

[76] F. Verstraete, V. Murg, and J. I. Cirac, Adv. Phys. **57**, 143 (2008).

[77] D. Pérez-García, M. M. Wolf, M. Sanz, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 167202 (2008).

[78] S. Singh, R. N. C. Pfeifer, and G. Vidal, Phys. Rev. A **82**, 050301(R) (2010).

[79] D. Pérez-García, M. Sanz, C. E. González-Guillén, M. M. Wolf, and J. I. Cirac, New J. Phys. **12**, 25010 (2010).

[80] N. Chepiga and F. Mila, SciPost Phys. **6**, 33 (2019).

[81] T. Bodineau, V. Lecomte, and C. Toninelli, J. Stat. Phys. **147**, 1 (2012).

[82] This only works because here we have OBC. For PBC, one would have to consider the spatial correlations to learn about the structure of the state.

[83] A. De Pasquale, G. Costantini, P. Facchi, G. Florio, S. Pascazio, and K. Yuasa, Eur. Phys. J.: Spec. Top. **160**, 127 (2008).

[84] J. P. Garrahan and D. Chandler, Phys. Rev. Lett. **89**, 035704 (2002).

[85] J. P. Garrahan and D. Chandler, Proc. Natl. Acad. Sci. USA **100**, 9710 (2003).

[86] J. Jackle and A. Kronig, J. Phys.: Condens. Matter **6**, 7633 (1994).

[87] M. Sellitto, Phys. Rev. E **100**, 040102(R) (2019).

[88] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).

[89] V. Lecomte, J. P. Garrahan, and F. van Wijland, J. Phys. A **45**, 175001 (2012).

[90] S. Ji, C. Ates, and I. Lesanovsky, Phys. Rev. Lett. **107**, 060406 (2011).

Chapter 5

Slow dynamics and large deviations in classical stochastic Fredkin chains

The following work is from the publication *"Slow dynamics and large deviations in classical stochastic Fredkin chains"* by Luke Causer, Juan P. Garrahan and Austen Lamacraft, published in Physical Review E 106, 014128 (2022).

This paper generalizes the quantum many-body system known "Fredkin spin chains" to classical stochastic dynamics. In a classical sense, it can be considered an ASEP equipped with kinetic constraints which only allow jumps if the neighbouring sites take certain configurations. The steady state properties and non-equilibrium dynamics are uncovered, as was done for the XORFA model. As is the case for the ground state of the quantum model, the steady state admits an equilibrium phase transition. Furthermore, when quenched from particular initial states, the Fredkin model undergoes hierarchical relaxation, similar to other KCMs such as the *East* model.

The dynamical LDs for the model reveals dynamical phase transitions for each of the equilibrium phases. Where possible, a finite-size scaling analysis is done, which indicate that for each equilibrium phase, there exists a first-order phase transition. Furthermore, a scaling exponent in system size is extracted where possible, demonstrating different university classes which are intrinsically related to the equilibrium fluctuations. One of the equilibrium phases has a hierarchy of dynamical phase transitions, each corresponding to intermediate levels of dynamical activity. Indeed, these phases are related to the hierarchical relaxation of the model due to its strong localization properties.

The work done here is an important step to understanding the quantum dynamics of a "tilted" Fredkin chain, similar to that done for the "Quantum East model" [146]. While not shown in this thesis, the quantum version of the model has various properties related to many-body localization, along with anomalous eigenstates with small amounts of entropy and non-thermal properties (often referred to as "many-body quantum scars"). This behaviour can be explained by the localization properties associated with the inactive dynamical phases, and the strong kinetic constraint.

Corrections to the manuscript:

1. The label of the x-axis in Figure 3(b) should read "Site, i". The caption should read $\langle n^{\mathrm{DW}}_{N/2+i} \rangle$.

2. The sum in Equation A1 should be in the range $i = 1, \ldots, N$.

Slow dynamics and large deviations in classical stochastic Fredkin chains

Luke Causer,[1,2] Juan P. Garrahan,[1,2] and Austen Lamacraft[3]
[1]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*
[2]*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, United Kingdom*
[3]*TCM Group, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, United Kingdom*

(Received 24 March 2022; accepted 5 July 2022; published 21 July 2022)

The Fredkin spin chain serves as an interesting theoretical example of a quantum Hamiltonian whose ground state exhibits a phase transition between three distinct phases, one of which violates the area law. Here we consider a classical stochastic version of the Fredkin model, which can be thought of as a simple exclusion process subject to additional kinetic constraints, and study its classical stochastic dynamics. The ground-state phase transition of the quantum chain implies an equilibrium phase transition in the stochastic problem, whose properties we quantify in terms of numerical matrix product states (MPSs). The stochastic model displays slow dynamics, including power-law decaying autocorrelation functions and hierarchical relaxation processes due to exponential localization. Like in other kinetically constrained models, the Fredkin chain has a rich structure in its dynamical large deviations—which we compute accurately via numerical MPSs—including an active-inactive phase transition and a hierarchy of trajectory phases connected to particular equilibrium states of the model. We also propose, via its height field representation, a generalization of the Fredkin model to two dimensions in terms of constrained dimer coverings of the honeycomb lattice.

DOI: 10.1103/PhysRevE.106.014128

I. INTRODUCTION

The Fredkin spin chain [1,2] is a one-dimensional lattice model with local three-body interactions, whereby hardcore particles can hop to adjacent sites if allowed by constraints involving next-to-nearest neighbors. This model has been of interest in the quantum many-body community over the last few years for a number of reasons. In its original formulation [1,2], the Fredkin chain can be expressed exactly as an equal superposition of all Dyck paths, i.e., random walk (RW) excursions, with appropriate endpoints, with an entanglement entropy which scales logarithmically in system size, thus violating the area law [1–4]. Furthermore, the model has slow unitary evolution [5–8] due to dynamical jamming. With the addition of particular potential energy terms, the model features a ground-state phase transition between states of bounded and extensive entanglement entropy [9,10]. These interesting properties have brought about further studies into generalized Fredkin models [11], including versions which present quantum scars [12].

Dynamical constraints, such as those present in the Fredkin model, are responsible more generally for many interesting phenomena in many-body dynamics. A striking example of this are the kinetically constrained models (KCMs) of structural glasses [13,14]—simple lattice models equipped with local dynamical constraints, leading to slow relaxation and dynamical heterogeneity [15,16]. Such models can also be considered as systems under closed unitary [17–20] and open dissipative [21–23] quantum dynamics. A recent example of these is the quantum PXP model [24,25] of Rydberg atoms in optical lattices under blockade conditions, which has been shown to exhibit nonthermal eigenstates (often called quantum scars [26]). Another area where dynamical constraints lead to interesting nonequilibrium dynamics is in deterministic cellular automata [27–43] (for a review, see Ref. [44]). Recently, cellular automata circuits have also been used to study Fredkin-like systems [45,46], revealing a universality class of hydrodynamics. While the connection of the Fredkin quantum spin chain to stochastic dynamics has been previously mentioned [2], it has not yet been extensively explored (other than briefly in Ref. [47]). Here we provide such systematic study of both typical dynamics and rare fluctuations.

Classically, the Fredkin model resembles the simple exclusion process (or SEP, for reviews see Refs. [48,49]). Like the SEP, it describes particles hopping stochastically to neighboring empty lattice sites with at most one particle per site. The key difference is the presence of further local kinetic constraints to motion. These, together with specific boundary conditions, specifically that of an open segment with fixed boundaries, restrict the dimensionality of the state space. For example, for a length $N = 2M$ chain half filled with M particles, the dimensionality is the Catalan number $C_N = \frac{1}{M+1}\binom{2M}{M}$ rather than the binomial coefficient $\binom{2M}{M}$. Although the difference in configurational entropy is not extensive, this constrained state space plays an important role in the dynamics, as we explain below.

SEPs and KCMs display interesting dynamical properties which can be studied with large deviation (LD) methods (for reviews, see Refs. [50–53]). A central result in the dynamics of these systems is the existence of phase transitions in the

space of trajectories, indicated by singularities in the LD functions that quantify the dynamical fluctuations in the long-time limit, both in terms of time-integrated currents [54–57], or dynamical activities [58–62]. In the case of the Fredkin model, a preliminary study [47] indicated that it also displays LD transitions. Here we make this finding concrete by studying LD functions using matrix product states (MPSs).

The paper is organized as follows. In Sec. II, we start by defining the model and reviewing its basic properties. We highlight its relationship to Catalan combinatorics and RW excursions [63]. In Sec. III, we consider the equilibrium states which follow from the properties of the ground state of the quantum problem [1,9]. We study the properties of the equilibrium phases in detail by means of numerical density matrix renormalization group (DMRG) [64]. An interesting observation is that there are three distinct equilibrium phases and a transition between them, despite the fact that this is a one-dimensional system with local dynamical rules. This apparent contradiction with the Landau principle is a consequence of the constrained configuration space of the model. In Sec. IV, we study the relaxation dynamics. As in the case of the quantum model [8], the stochastic Fredkin spin chain exhibits slow dynamics. We provide evidence for power-law decaying autocorrelations, and for a pattern of hierarchical relaxation when quenched from extremal initial states into the different equilibrium phases. In Sec. V, we study the LD statistics of dynamical observables by means of numerical MPSs. As in other constrained models, the phase transitions at the LD level underpin the slow dynamics and fluctuations seen in typical relaxation trajectories. We reveal the existence of an active-inactive transition, as in other KCMs, but also a hierarchy of trajectory transitions connected to hierarchical relaxation dynamics. In Appendix B, we speculate on a possible generalization of the Fredkin model to a two-dimensional setting defined in terms of fully packed dimers on the honeycomb lattice (that is, rhombus coverings of the plane). We give our conclusions in Sec. VI.

II. MODEL

The Fredkin model is defined in terms of particles hopping on a lattice of N sites with binary occupation, $n_j = 0$ (for empty or down) or 1 (for occupied or up) with $j = 1, \cdots, N$. The system evolves under stochastic continuous-time Markov dynamics with generator

$$\mathbb{W} = \sum_{i=2}^{N-2} f_i\{c[\sigma_i^+\sigma_{i+1}^- - (1-n_i)n_{i+1}] + (1-c)[\sigma_i^-\sigma_{i+1}^+ - n_i(1-n_{i+1})]\}, \tag{1}$$

where $\sigma_i^{\pm}$ are Pauli creation and annihilation operators on site i. The factor in curly brackets in each term is the same as the local generator of the asymmetric SEP (ASEP) [48,49], with rates for hops to the left or right given by c and $1-c$, respectively. What distinguishes the Fredkin model from the ASEP is the kinetic constraint

$$f_i = n_{i-1} + (1 - n_{i+2}), \tag{2}$$

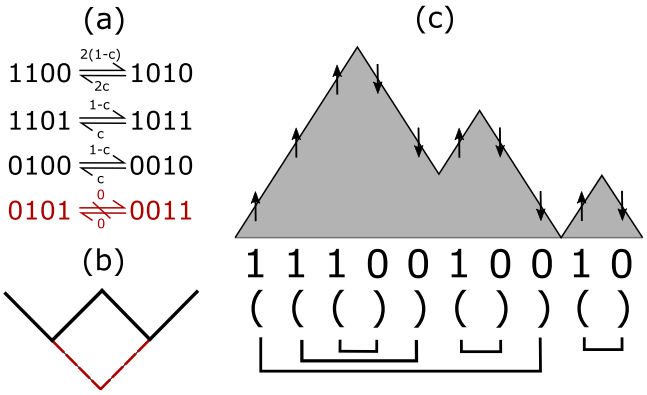


FIG. 1. Fredkin spin chains. (a) The local stochastic transition rates for neighboring occupied and unoccupied sites, given by all choices of their neighbors. The fourth transition is not allowed. (b) The disallowed configuration change in the height representation. The troughs $(\cdots 0011 \cdots)$ are locally immobile. (c) An example configuration in the chosen symmetry sector. The top shows the RW representation of the height field, which must always satisfy $h > 0$. The middle is the corresponding particle representation. The bottom is in terms of Dyck words, where opening "(" must always be matched with a closing ")".

which means that hopping between sites i and $i+1$ is not allowed if $n_{i-1} = 0$ and $n_{i+2} = 1$, see Fig. 1(a) [65]. In Eq. (1), we are considering open boundary conditions on a segment $[1, N]$ with no injection/ejection of particles at the boundaries. The fixed sites at the edges, which are not acted on by the generator, we fix to be $n_1 = 1$ and $n_N = 0$.

Note that at $c = 1/2$, Eq. (1) is equivalent to the quantum Hamiltonian of the original Fredkin model defined in Ref. [1], up to a minus sign and boundary terms. For $c \neq 1/2$, the generator Eq. (1) obeys detailed balance despite the asymmetry in the hopping rates [66], meaning that for all values of c we expect to find an equilibrium stationary state of $\mathbb{W}$. Notice that under a similarity transformation (see below), it becomes equivalent to the deformed Fredkin model of Ref. [9].

The model discussed here has various symmetries. The most obvious one is the conservation of the total number or particles (or occupied sites): $M = \sum_i n_i$. This property is shared with the SEP. The constraint Eq. (2) gives rise to a further subdivision of each subspace of fixed M, which is most easily understood by a representation of the allowed moves in terms of matched brackets [1]. In this representation, particles and holes correspond to opening and closing parentheses, and the dynamics respects normal matching rules. Thus the move

$$\cdots 0101 \cdots \longleftrightarrow \cdots 0011 \cdots, \tag{3}$$

$$\cdots)()(\cdots \longleftrightarrow \cdots))((\cdots \tag{4}$$

is forbidden because both sides cannot simultaneously be matched configurations [this forbidden transition is shown in Fig. 1(b)]. Thus a complete specification of a subspace of allowed configurations involves specifying the M pairs of matched brackets, a unmatched opening brackets (particles), and b unmatched closing brackets (holes) for a total $N = 2M + a + b$.

For concreteness, here we will focus on the case of half filling by fully matched particles and holes, i.e., $M = N/2$ $a = b = 0$. In this case, the accumulated number of particles starting from the left is never smaller than the accumulated number of holes (that is, the sector that is dynamically connected to having all particles to the left and all holes to the right, see below), cf. Ref. [1]. We call this sector $\mathcal{D}$.

It is convenient to represent a configuration $x = n_{1:N}$ also in terms of a *height field* defined as

$$h_i(x) = \sum_{j=1}^{i} Z_j = h_{i-1}(x) + Z_i, \tag{5}$$

with boundary condition $h_0 = 0$, and where $Z_i = 2n_i - 1$. For all configurations $x \in \mathcal{D}$, we have $h_i(x) \geqslant 0$ for all i. If we think of the space direction as time and a particle (hole) representing a step up (down), then $\mathcal{D}$ is the space of all paths that correspond to *RW excursions* [63], that is, random paths that return to the origin while never crossing the horizontal axis. (In contrast, for the SEP in the height representation at half filling, the space of dynamically connected configurations is that of RW bridges, which are also constrained to return to the origin but can cross the horizontal axis.). Excursions are also known as Dyck paths. An example configuration is shown in Fig. 1(c) with each of the representations.

III. EQUILIBRIUM STATICS

To determine the equilibrium properties of the model, we need to find the state $|\mathrm{ss}\rangle$ annihilated by Eq. (1). Let us consider as an observable the area under the height profile of a configuration x,

$$A(x) = \sum_{i=1}^{N} h_i(x) = \sum_{i=1}^{N} (N + 1 - i)\, Z_i. \tag{6}$$

It is then easy to see that the the stationary state to the dynamics Eq. (1) is given by [9]

$$|\mathrm{ss}\rangle = \mathcal{N}_c \sum_{x\in\mathcal{D}} \left(\frac{c}{1-c}\right)^{\frac{1}{2}A(x)} |x\rangle\,, \tag{7}$$

with $\mathcal{N}_c$ a c-dependent normalization constant to make $\langle -|\mathrm{ss}\rangle = 1$, where $\langle -| = \sum_{x\in\mathcal{D}} \langle x|$ is called the *flat state*.

The connection to RW excursions means that this probability distribution is related to the Airy function [63,67]. The properties of the stationary state at arbitrary c can also be understood from the properties of the ground state of the corresponding quantum model [1,9]. That is, if we perform a similarity transformation of Eq. (1) (cf. the ASEP with the same boundary conditions, e.g., Ref. [68]),

$$\mathbb{H} = -\mathbb{P}^{-1/2}\, \mathbb{W}\, \mathbb{P}^{1/2}, \tag{8}$$

where $\mathbb{P}^{1/2}$ is the diagonal matrix of the square root of configuration probabilities,

$$\langle x|\mathbb{P}^{1/2}|x\rangle = \mathcal{N}_c^{1/2} \left(\frac{c}{1-c}\right)^{\frac{1}{4}A(x)}, \tag{9}$$

we get the Hamiltonian

$$\mathbb{H} = -\sum_{i=2}^{N-2} f_i[\sqrt{c(1-c)}(\sigma_i^+\sigma_{i+1}^- + \sigma_i^-\sigma_{i+1}^+) - c(1-n_i)n_{i+1} - (1-c)n_i(1-n_{i+1})], \tag{10}$$

whose ground state is $|\psi\rangle = \mathbb{P}^{1/2} \sum_{x\in\mathcal{D}} |x\rangle$. The transformation to a Hermitian form shows that, despite the asymmetric hopping when $c \neq 1/2$, the Fredkin model obeys detailed balance and consequently the stationary state Eq. (7) is an equilibrium one.

The properties of the ground state of Eq. (10) are well understood from previous studies [10,11]. Here we restate them from the point of view of the equilibrium state of the stochastic model, using MPSs (see reviews, e.g., Refs. [69–71]).

A. Exact equilibrium MPS at $c = 1/2$

From the connection to RW excursions at $c = 1/2$ the equilibrium state $|\mathrm{ss}\rangle$ can be written exactly as an MPS,

$$|\mathrm{ss}\rangle = \sum_{\{n_{1:N}\}} \left(i|B_{n_1}^{(1)} \cdots B_{n_N}^{(N)}|f\right)|n_{0:1}\rangle, \tag{11}$$

where $B_n^{(j)}$ are site-dependent tensors and $(i|$ and $|f)$ appropriate boundary vectors in the auxiliary (or bond) space of the MPS (we use rounded brackets to distinguish them from vectors in configuration space).

Consider first the slightly simpler problem of the symmetric SEP (SSEP), whose generator is given by an operator like Eq. (1) but without a constraint, $f_i = 1$. If we consider the same boundary conditions as for the Fredkin model, but with extra terms in Eq. (1) that allow particle hops between sites $j = 1, 2$ and $N-1, N$ (no longer prevented in the absence of a constraint), then the SSEP configurations at half filling are those of RW bridges. If the height field h_j describes the position of the RW after step j, the exact transition probabilities at step j for generating bridges of N steps are

$$T_j^{\mathrm{br}}(h \to h \pm 1) = \frac{1}{2}\left(1 \mp \frac{h}{N+1-j}\right) \tag{12}$$

for $|h| \leqslant N + 1 - j$, or zero otherwise. (These are obtained from the naive symmetric RW transition probabilities via a Doob transform, see, e.g., Ref. [72].) The equilibrium MPS for the SSEP is then given by the $(2N+1) \times (2N+1)$ matrices

$$B_0^{(j),\mathrm{SSEP}} = \sum_{h=-N}^{N} |h)(h-1|\, T_j^{\mathrm{br}}(h \to h-1), \tag{13}$$

$$B_1^{(j),\mathrm{SSEP}} = \sum_{h=-N}^{N} |h)(h+1|\, T_j^{\mathrm{br}}(h \to h+1), \tag{14}$$

with boundaries $(i| = (0|$ and $|f) = |0)$. It is easy to see that the matrices above satisfy $B_0^{(j),\mathrm{SSEP}} B_1^{(j+1),\mathrm{SSEP}} - B_1^{(j),\mathrm{SSEP}} B_0^{(j+1),\mathrm{SSEP}} = 0$ for all j, which means that the MPS Eq. (11) with tensors Eqs. (13) and (14) is annihilated by the SSEP generator.

The construction for the equilibrium state of the stochastic Fredkin chain at $c = 1/2$ is similar, but the relevant paths are

RW excursions. In this case the Doob transition probabilities that guarantee an excursion are (cf., e.g., Ref. [72])

$$T_j^{\rm ex}(h \to h \pm 1) = \begin{cases} \frac{1}{2}\left(1 + \frac{1}{h+1}\right)\left(1 - \frac{h}{N+1-j}\right) \\ \frac{1}{2}\left(1 - \frac{1}{h+1}\right)\left(1 + \frac{h+2}{N+1-j}\right) \end{cases} \tag{15}$$

for $0 \leqslant h \leqslant N+1-j$, or zero otherwise. The corresponding matrices now have bond dimension $N+1$ and read

$$B_0^{(j)} = \sum_{h=0}^{N} |h)(h-1|\, T_j^{\rm ex}(h \to h-1), \tag{16}$$

$$B_1^{(j)} = \sum_{h=0}^{N} |h)(h+1|\, T_j^{\rm ex}(h \to h+1), \tag{17}$$

with the same boundary vectors $(i| = (0|$ and $|f) = |0)$. The relevant relations in this case are $B_1^{(j-1)}B_0^{(j)}B_1^{(j+1)} - B_1^{(j-1)}B_1^{(j)}B_0^{(j+1)} = 0$ and $B_0^{(j-1)}B_1^{(j)}B_0^{(j+1)} - B_1^{(j-1)}B_0^{(j)}B_0^{(j+1)} = 0$ for all j. Given these, one can show that the MPS Eq. (11) with tensors Eqs. (16) and (17) is annihilated by the Fredkin generator Eq. (1). In fact, the MPS is annihilated by every local term in the spatial sum that defines Eq. (1), so $\mathbb{W}$ can be said to be a *parent generator* (cf. parent Hamiltonian [71]) of the MPS Eq. (11).

Note that from the definition of the tensors $B_n^{(j)}$ above in terms of transition probabilities, the MPS is in right canonical form, and Eq. (11) therefore satisfies $\langle -|{\rm ss}\rangle = 1$. Away from $c = 1/2$, we can also write Eq. (7) as an MPS if we reweigh the coefficients in Eqs. (16) and (17) as

$$T_j^{\rm ex}(h \to h \pm 1) \to \left(\frac{c}{1-c}\right)^{-\frac{1}{2}(h\pm 1)} T_j^{\rm ex}(h \to h \pm 1).$$

These reweighed coefficients are not transition probabilities in the height (they do not add up to one), meaning that the resulting MPS is not in canonical form. Finding the normalization $\mathcal{N}_c$ in this case is nontrivial.

B. Equilibrium phase diagram from numerical MPS

To overcome the difficulty above, to study the equilibrium properties for all c we resort to numerical MPS approximations. This we implement with the ITensor library [73] and make use of the DMRG [64,74,75] to find the leading eigenvector of Eq. (10). We employ an adaptive bond dimension, which is at most $D = 2000$ with a truncation cutoff error $\epsilon = 10^{-12}$. Furthermore, we exploit the U(1) symmetry which conserves the number of particles and initialize the MPS with a product state which lies in $\mathcal{D}$. We then carefully check the relevant observables to ensure they satisfy the properties associated with $\mathcal{D}$, such as a positive height field.

By looking at various observables at stationarity, it becomes clear that there are three distinct equilibrium phases in the Fredkin model: (i) $c < 0.5$, (ii) $c = 0.5$, and (iii) $c > 0.5$. We denote the expectation value of an observable O with respect to the equilibrium state as $\langle O\rangle$, with

$$\langle O\rangle = \langle -|O|{\rm ss}\rangle = \langle \psi|O|\psi\rangle\,. \tag{18}$$

The appropriate order parameter to characterize the equilibrium phases is the average area $\langle A\rangle$. In Fig. 2(a), we show $\langle A\rangle$ as a function of c for a range of system sizes $N \in [20, 400]$. For $c < 1/2$, the area becomes minimal, while for $c > 1/2$, the area is maximal. If we consider the area as a function of system size N we find that $\langle A\rangle$ grows as a power law $\langle N\rangle \sim N^{-\beta}$, as shown in Fig. 2(b). This reveals three distinct behaviors: the exponent β takes the values $\beta = 1, 3/2$, and 2 for $c < 1/2$, $c = 1/2$ and $c > 1/2$, respectively [9].

For each phase, we show the average of the spatial occupation profile, $\langle n_i\rangle$, and the average height field, $\langle h_i\rangle$, in Figs. 2(c) and 2(d), respectively. For $c < 1/2$, the particles take an antiferromagnetic arrangement, Fig. 2(c) (red circles), thus minimizing the height and therefore the area, Fig. 2(d) (red circles). We sometimes refer to this as the *flat phase* (in analogy with interacting dimers [76,77]).

At $c = 1/2$, all configurations occur with equal probability, cf. Eq. (7). In terms of the RW representation of configurations, this corresponds to the set of RW excursions. The average occupation, Fig. 2(c) (blue squares) interpolates between 1 and 0, and in the thermodynamic limit, $N \to \infty$, the average occupation density in the bulk is 1/2 [11]. In turn, the average height field takes a semicircular form, Fig. 2(d) (blue squares). Note that this is a phase of large fluctuations and this average height field is not representative of typical sample profiles. This is in contrast to the other two phases which are exponentially dominated by extremal area configurations, cf. Eq. (7).

For $c > 1/2$, the particles (holes) localize to the left (right) edge of the system [10], with a sharp change in average occupation, Fig. 2(c) (green triangles), and with an average height profile in the shape of a tent (with a rounded top, a finite residue of the fluctuations of the $c = 1/2$ phase), Fig. 2(d) (green triangles). This behavior is similar to that seen in the ASEP in an open segment with fixed boundaries [68]. Simple arguments (see Appendix A) give the profile [10]

$$\langle n_j\rangle = \frac{1}{\exp((j - N/2)/\lambda) + 1}, \tag{19}$$

with an inverse localization length λ,

$$\lambda = \ln\left(\frac{c}{1-c}\right)^{-1}. \tag{20}$$

We sometimes refer to the $c > 1/2$ phase as the *tilted phase* (also in analogy with interacting dimers [76,77]). Note that this shares no connection with the tilted generator later introduced in Sec. V.

An observable which will be of importance later is the dynamical activity $\langle k\rangle$, which measures the average number of configuration changes per unit time in stochastic trajectories [59,60,78]. At equilibrium, it can be measured as the average escape rate, $\langle k\rangle = \langle -|\mathbb{R}|{\rm ss}\rangle$, where $\mathbb{R}$ is the diagonal part of Eq. (1). We show this in Fig. 2(e) as a function of c for various system sizes $N \in [20, 400]$. It is immediately clear that the dynamical activity scales with system size (up to small finite size effects) for $c \leqslant 1/2$ where occupation is spread

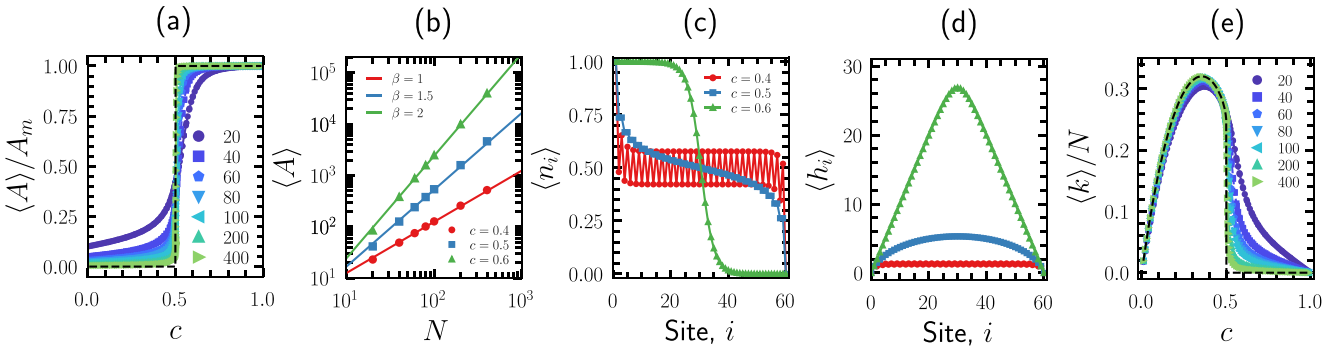


FIG. 2. Equilibrium properties of the Fredkin model. (a) The average area (scaled by maximum area) $\langle A\rangle/A_m$ as a function of c for various systems sizes $N \in [20, 400]$. The dashed line shows the extrapolated value for $N \to \infty$. (b) The average area (symbols) for $c = 0.4$ (red/dark grey), $c = 0.5$ (blue/medium grey) and $c = 0.6$ (green/light grey). The lines show the power laws $\langle N\rangle \sim N^{-\beta}$ with $\beta = 1,\ 1.5,\ 2$, respectively. (c), (d) The spin occupation $\langle n\rangle_i$ and height profiles $\langle h_i\rangle$ for each equilibrium phase with a system size $N = 60$. (e) The average dynamical activity (per unit time and system size) as a function of c for various systems sizes $N \in [20, 400]$. The dashed line shows the extrapolated value for $N \to \infty$. All results are calculated using numerical DMRG.

out in equilibrium, cf. Fig. 2(c), leading to less constrained and therefore more dynamics throughout the entire system. Conversely, the activity for $c > 1/2$ is subextensive in system size as expected due to the much more inactive conditions given to the localization of the equilibrium state, cf. Fig. 2(c): motion is limited to the center of the lattice (the tip of the tent), where particle hops are not restricted by exclusion. By fitting the activity with a linear form $\langle k\rangle = a + bN$ (for $c \leqslant 1/2$), one can extrapolate to infinite size to determine $\langle k\rangle/N$ in the thermodynamic limit. We show this as the black dashed line, peaking around $c \approx 0.36$. Notice that for $c > 1/2$, the activity goes as $O(1)$ and is suppressed by the scaling in N. The differences in the active ($c \leqslant 1/2$) and inactive ($c > 1/2$) dynamics are directly related to the dynamical LDs in Sec. IV.

C. Localization of the tilted phase

The equilibrium state for $c > 1/2$ is exponentially dominated by maximal area configurations, that is, configurations in which particles cluster toward the left edge of the system, and holes cluster at the right edge. Figure 3(a) shows the average occupation profile $\langle n_i\rangle$ for various $c > 1/2$: for sites beyond the halfway point, $i > N/2$, we observe an exponential decay of the average occupation, $\langle n_i\rangle \sim e^{-i/\lambda}$. [Note that the same occurs for the density of holes, $1 - \langle n_{N+1-i}\rangle$, coming from the right, due to fact the generator Eq. (1) is invariant under $i \to N + 1 - i$ and $|0\rangle \leftrightarrow |1\rangle$.]

This localization can be further characterized by the density of domain walls (DWs):

$$\left\langle n_i^{\mathrm{DW}}\right\rangle = \langle n_i(1 - n_{i+1})\rangle + \langle (1 - n_i)n_{i+1}\rangle\,. \tag{21}$$

This is shown in Fig. 3(b): The DW density is close to 1 at the center of the lattice, and decays exponentially when moving away from it in both directions, $\langle n_i^{\mathrm{DW}}\rangle \sim e^{-|\frac{N}{2}-i|/\lambda}$. Notice that the localization is consistent for increasing system size. As we discuss further in the next sections, exponential localization of DWs at the center of the lattice has important implications for the dynamics in the tilted phase: particle hops can only occur when there are DWs, and thus activity is exponentially suppressed away from the midpoint, and is subextensive in system size, cf. Fig. 2(e).

The localization length λ decreases with increasing c. We show this in Fig. 3(c) for both particle and DW densities. The agreement with the theoretical prediction Eq. (20) is excellent. The numerically extracted lengths here are from DMRG with $N = 100$. For smaller system sizes, the localization length becomes comparable to system size for $c \approx 1/2$, and one might expect to see small deviations from the theoretical prediction.

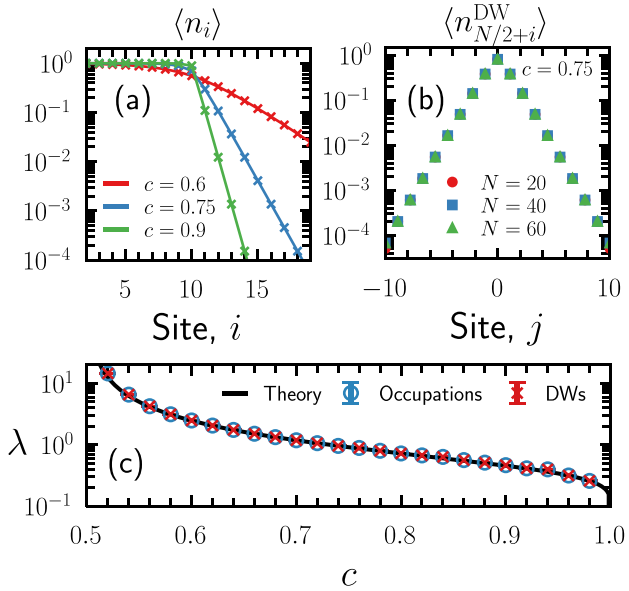


FIG. 3. Localization in the Fredkin chain. (a) The occupation profile $\langle n_i\rangle$ of the steady state for $c > 0.5$ and $N = 20$. The occupations exhibit an exponential decay for $i > N/2$. (b) The average domain-wall occupations $\langle n^{\mathrm{DW}}_{N/2-i}\rangle$ for $c = 0.75$ and $N = 20, 40, 60$. We see the same exponential decay of domain-wall density as we move away from the center of the lattice. (c) The localization length λ as a function of c. The line shows the result from the theory, Eq. (20), and the blue circles and red crosses the numerically extracted lengths from the occupation and DW profiles, respectively. The numerical data is from DMRG with $N = 100$.

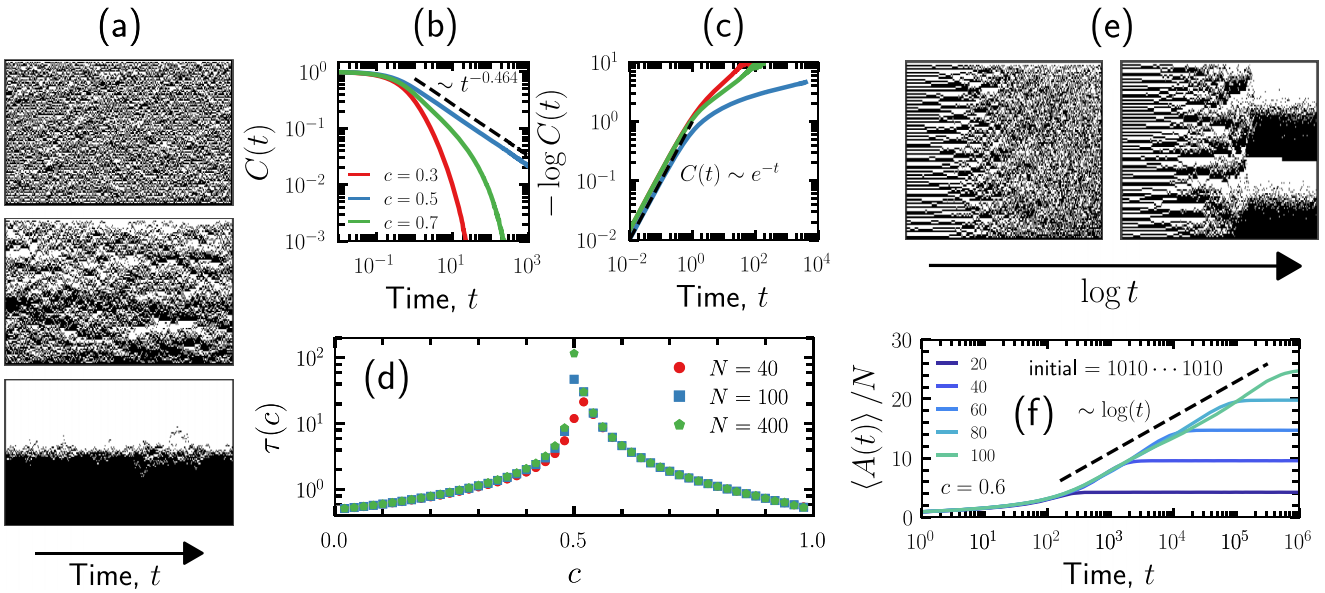


FIG. 4. Stochastic trajectories and dynamics. (a) Representative trajectories with initial states sampled from equilibrium for $c = 0.4$ (top), $c = 0.5$ (middle), and, $c = 0.6$ (bottom), respectively, for system size $N = 100$ and time $t = 10^3$. (b) The autocorrelation functions Eq. (22) for each of the three distinct equilibrium phases. At large times, the autocorrelator for $c = 0.5$ decays as the power law $t^{-0.464}$ (from size $N = 100$). (c) The same autocorrelation functions plotted on a double-log ordinate scale. For small times, they show exponential decay in the three phases. For large times, they take a stretched-exponential form for $c < 1/2$ and $c > 1/2$. (d) The numerically estimated timescales Eq. (23) as a function of c (for $N = 40$, 100, and 400). (e) Example trajectories after a quench from the initial state $1010\cdots1010$ for $c = 0.5$ (left) and $c = 0.6$ (on a logarithmic time scale). The former relaxes to equilibrium quickly, while the latter shows hierarchical relaxation (both panels for $N = 100$ and $t = 10^5$). (f) The area (scaled by system size) $\langle A\rangle /N$ after the same quench, for various system sizes $N \in [20, 100]$ and $c = 0.6$. The dashed line shows $\log t$. All data is obtained using continuous-time Monte Carlo.

IV. TYPICAL DYNAMICS

A. Dynamics in equilibrium

Figure 4(a) shows representative trajectories in the stationary dynamics of each of the three equilibrium phases (with the initial states sampled from equilibrium). The largest fluctuations occur for $c = 0.5$. Dynamics in equilibrium can be quantified through the (density) autocorrelation function,

$$C(t) = \frac{1}{N}\sum_{i=1}^{N} \frac{\langle n_i(0)n_i(t)\rangle - \langle n_i\rangle^2}{\langle n_i\rangle - \langle n_i\rangle^2}, \tag{22}$$

which provides a measure of the memory of a initial configuration after time t in an equilibrium trajectory. We show $C(t)$ for the three equilibrium phases in Figs. 4(b) and 4(c). It is apparent from Fig. 4(b) that for $c = 1/2$ the autocorrelation decays asymptotically as a power law, with a numerically extracted exponent of just under a half. This power-law decay can also be observed for intermediate times at $c > 1/2$ [corresponding to fluctuations of the top of the tent, cf. Fig. 2(d)], with this intermediate regime becoming longer as c gets closer to $1/2$. While at short times, decay is exponential, see Fig. 4(c), for long times relaxation is stretched exponential in both the flat and tilted phases. These are signatures of slow dynamics.

We can extract a timescale for relaxation of correlations from $C(t)$ from its integral:

$$\tau_{\rm eq} = \int_0^\infty C(t)dt. \tag{23}$$

This is shown in Fig. 4(d) for a range of c. This equilibrium timescale spikes at $c = 1/2$, as expected from the slow law decay of $C(t)$. Notice that the spike is less sharp for smaller system sizes due to the finite-size effects from the boundaries.

B. Relaxation toward equilibrium

Also of significance is the relaxation toward the equilibrium state when starting from nonequilibrium conditions. We explore this behavior by considering dynamics following from an initial state of minimal area, $x_0 = 1010\cdots1010$, corresponding to a quench from deep in the flat phase ($c_0 \gtrsim 0$) to finite $c > 0$. When $c < 1/2$, equilibrium is achieved quickly as the initial state is not far from typical states in the flat phase. Interesting nonequilibrium dynamics occurs when quenching to $c = 1/2$ or to the tilted phase at $c > 1/2$. In Fig. 4(e), we show two relaxation trajectories, one for $c = 1/2$ (left) and another for $c = 0.6$ (right) [79]. The system size is $N = 100$ and the overall time of trajectories $t = 10^5$ (note that the time axis is shown on a logarithmic scale). For the case of a quench to $c = 1/2$, after a slow early regime, equilibrium is reached in reasonable times.

For a quench to $c > 1/2$, we observe a slow hierarchical relaxation, with a progressive coarsening of clusters of particles and holes. The target state is a tilted one, cf. Fig. 2(d), and in the height representation this hierarchical process is the merging of smaller tents in the profile into larger ones. Due to the constraint, Eq. (2), local configurations of $\cdots0011\cdots$, corresponding to troughs in the height field, are locally trapped, and require particles from the right edge of clusters to diffuse

away to allow clusters to merge. This process is exponentially scarce in the separation distance, as occupations are exponentially localized, cf. Sec. III C.

The time to complete each stage of relaxation in the tilted phases grows exponentially with the stage. This hierarchy is evident in the growth of the average area normalized by system size, $\langle A(t)\rangle/N$, as shown in Fig. 4(f), where we see the area increasing logarithmically in time. This reveals the hierarchical nature of the relaxation process: while smaller systems may have relaxed to equilibrium, larger systems require the merging of larger clusters, and so the growth of the area continues.

V. DYNAMICAL LARGE DEVIATIONS

We now study the statistical properties of the stochastic trajectories $\omega_t = x_{0:t}$ of the Fredkin model, in particular, the LD statistics of dynamical observables.

If $K(\omega_t)$ is a trajectory observable, the probability of it taking a value K is

$$P_t(K) = \sum_{\omega_t} \pi(\omega_t)\delta[K(\omega_t) - K], \tag{24}$$

where $\pi(\omega_t)$ is the probability of the trajectory ω_t being realized under the stochastic dynamics. For a dynamical observable K that is time extensive, in the long-time limit, the probability of K obeys a LD principle [50–53],

$$P_t(K) \asymp e^{-t\varphi(K/t)}, \tag{25}$$

where the function $\varphi(k)$ is called the LD rate function. The above asymptotic equality holds as long as the spectral gap is nonvanishing (which it is in the Fredkin model for finite size N [2]). A LD principle also holds for the moment generating function,

$$Z_t(s) = \sum_K P_t(K)\, e^{-sK} = \sum_{\omega_t} \pi(\omega_t)\, e^{-sK(\omega_t)} \asymp e^{t\theta(s)},$$

where $\theta(s)$ is the scaled cumulant generating function (SCGF) whose derivatives at $s = 0$ give the cumulants of K, scaled by time [50]. In analogy with what occurs in equilibrium thermodynamics, the rate function and SCGF are related by a Legendre transform:

$$\theta(s) = -\min_k[sk + \varphi(k)]. \tag{26}$$

We consider as observable K the dynamical activity. Its SCGF is given by largest eigenvalue of the *tilted generator*, $\mathbb{W}_s$, which for the Fredkin model reads

$$\begin{aligned}\mathbb{W}_s = \sum_{i=2}^{N-2} f_i\{&c[e^{-s}\sigma_i^+\sigma_{i+1}^- - (1-n_i)n_{i+1}] \\ &+ (1-c)[e^{-s}\sigma_i^-\sigma_{i+1}^+ - n_i(1-n_{i+1})]\},\end{aligned} \tag{27}$$

with s being *counting field*. As $\mathbb{W}_s$ is in general non-Hermitian, the leading eigenvalue $\theta(s)$ has right and left eigenvectors $|r_s\rangle$ and $\langle l_s|$.

We can write the generator in a Hermitian form with the same similarity transformation as before, Eq. (8),

$$\begin{aligned}\mathbb{H}_s = -\sum_{i=2}^{N-2} f_i[&e^{-s}\sqrt{c(1-c)}(\sigma_i^+\sigma_{i+1}^- + \sigma_i^-\sigma_{i+1}^+) \\ &- c(1-n_i)n_{i+1} - (1-c)n_i(1-n_{i+1})],\end{aligned} \tag{28}$$

with ground state $\mathbb{H}_s|\psi_s\rangle = -\theta(s)|\psi_s\rangle$, related to the leading eigenvectors of $\mathbb{W}_s$ by

$$|\psi_s\rangle = \sum_x \sqrt{l_s(x)r_s(x)}\,|x\rangle\,, \tag{29}$$

where $l_s(x) = \langle l_s|x\rangle$ and $r_s(x) = \langle x|r_s\rangle$.

A. Active-inactive trajectory transitions at $c \leqslant 1/2$

From the ground state of Eq. (28), we can study statistical properties of the trajectory ensemble of the Fredkin model for long-time trajectories. We do this by means of numerical tensor networks along the lines of similar recent work in KCMs [80–85]. Figure 5 shows the LD statistics obtained numerically. The top row gives this for the flat phase at $c = 0.4$, and the middle row for the $c = 1/2$ phase. These results are for system sizes in the range $N \in [20, 400]$ obtained using DMRG.

Figure 5(a) shows the SCGF as a function of $s = 0$ for a range of sizes. For small $s \gtrsim 0$, the SCGF follows linear response (LR), $\theta(s) \approx -sk(0)$, where $k(s) = -\theta'(s)$ is the average dynamical activity in the ensemble tilted by s, shown in column (b). The LR prediction is shown by the dashed black line for $N \to \infty$, calculated by fitting the dynamical activity for finite sizes at $s = 0$ with a power law and extrapolating. Notice that at some $s_c(N) > 0$, which becomes smaller for increasing N, the behavior deviates from LR to one which no longer scales with N (this is most apparent for $c < 1/2$). The step in the average activity, Fig. 5(b), top and center, indicates a phase transition between dynamical phases of high and low activity. The change in activity tends to a discontinuity with increasing size, indicative of a first-order transition.

The point $s_c(N)$ at which the crossover occurs at finite size can be estimated from the peak in the corresponding dynamical susceptibility, $\chi(s) = \theta''(s)$, shown in column (c) of Fig. 5. As the system size is increased, the crossover point shifts towards $s = 0$ and becomes sharper. The change in dynamics can be seen in the broadening of the LD rate function $\varphi(k)$ around the equilibrium average, shown in Fig. 5(d). The rate functions show the characteristic Maxwell construct of a first-order transition between two phases, an active one with large k and an inactive one with vanishing k. Note that while the transition in activity looks less sharp for $c = 0.5$, we expect to recover the usual first-order behavior for increasing system sizes as seen by the broadening of the rate function.

For $c \leqslant 1/2$, the location of the crossover can be fit by a power law $s_c(N) \sim N^{-\alpha}$. The upper panel of Fig. 5(e) shows this for $c = 0.4$ and $c = 1/2$. The lower panel of Fig. 5(e) shows the dynamical exponents α as a function of c. When c is far from $1/2$, we have approximately $\alpha \approx 1.2$. The exponent increases quickly as we approach $c = 1/2$, to around $\alpha \approx 2.5$, a value similar to that found in other exclusion processes [83]. It could be that for values close to (but not equal to) $c = 1/2$,

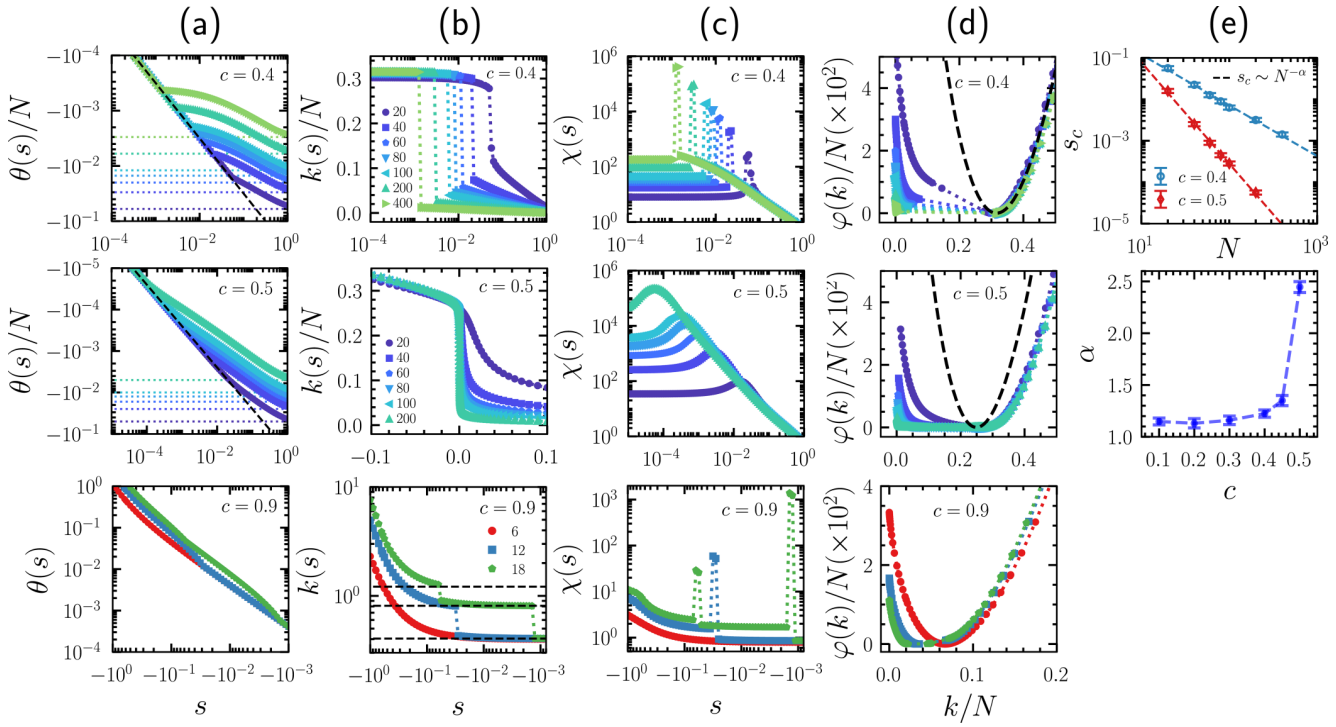


FIG. 5. Finite-size scaling of dynamical LD transitions. The dynamical LD statistics for each equilibrium phase. The top (middle) row of (a)–(d) shows $c = 0.4$ ($c = 0.5$) with $N \in [20, 400]$ obtained via DMRG and the bottom row shows $c = 0.9$ obtained through ED. (a) The SCGF $\theta(s)$ as a function of s. The upper and middle panels are scaled by system size, with dotted lines showing the value for $s \to \infty$ and the dashed line showing the linear response prediction (see the main text). (b) The average dynamic activity $k(s)$ as a function of s. The top and bottom panels are shown on log-log scales, whereas the middle one is shown in linear scale. The dashed lines in the bottom panel correspond to integer multiples of $k(0)$. (c) The dynamical susceptibility $\chi(s) = \theta''(s)$ as a function of s. (d) The LD rate function scaled by system size $\varphi(k)/N$ as a function of activity k. The black dashed lines show a Poisson distribution with mean $k(0)/N$ in the thermodynamic limit $N \to \infty$, extrapolated from finite-size DMRG data. (e) We estimate the critical point as a function of system size from the peaks of the dynamical susceptibility for $c = 0.4, 0.5$ in the top panel. The dashed lines shows a fitted power law $s_c \sim N^{-\alpha}$, with the bottom panel showing the obtained α for various c.

the measured exponent would be lower if we accounted for larger system sizes.

B. Dynamical phases for $c > 1/2$

Obtaining accurate estimates of $\theta(s)$ for $c > 1/2$ at large system sizes is difficult due to a proliferation of dynamical phases. In particular, it is hard for DMRG to converge to the correct phase due to a large density of states. For this reason, for $c > 1/2$ we limit our studies to system sizes $N = 6, 12, 18$ with large $c = 0.9 \gg 1/2$, which allows us to effectively study the hierarchy of dynamical phases using exact diagonalization (ED) [86]. The bottom row of Fig. 5 shows these results.

Since the typical dynamics ($s = 0$) of the tilted $c > 1/2$ phase is itself inactive, cf. Fig. 2(e), we expect transitions to the active phase to occur at $s < 0$ for finite size systems. In fact, Fig. 5(b) shows several points where the behavior of the SCGF changes. The number of these points seems to increase with system size. In each case, this change in behavior corresponds to transitions in the dynamics. At each of these points, we see a sharp drop in the activity, this becoming sharper with increasing N. The values at which the activity plateaus are multiple integers of the equilibrium activity, $k(s = 0)$, and are shown by the black dashed lines. With the limited range of sizes accessible via ED, it is not possible to do a finite-size scaling analysis as we did for $c < 1/2$. From the systems studied, we observe that the first away from equilibrium inactive behavior happens at increasing s (that is, getting closer to 0) for increasing N, which shows in the flattening of the rate function, see bottom panel in Fig. 5(d).

C. Structural properties of the dynamical large deviations

The difference in the behavior of the various dynamical phases also manifests in the structural properties of the configurations visited by the trajectories. The eigenvector $|\psi_s\rangle$ obtained from either DMRG and ED contains the probability amplitudes for each configuration, making it easy to calculate averages of configuration observables $\mathcal{O}(x)$ in the tilted ensemble [87]:

$$\langle \mathcal{O} \rangle_s = \langle L_s | \, \mathcal{O} \, | R_s \rangle = \sum_x \mathcal{O}(x) \psi_s(x)^2. \tag{30}$$

In Fig. 6, we show the local occupations $\langle n_i \rangle_s$ (top panels), and the average area $\langle A \rangle_s$ (middle and bottom panels), for (a) $c = 0.2$, (b) $c = 1/2$, and (c) $c = 0.9$. It is clear that the limit of large activity ($s < 0$ with $|s|$ large) particles spread out

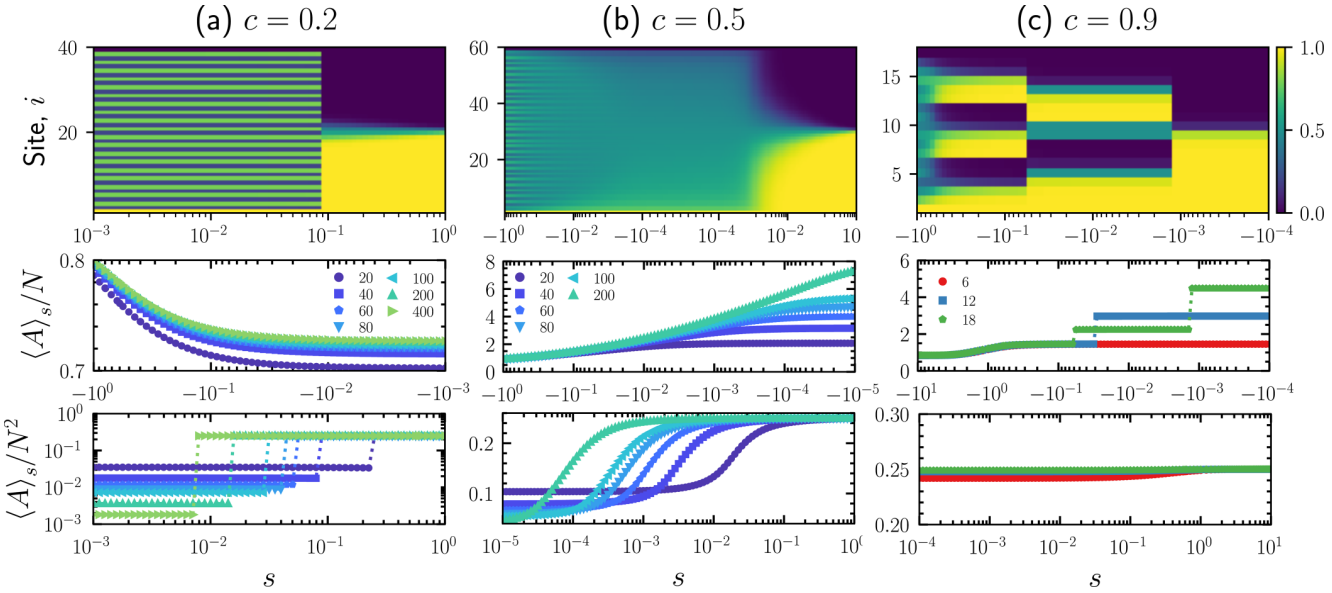


FIG. 6. Structural properties of the LDs. We show observables for each equilibrium phase with (a) $c = 0.2$, (b) $c = 0.5$, and (c) $c = 0.9$. The top row shows the average occupations $\langle n_i \rangle_s$ for site i (with differing system sizes and ranges of s). The middle row shows the area scaled by system size $\langle A \rangle_s / N$ for $s < 0$. Finally, we show the area scaled by system size squared $\langle A \rangle_s / N^2$ for $s > 0$ in the bottom row.

to maximize the activity. This is evident by the average area $\langle A \rangle_s$, which scales linearly with system size N, resembling the structures associated with the equilibrium flat phase for $c < 1/2$. Thus, the active phase for all values of c is also a structurally flat one. Conversely, in the inactive limit for all values of c (large $s > 0$), particles cluster at the left edge of the system and maximize the area, which scales as N^2. Thus, irrespective of the equilibrium static phase, the inactive phase of the dynamics is structurally tilted.

Interestingly, we observe very sharp transitions for $c \neq 1/2$ even at smaller sizes—this is unusual when compared to other constrained models [80,83]. This could be due to the sharp transition in activity at equilibrium, cf. Fig. 2. Indeed, for $c > 1/2$ we notice sharp structural changes at the location of the sharp points of Fig. 5. It is clear that the corresponding structures are related to the assembly of excited states at equilibrium ($s = 0$) obtained by joining multiple ground states of smaller system sizes (compared to what occurs in the excited states of the quantum East model [19]). Of course this makes sense, as despite the scarcity of the configurations associated with these states, they have large lifetimes (as discussed in Sec. III) with impactful consequences on the relaxation behavior.

D. Entanglement entropy

We now consider the bipartite von Neumann entanglement entropy of the MPS approximations to Eq. (28). We partition the system into two subsystems A and B, which denote the spins $i \in [1, N/2]$ and $i \in [N/2 + 1, N]$ respectively. The bipartite entanglement entropy between the two partitions is then calculated by

$$S_E(s) = -\mathrm{Tr}[\rho_\mathrm{A} \log \rho_\mathrm{A}], \tag{31}$$

where $\rho_A = \mathrm{Tr}_B[\rho]$ denotes the reduced density matrix for A, and $\rho = |\psi_s\rangle \langle \psi_s|$ is the density matrix for the full system.

The Hamiltonian Eq. (28) exhibits a ground-state phase transition in the bipartite entanglement entropy for $s = 0$. In particular, it scales as $S_E(0) \sim \log N$ for $c = 1/2$ and $S_E(0) \sim 1$ for $c \neq 1/2$ [1–4]. We now extend this analysis to $s \neq 0$. Figure 7 shows the entanglement entropy for increasing system size $N \in [20, 200]$ for $c = 0.4$ (top) and $c = 0.5$ (bottom) and a range of s. Notice that for $c < 1/2$, the entropy obeys an area law for all s (although we observe spikes around the transition from active to inactive dynamics). For $c = 1/2$,

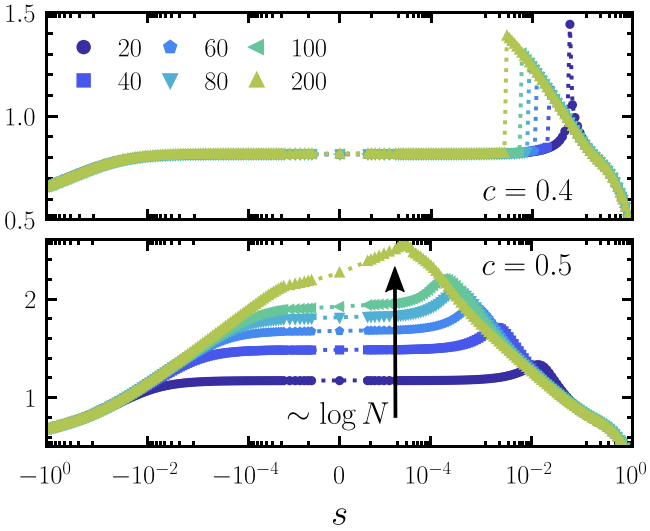


FIG. 7. Entanglement entropy of the LD eigenvectors. The bipartite entanglement entropy $S_E(s)$ for $c = 0.4$ (top) and $c = 0.5$ (bottom) and system sizes $N \in [20, 200]$. For $c = 0.5$, the entanglement entropy scales as approximately $S_E \sim \log N$.

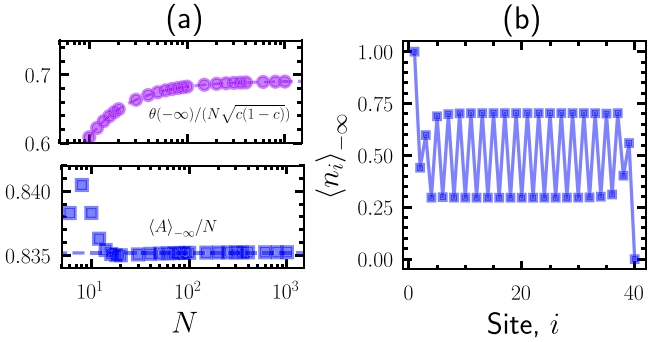


FIG. 8. Extreme active limit. (a) The rescaled SCGF $\tilde{\theta}/N$ (top) and the area $\langle A\rangle_{-\infty}/N$ (bottom) as functions of N measured via DMRG. We fit the SCGF as $a + bN^{-1}$, allowing us to extrapolate the value in the thermodynamic limit $N \to \infty$ (see main text), shown by the dashed line. The value for the area quickly settles with increasing system size, indicated by the dashed line. (b) The occupation profiles $\langle n_i\rangle_{-\infty}$ for $N = 40$.

we observe for large magnitude $s < 0$ the states clearly also obey an area law. As s approaches $s = 0$, the entanglement entropy appears to grow significantly toward $S_E(0)$, and looks to scale logarithmically. It is important to note, however, we only show a small range of system sizes, and it is most likely that for some fixed $s < 0$, the entanglement entropy will be bounded as $N \to \infty$, and thus obeys an area law. This can be seen by the branching behavior seen in Fig. 7. An important consequence is that for large enough N, one is able to construct a state with arbitrarily high entropy by tuning the value of s toward $s = 0$. For the inactive phase $s > 0$, ψ_s clearly also obeys an area law, again with the entanglement entropy spiking as s approach $s = 0$.

E. Limits of maximal and minimal activity

The limit of maximum activity is that at $s \to -\infty$. In this limit, the diagonal parts of $\mathbb{W}_s$ (and $\mathbb{H}_s$) are suppressed and only the off-diagonals are left. Notice that for $\mathbb{H}_s$, the dependence on c falls out as a prefactor. As the tilting in $\mathbb{W}_s$ grows exponentially with $-s$ for negative s, we rescale the SCGF as

$$\tilde{\theta} = \lim_{s\to-\infty} \frac{e^s\theta(s)}{\sqrt{c(1-c)}}, \tag{32}$$

when taking the limit. The (rescaled) eigenvalue $\tilde{\theta}$ coincides with the (similarly rescaled) dynamical activity. We show this in Fig. 8(a) as a function of $N \in [10, 400]$ (circles, shown divided by N), and fit it with the function of $aN + b$ (blue dashed line, shown divided by N). By extrapolating to infinity, we find that

$$\lim_{N\to\infty} \tilde{\theta}/N \approx 0.691. \tag{33}$$

The average area $\langle A\rangle_{-\infty}$, see Fig. 8(a) takes an almost constant value, with small fluctuations for small system sizes:

$$\lim_{N\to\infty} \langle A\rangle_s /N \approx 0.835. \tag{34}$$

Notice that the area scales linearly with system size and is similar to the equilibrium states found for $c < 1/2$. This is further seen from the average occupations $\langle n_i\rangle_{-\infty}$, see Fig. 8(b), showing the antiferromagnetic pattern of the flat equilibrium phase.

The opposite limit of $s \to \infty$ gives the most inactive state. In this limit, only the diagonal escape rate part of Eq. (27) [or Eq. (28)] remains and each configuration $x \in \mathcal{D}$ is an eigenstate. The configurations with the smallest escape rates dominate. Depending on c and N, this is either the maximal area (i.e., fully tilted) configuration, $1111\cdots0000$, which has escape rate $R = 2(1 - c)$, or the minimal area configuration $1010\cdots1010$, which has escape rate $R = c(N - 2)$. The latter dominates if $N > 2c^{-1}$, and the former dominates if $N < 2c^{-1}$ (with degeneracy at $N = 2c^{-1}$).

VI. CONCLUSIONS

Here we have provided a detailed study of the statics and dynamics of the stochastic Fredkin model. Despite being one-dimensional and having local transition rules, this model displays phase transitions between three distinct equilibrium phases. This is a consequence of the constraints in the dynamics which restrict the state space to that of RW excursions, with these static transitions controlled by the asymmetry in the particle hopping rates. Two of these phases are ordered, one being flat and another one tilted (in terms of the height field representation), with an intermediate disordered and fluctuating phase. This phase behavior is in some ways reminiscent of interacting two-dimensional dimer coverings [76,77,88].

The constraints in the local transitions of the Fredkin model lead to a rich dynamics, both in equilibrium and in the relaxation after a quench. This richness can be seen as a consequence of a nontrivial phase structure of the ensemble of stochastic trajectories. Using numerical MPSs with DMRG, we compute the LDs of the dynamical activity and show the existence of active-inactive space-time phase transitions, something that is also observed in other KCMs. The overall picture is one where the static phases extend into dynamical ones, with the flat phase also being a dynamical active phase, and the tilted phase a dynamical inactive one, with first-order transitions between them.

There are many possible continuations of the work here. One is to go beyond one dimension. As an initial step, in Appendix B we propose a two-dimensional generalization of the Fredkin model: By focusing on the fact that Fredkin configurations are RW excursions, we proposed a two-dimensional model in terms of packed dimers on the honeycomb lattice with constraints in the dynamics which enforce configurations to be excursion surfaces. It will be interesting to study this and similar stochastic models in future work. Another interesting area of exploration would be to study the Hamiltonian Eq. (28) under unitary dynamics, in analogy with recent work that studied other quantum KCMs. As occurs with the quantum East model [19], we expect the constraints in Fredkin models to provide mechanisms for localization and nonthermal eigenstates. We hope to report on this in the near future.

ACKNOWLEDGMENTS

We acknowledge financial support from EPSRC Grants No. EP/R04421X/1 and No. EP/P034616/1 and the

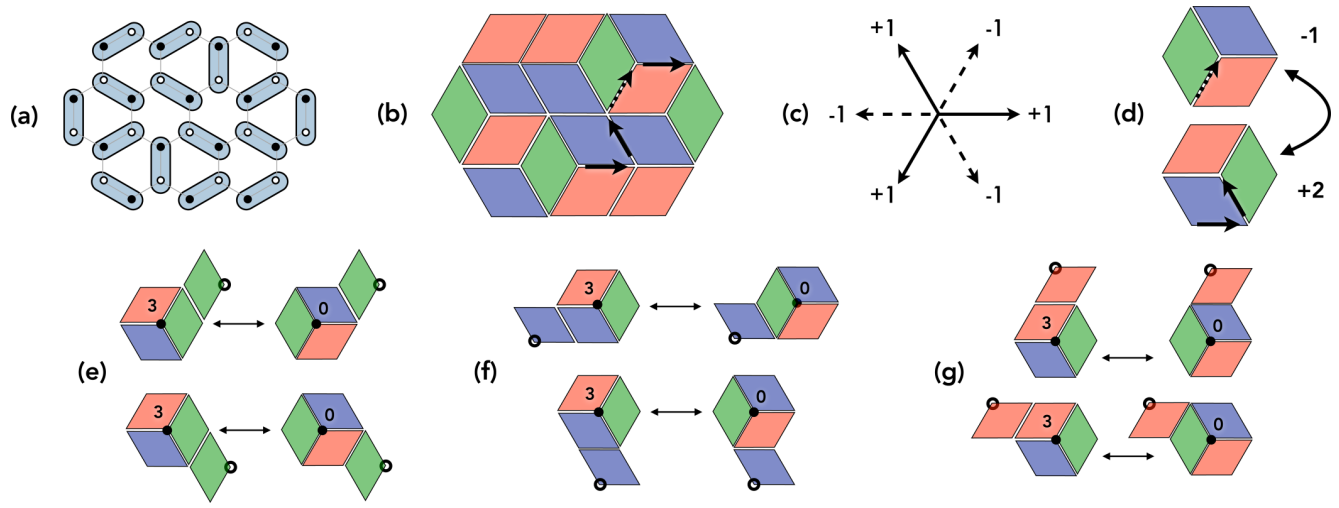


FIG. 9. Two-dimensional generalization of the Fredkin model. (a) Dimer covering of the honeycomb lattice. (b) Equivalent rhombus tiling of the plane. (c) Definition of the height field: Given a tiling, moving along the edges of the rhombi the height increases of decreases by one unit as shown. For example, the path in (b) shows that the start and end points have a height difference of +2. (d) The elementary local moves that preserve the perfect tiling character (i.e., no tiling defects or no monomers in the dimer representation) are rotations of a triplet of tiles forming an elementary hexagon. These are the dimer/tiler equivalent of the particle-hole exchange in the SEP. These moves change the height field of the central site by three units. (e)–(g) Constrained moves: Requiring the presence of the extra tile guarantees that the height of the central site (indicated by the filled circle) never goes below the lowest height of the arrangement (indicated by the open circle). These are the two-dimensional equivalents of the allowed moves in the Fredkin chain, see Fig. 1(a).

Leverhulme Trust Grant No. RPG-2018-181. We acknowledge access to the University of Nottingham Augusta HPC service.

APPENDIX A: THE DENSITY PROFILE FOR $c > 1/2$

Simple statistical mechanical considerations can be used to compute the density profile for $c > 1/2$ in the thermodynamic limit. Recall that the probability of a configuration x is weighted by a factor that depends on the area $A(x)$ under the path

$$P(x) \propto \left(\frac{c}{1-c}\right)^{\frac{1}{2}A(x)} = \exp(-\beta A(x)),$$

where $\beta \equiv \frac{1}{2}\log[(1-c)/c]$. The entropy associated with a configuration is just the sum of the binary entropies:

$$S(x) = -\sum_i [n_i \log n_i + (1-n_i)\log(1-n_i)].$$

After writing the area as

$$A(x) = \sum_{i=j}^{N} h_j(x) = \sum_{i=j}^{N} (N+1-j)\, Z_j \qquad \text{(A1)}$$

($Z_i = 2n_i - 1$), we arrive at the free energy,

$$\sum_{j=1}^{N} [(\xi - j)Z_j] - \beta^{-1} S,$$

where ξ is a Lagrange multiplier introduced to fix the overall particle number $\sum_j n_j$. Extremizing the free energy gives

$$Z_j = \tanh(\beta[\xi - j]).$$

Thus Z_j has a DW profile with a location ξ that is determined by the particle number ($\xi = N/2$ for half filling), and a width

$$\lambda \equiv (2\beta)^{-1} = \left[\ln\left(\frac{c}{1-c}\right)\right]^{-1}.$$

APPENDIX B: POSSIBLE TWO-DIMENSIONAL GENERALIZATION

The height representation of the Fredkin model suggests several possible generalizations to two dimensions by analogy with dimer coverings. One possibility is the following.

Consider a fully packed dimer covering of the honeycomb lattice, see Fig. 9(a), where each link connecting any two neighboring sites of the lattice is occupied by a dimer. Such coverings have a height representation in terms of a height field $h_{i,j}$, which becomes apparent in the equivalent rhombus tiling of the plane, see Fig. 9(b): From some origin (0,0) where $h_{0,0} = 0$, the height of a site is computed by moving along the edges of the rhombi with Δh at each step according to Fig. 9(c). For example, in the covering of Fig. 9(b), the two initial and final sites connected by the path with the arrows differ in height by $\Delta h = 2$. For fully packed dimer configurations (also called perfect tilings), it is easy to verify that any path that connects two sites gives the same height difference and the height field is uniquely defined. Honeycomb dimer coverings (rhombus tilings) therefore define surfaces in two dimensions. In a configuration with an equal amount of the three kind of tiles, the height field is pinned at zero at the boundaries (for example, in three sites at angles of $2\pi/3$ within a hexagonal region). This is a two-dimensional version of the one-dimensional height field from a lattice of particles

and holes at half filling which is bound to return to the origin.

In the one-dimensional case, the elementary local move that preserves the filling fraction is to exchange a particle with an adjacent hole. The analogous move for a rhombus tiling is shown in Fig. 9(d) and corresponds to rotating a triplet of tiles forming an elementary hexagon. This move changes the height of the central site by $\Delta h = \pm 3$. To prevent the height field from becoming negative, which is the defining property of the dynamics of the Fredkin model, transitions like those of Fig. 9(d) have to be constrained, cf. Fig. 1(a). Figures 9(e)–9(g) show the corresponding allowed transitions in the two-dimensional case: the exchange of tiles is only possible if either of the extra green/blue/red tiles as in arrangement Figs. 9(e)–9(g), respectively, is present, and not allowed otherwise. This constraint implies that in the transition the height of the site at the center of the hexagon cannot go below that of the site indicated by a circle. With this dynamical rule, it is guaranteed that the height field of the dimer/rhombus arrangement never becomes negative at any point, a two-dimensional version of the RW excursions that define the configurations of the Fredkin model. Furthermore, giving different rates to the forward and backward moves in Figs. 9(e)–9(g) should lead to flat and tilted phases weighted by the volume under the surface.

[1] O. Salberger and V. Korepin, arXiv:1605.03842
[2] R. Movassagh, Ann. Math. Sci. Appl. **3**, 531 (2018).
[3] F. Sugino and V. Korepin, Int. J. Mod. Phys. B **32**, 1850306 (2018).
[4] L. Dell'Anna, SciPost Phys. **7**, 053 (2019).
[5] L. Dell'Anna, O. Salberger, L. Barbiero, A. Trombettoni, and V. E. Korepin, Phys. Rev. B **94**, 155140 (2016).
[6] X. Chen, E. Fradkin, and W. Witczak-Krempa, Phys. Rev. B **96**, 180402(R) (2017).
[7] L. Dell'Anna, L. Barbiero, and A. Trombettoni, J. Stat. Mech. Theory Exp. (2019) 124025.
[8] K. Adhikari and K. S. D. Beach, Phys. Rev. B **104**, 115149 (2021).
[9] O. Salberger, T. Udagawa, Z. Zhang, H. Katsura, I. Klich, and V. Korepin, J. Stat. Mech. Theory Exp. (2017) 063103.
[10] T. Udagawa and H. Katsura, J. Phys. A: Math. Theor. **50**, 405002 (2017).
[11] K. Adhikari and K. S. D. Beach, Phys. Rev. B **99**, 054436 (2019).
[12] C. M. Langlett and S. Xu, Phys. Rev. B **103**, L220304 (2021).
[13] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[14] J. P. Garrahan, P. Sollich, and C. Toninelli, in *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media*, International Series of Monographs on Physics, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos (Oxford University Press, Oxford, UK, 2011), Chap. 10, pp. 341–366.
[15] J. P. Garrahan and D. Chandler, Phys. Rev. Lett. **89**, 035704 (2002).
[16] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[17] M. van Horssen, E. Levi, and J. P. Garrahan, Phys. Rev. B **92**, 100305(R) (2015).
[18] A. Prem, J. Haah, and R. Nandkishore, Phys. Rev. B **95**, 155133 (2017).
[19] N. Pancotti, G. Giudice, J. I. Cirac, J. P. Garrahan, and M. C. Bañuls, Phys. Rev. X **10**, 021051 (2020).
[20] B. Pozsgay, T. Gombor, A. Hutsalyuk, Y. Jiang, L. Pristyák, and E. Vernier, Phys. Rev. E **104**, 044106 (2021).
[21] B. Olmos, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. Lett. **109**, 020403 (2012).
[22] B. Olmos, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. E **90**, 042147 (2014).
[23] D. C. Rose, K. Macieszczak, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. E **105**, 044121 (2022).
[24] P. Fendley, K. Sengupta, and S. Sachdev, Phys. Rev. B **69**, 075106 (2004).
[25] I. Lesanovsky, Phys. Rev. Lett. **106**, 025301 (2011).
[26] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Phys. Rev. B **98**, 155134 (2018).
[27] T. Prosen and C. Mejía-Monasterio, J. Phys. A: Math. Theor. **49**, 185003 (2016).
[28] A. Inoue and S. Takesue, J. Phys. A: Math. Theor. **51**, 425001 (2018).
[29] T. Prosen and B. Buča, J. Phys. A: Math. Theor. **50**, 395002 (2017).
[30] B. Buča, J. P. Garrahan, T. Prosen, and M. Vanicat, Phys. Rev. E **100**, 020103(R) (2019).
[31] A. J. Friedman, S. Gopalakrishnan, and R. Vasseur, Phys. Rev. Lett. **123**, 170603 (2019).
[32] S. Gopalakrishnan, Phys. Rev. B **98**, 060302(R) (2018).
[33] S. Gopalakrishnan, D. A. Huse, V. Khemani, and R. Vasseur, Phys. Rev. B **98**, 220303(R) (2018).
[34] K. Klobas, M. Medenjak, T. Prosen, and M. Vanicat, Commun. Math. Phys. **371**, 651 (2019).
[35] K. Klobas and T. Prosen, SciPost Phys. Core **2**, 010 (2020).
[36] V. Alba, J. Dubail, and M. Medenjak, Phys. Rev. Lett. **122**, 250603 (2019).
[37] V. Alba, Phys. Rev. B **104**, 094410 (2021).
[38] T. Iadecola and S. Vijay, Phys. Rev. B **102**, 180302(R) (2020).
[39] J. W. P. Wilkinson, K. Klobas, T. Prosen, and J. P. Garrahan, Phys. Rev. E **102**, 062107 (2020).
[40] T. Gombor and B. Pozsgay, Phys. Rev. E **104**, 054123 (2021).
[41] J. W. P. Wilkinson, T. Prosen, and J. P. Garrahan, Phys. Rev. E **105**, 034124 (2022).
[42] K. Klobas, B. Bertini, and L. Piroli, Phys. Rev. Lett. **126**, 160602 (2021).
[43] B. Pozsgay, J. Phys. A: Math. Theor. **54**, 384001 (2021).
[44] B. Buča, K. Klobas, and T. Prosen, J. Stat. Mech.: Theory Exp. (2021) 074001.
[45] H. Singh, B. A. Ware, R. Vasseur, and A. J. Friedman, Phys. Rev. Lett. **127**, 230602 (2021).
[46] J. Richter and A. Pal, Phys. Rev. Res. **4**, L012003 (2022).
[47] T. Brewer, Efficient simulation of rare events in one-dimensional systems using a parallelised cloning algorithm, Ph.D. thesis, University of Bath, 2019.
[48] R. A. Blythe and M. R. Evans, J. Phys. A: Math. Theor. **40**, R333 (2007).
[49] K. Mallick, Physica A **418**, 17 (2015).
[50] H. Touchette, Phys. Rep. **478**, 1 (2009).

[51] J. P. Garrahan, Physica A **504**, 130 (2018).
[52] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).
[53] D. T. Limmer, C. Y. Gao, and A. R. Poggioli, Eur. Phys. J. B **94**, 145 (2021).
[54] T. Bodineau and B. Derrida, C. R. Acad. Sci. **8**, 540 (2007).
[55] C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van Wijland, Phys. Rev. E **78**, 021122 (2008).
[56] T. Bodineau, V. Lecomte, and C. Toninelli, J. Stat. Phys. **147**, 1 (2012).
[57] D. Karevski and G. M. Schütz, Phys. Rev. Lett. **118**, 030601 (2017).
[58] M. Merolle, J. Garrahan, and D. Chandler, Proc. Natl. Acad. Sci. USA **102**, 10837 (2005).
[59] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[60] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
[61] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A: Math. Theor. **42**, 075007 (2009).
[62] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).
[63] S. N. Majumdar and A. Comtet, J. Stat. Phys. **119**, 777 (2005).
[64] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[65] Note that the Fredkin model is different from the so-called facilitated SEP [89], where jumps are only allowed if there is a particle behind the one that is hopping. For example, in that model the third transition in Fig. 1(a) is not allowed.
[66] This is analogous to what occurs in the ASEP with the same kind of boundaries, see, e.g., Ref. [68].
[67] T. Agranov, P. Zilber, N. R. Smith, T. Admon, Y. Roichman, and B. Meerson, Phys. Rev. Res. **2**, 013174 (2020).
[68] J. De Gier and F. H. Essler, J. Stat. Mech.: Theory Exp. (2006) P12011.
[69] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).
[70] U. Schollwöck, Ann. Phys. **326**, 96 (2011).
[71] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Rev. Mod. Phys. **93**, 045003 (2021).
[72] R. Chetrite and H. Touchette, Ann. Henri Poincaré **16**, 2005 (2015).
[73] M. Fishman, S. R. White, and E. M. Stoudenmire, arXiv:2007.14822.
[74] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[75] I. P. McCulloch, J. Stat. Mech. (2007) P10014.
[76] S. Papanikolaou, E. Luijten, and E. Fradkin, Phys. Rev. B **76**, 134514 (2007).
[77] C. Castelnovo, C. Chamon, C. Mudry, and P. Pujol, Ann. Phys. **322**, 903 (2007).
[78] C. Maes, Phys. Rep. **850**, 1 (2020).
[79] The lack of visible fluctuations for $c = 0.5$ is a consequence of the logarithmic scale in time.
[80] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).
[81] P. Helms, U. Ray, and G. K.-L. Chan, Phys. Rev. E **100**, 022101 (2019).
[82] P. Helms and G. K.-L. Chan, Phys. Rev. Lett. **125**, 140601 (2020).
[83] L. Causer, I. Lesanovsky, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **102**, 052132 (2020).
[84] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **103**, 062144 (2021).
[85] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. Lett. **128**, 090605 (2022).
[86] We construct the basis so only the dynamically connected configurations within the sector are included.
[87] The meaning of this is the following: At a given value of s, what we are exploring is the ensemble of trajectories ω_t with probabilities reweighed from those of the original dynamics as $\pi_s(\omega_t) = Z_t^{-1}(s)\pi(\omega_t)e^{-sK(\omega_t)}$. For long-time trajectories, $t \gg 1$, the probability of finding a configuration $x_{t'}$ at an intermediate time $1 \ll t' \ll t$ (i.e., far from both initial and final time boundaries) is given by $l_s(x)r_s(x) = \psi_s(x)^2$. Expectation values such as $\langle A \rangle_s$ provide information about the properties of configurations at the midpoint of long-time atypical trajectories in the ensemble characterized by s.
[88] A. Stannard, J. C. Russell, M. O. Blunt, C. Salesiotis, M. d. C. Gimenez-Lopez, N. Taleb, M. Schroeder, N. R. Champness, J. P. Garrahan, and P. H. Beton, Nat. Chem. **4**, 112 (2012).
[89] A. Gabel, P. L. Krapivsky, and S. Redner, Phys. Rev. Lett. **105**, 210603 (2010).

Chapter 6

Optimal sampling of dynamical large deviations via matrix product states

The following work is from the publication *"Optimal sampling of dynamical large deviations via matrix product states"* by Luke Causer, Mari Carmen Bañuls and Juan P. Garrahan, in Physical Review E 103, 062144 (2021).

Previous efforts, including the previous two chapters, make use of MPSs to characterize the dynamical LDs for KCMs. The MPS contains information about the long-time averaged properties associated with the dynamical LDs. However, they do not contain any of the dynamical information of the rare events (other than the average number of jumps). To retrieve this information, one must consider the *trajectories* associated with the rare events. Generating these trajectories on-hand is a much harder task. This problem is tackled here by using MPS to estimate the so-called "Doob" dynamics, the true dynamics responsible for generating the rare trajectories.

The constructed dynamics only approximates the Doob dynamics, but can be used in tandem with trajectory sampling algorithms, such as TPS, to efficiently sample rare-events directly from the Doob dynamics. Our work shows this can be easily achieved with little computational power. The method can be used to sample from trajectory ensembles at *finite times*, in contrast to the purely MPS approaches, which only provide time-averaged information at infinite times. This approach allowed for an investigation of the temporal scaling of first-order dynamical phase transitions.

Corrections to the manuscript:

1. Equation 7 should read $P_t(K) \sim e^{-t\varphi(K/t)}$.

Editors' Suggestion

Optimal sampling of dynamical large deviations via matrix product states

Luke Causer,[1,2] Mari Carmen Bañuls,[3,4] and Juan P. Garrahan[1,2,5]

[1]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

[2]*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

[3]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany*

[4]*Munich Center for Quantum Science and Technology (MCQST), Schellingstrasse 4, D-80799 München, Germany*

[5]*All Souls College, Oxford, United Kingdom*

(Received 1 April 2021; accepted 7 May 2021; published 28 June 2021)

The large deviation statistics of dynamical observables is encoded in the spectral properties of deformed Markov generators. Recent works have shown that tensor network methods are well suited to compute accurately the relevant leading eigenvalues and eigenvectors. However, the efficient generation of the corresponding rare trajectories is a harder task. Here, we show how to exploit the matrix product state approximation of the dominant eigenvector to implement an efficient sampling scheme which closely resembles the optimal (so-called "Doob") dynamics that realizes the rare events. We demonstrate our approach on three well-studied lattice models, the Fredrickson-Andersen and East kinetically constrained models, and the symmetric simple exclusion process. We discuss how to generalize our approach to higher dimensions.

DOI: 10.1103/PhysRevE.103.062144

I. INTRODUCTION

The complex behavior of the nonequilibrium dynamics of stochastic systems can be characterized by studying trajectory ensembles, that is, the set of all possible trajectories alongside the probability that they occur under the evolution defined via a stochastic master operator. This is analogous to standard thermodynamics, where static properties are entirely determined by the equilibrium ensemble of all microstates and the probabilities [1]. Often the dynamical behavior of interest is dominated not by trajectories that are typical under the dynamics, but by "rare events" which are exponentially (in time and in system size) scarce. Studying these rare events is made possible by using the framework of large deviations (LDs) [2–7], where in large time limits time-extensive dynamical observables obey a LD principle, and their statistics is encoded in functions which play for dynamics the role that thermodynamic potentials play for statics (see below for definitions).

LD functions can be obtained in principle from a deformation or *tilting* of the dynamical generator (in the case of continuous-time dynamics) or the Markov matrix (in the case of discrete-time dynamics), through its largest eigenvalue. Obtaining this eigenvalue is not always an easy—or even possible—task, and often one needs to resort to numerical methods. Methods to overcome this difficulty often include techniques based on population dynamics, namely cloning or splitting [8–11], and importance sampling [12–16] which provide information about the configurations frequently visited by the rare events. Notice that even if one manages to diagonalize the tilted generator (or the Markov matrix), the generation of rare trajectories is nontrivial: While rare trajectories are "generated" by the tilted operator, this is not a proper stochastic operator and these trajectories cannot be directly sampled.

The efficient sampling of rare events can be achieved by searching for another stochastic dynamics which generates trajectories with desirable probabilities that are the same as (or a close approximation to) those of the tilted generator (with any small discrepancy corrected via importance sampling techniques). Methods for doing so currently include optimal control [17,18] and machine learning approaches, where one attempts to "learn" this convenient sampling dynamics [19–21]. The optimal choice for a reference dynamics is the so-called *generalized Doob dynamics* [22–27], which generates trajectories with the exact tilting corresponding to the deformed generator. The Doob dynamics thus produces rare trajectories of the original dynamics "on demand." To construct such optimal dynamics, however, requires knowledge of the leading eigenvector of the tilted generator.

Variational tensor network (TN) techniques [28–34], originally devised as a tool to study quantum many-body systems, are also convenient for studying classical statistical systems [35–38]. More recently, they have been shown to be useful in the context of LDs in stochastic dynamics [39–43]. In particular, it is often both possible and easy to approximate the leading eigenstate of the tilted generator of a one-dimensional stochastic lattice system using a matrix product state (MPS) ansatz, even those with dynamical (i.e., LD) phase transitions. Recent works have made use of this eigenstate to determine the statistical properties of the dynamics [39–43]. However, such TN approach has not yet been exploited to sample efficiently rare trajectories. This is what we do in this paper. We

present a scheme to use the MPS approximation to the leading eigenvalue of the tilted generator to construct a dynamics which very closely resembles the optimal Doob dynamics, and we show how we can use this dynamics to efficiently sample rare events.

We focus on three paradigmatic models. The first two correspond to kinetically constrained models (KCMs) [6,44,45], specifically the Fredrickson-Andersen (FA) and the East [46] model, two well-studied models known for their connection to structural glasses [47,48]. The third model is the symmetric simple exclusion processes (SSEPs) [49,50]. All these models have interesting LD statistics, including trajectory phase transitions controlled by their activities and/or currents (in the case of the SSEP) [3,5,51–54].

The paper is organized as follows. In Sec. II, we review continuous-time Markov dynamics and LDs. We also recap how one can apply an MPS ansatz to study KCMs. In Sec. III, we define the Doob dynamics and introduce a scheme to approximate it with a reference dynamics, constructed using an MPS approximation to the leading eigenstate of the tilted generator. In Sec. IV we present the numerical results from our method applied to the three models. We show how our approach can effectively be used to accurately measure the statistics of time-extensive observables. We provide an outlook on possible generalizations and our conclusions in Sec. V.

II. LARGE DEVIATIONS AND MATRIX PRODUCT STATES

In this section we introduce continuous-time Markov dynamics, giving specific examples in the context of kinetically constrained models (KCMs) and exclusion processes. We then also review the framework of large deviations (LDs) and how variational matrix product states (MPS) can be used to determine the LD statistics.

A. Continuous-time Markov dynamics for KCMs and exclusion processes

We consider stochastic Markov dynamics which evolves continuously in time. Suppose we have some system with the set of configurations $\{x_1, x_2, \ldots, x_M\}$ where M is the size of the configuration space. The probability that the system is in some configuration x at the time t is encoded in the probability vector $|P(t)\rangle = \sum_x P(x,t)\,|x\rangle$ which evolves under the stochastic master equation

$$\frac{d}{dt}\,|P(t)\rangle = \mathbb{W}\,|P(t)\rangle\,. \tag{1}$$

Here, the generator of the dynamics $\mathbb{W}$ is given by

$$\mathbb{W} = \sum_{x,x'\neq x} w_{x\to x'}\,|x'\rangle\langle x| - \sum_x R_x\,|x\rangle\langle x|\,, \tag{2}$$

where $w_{x\to x'}$ are the transition rates from configuration x to x' and $R_x = \sum_{x'\neq x} w_{x\to x'}$ is the escape rate from x. The largest eigenvalue of the generator is zero, with the left eigenvector the flat state $\langle -| = \sum_x \langle x|$, and the right eigenvector the steady state $|\mathrm{ss}\rangle = \sum_x P(x)\,|x\rangle$, which describes the probability of finding any configuration at equilibrium. If our system obeys detailed balance, then we are guaranteed that any initial state will eventually relax to some equilibrium state given enough time. Here, we assume this to be the case.

We will focus on two broad areas of one-dimensional (1D) constrained systems. The first is KCMs (for reviews, see Refs. [6,44,45]), for which configuration changes are governed by a kinetic constraint which is explicitly encoded in the generator. For concreteness, we focus on the 1D spin facilitation Fredrickson-Andersen (FA) [55] and East [46] models. Both models are defined on a 1D lattice of N binary variables (spins) $n_j = 0, 1$ for $j = 1, \ldots, N$, and configuration changes are only allowed via single-spin flips. The Markovian generators for both models are given by

$$\mathbb{W}^{\mathrm{East/FA}} = \sum_{i=1}^{N} \mathbb{P}_i^{\mathrm{East/FA}}[c\sigma_i^+ + (1-c)\sigma_i^- - c(1-n_i) - (1-c)n_i] \tag{3}$$

where $\sigma_i^{\pm}$ are the Pauli raising/lowering operators acting on site i and $c \in (0, 0.5]$ controls the rates at which spins flip, given they satisfy the kinetic constraints

$$\mathbb{P}_i^{\mathrm{FA}} = n_{i-1} + n_{i+1}, \quad \mathbb{P}_i^{\mathrm{East}} = n_{i-1}, \tag{4}$$

where the first only allows a transition if the spin attempting to flip has a neighboring excitation, and the second only if the neighboring spin to the left is excited. (For the FA model the constraint is sometimes defined as the projector $n_{i-1} + n_{i+1} - n_{i-1}n_{i+1}$, but in practice it makes little difference with the definition above.)

The second area we consider are exclusion processes [49,50]—particles hopping around sites on a lattice, with a hardcore exclusion such that we can have at most one particle per site. We focus on the 1D symmetric simple exclusion process (SSEP), adopting the lattice notation we used for KCMs, where now $n_j = 1(0)$ implies the site is occupied (empty). In the SSEP, a particle can hop left or right to its neighboring sites, both with the same rate ($\gamma = 1/2$) if the neighboring site is not already occupied. The generator for the dynamics is

$$\mathbb{W}^{\mathrm{SSEP}} = \frac{1}{4}\sum_{i=1}^{N}\left(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y + \sigma_i^z\sigma_{i+1}^z - 1\right), \tag{5}$$

where σ_i^a are the Pauli operators acting on site i.

For the entirety of this paper, we will assume open boundary conditions (OBCs), which will later reduce the computational cost of tensor network contractions. This formally means that we set $n_0 = n_{N+1} = 0$. Furthermore, we impose certain restrictions on the state space. For the FA model, we simply exclude the disconnected zero state $n_i = 0$, $\forall i$. On the other hand, we set $n_1 = 1$ for the East model which ensures the state space remains fully connected on each dynamical site $i > 1$. Finally, we restrict SSEP such that the total number of particles $N_p = \sum_i n_i$ is fixed, with particle density $n_p = N_p/N$ which will be assumed to be $n_p = 1/2$.

B. Trajectories and large deviations

Consider some general trajectory $\omega_t = \{x_0 \to x_{t_1} \to \cdots \to x_{t_K}\}$ where the system moves into the configuration x_{t_i} at time t_i and has the total time $t > t_K$. The dynamical activity $\hat{K}$ [3–6,56] is a trajectory observable which measures

the number of configuration changes for a given trajectory. The probability of observing some activity K can then be calculated as the sum over all trajectories with K configuration changes, and the probability they occur,

$$P_t(K) = \sum_{\omega_t} \pi(\omega_t)\delta[\hat{K}(\omega_t) - K], \tag{6}$$

where $\pi(\omega_t)$ is the probability of observing ω_t. For large times, this obeys the large deviation (LD) principle [2–5]

$$P_t(K) \sim e^{t\varphi(K/t)}, \tag{7}$$

where $\varphi(K/t)$ is called the LD rate function and plays the role of entropy density for trajectories. Alternatively, one can consider the moment generating function (MGF) [2]

$$Z_t(s) = \sum_K P_t(K)e^{-sK} = \sum_{\omega_t} \pi(\omega_t)e^{-s\hat{K}(\omega_t)}, \tag{8}$$

which contains equivalent information to Eq. (7) and can be considered the partition function. From Eq. (8), we see that the weighting of each trajectory is the probability that the trajectory occurs, exponentially reweighted by its dynamical activity. The MGF also obeys a LD principle,

$$Z_t(s) \sim e^{t\theta(s)}, \tag{9}$$

where $\theta(s)$ is the scaled cumulant generating function (SCGF), whose derivatives evaluated at $s = 0$ give the cumulants of K scaled by time. The SCGF plays the role of the thermodynamical free energy of trajectories and is related to the LD rate function by a Legendre transform $\theta(s) = -\min_k[sk + \varphi(k)]$ [2].

The MGF Eq. (8) can be expressed as

$$Z_t(s) = \langle -|e^{t\,\mathbb{W}_s}|\text{in}\rangle\,, \tag{10}$$

where $|\text{in}\rangle$ is some initial probability vector and $\mathbb{W}_s$ is an operator which we name the *tilted generator*, and is a deformed version of Eq. (2) where we tilt with respect to the dynamical observable of interest [2–5]. For the case of the dynamical activity [3–5], we simply tilt the off-diagonals of $\mathbb{W}$ with the same factor to obtain

$$\mathbb{W}_s = \sum_{x,x'\neq x} e^{-s} w_{x\to x'}\,|x'\rangle\,\langle x| - \sum_x R_x\,|x\rangle\,\langle x|\,. \tag{11}$$

The largest eigenvalue of $\mathbb{W}_s$ is the SCGF $\theta(s)$, with associated left and right eigenvectors $\langle l_s|$ and $|r_s\rangle$. Since $\langle l_s|$ in general is not the flat state, $\mathbb{W}_s$ is not a proper stochastic generator for $s \neq 0$ [3–5]. If one could exactly diagonalize Eq. (11) to find its leading eigenvalue and eigenvectors, then they would entirely unravel the LD statistics. We now briefly recap how this can be achieved using numerical TN techniques [40–43].

C. Variational matrix product states

A matrix product state (MPS) is an ansatz for describing vector states of many-body systems [28–30,57,58],

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N}^{d} \text{Tr}\big(A_1^{i_1}A_2^{i_2}\cdots A_N^{i_N}\big)\,|i_1\,i_2\cdots i_N\rangle\,, \tag{12}$$

where each subsystem k has its own rank-3 tensor A_k with the dimensions $d \times D \times D$. The allowed entanglement within the state is controlled by the *bond dimension* D [32]. It is often convenient to represent tensor networks in a diagrammatic form using shapes to represent tensorial objects, and (connecting) lines to represent contractions over tensors. For example, the corresponding diagram for an MPS is

$|\Psi\rangle =$  , (13)

where each circle corresponds to one of the tensors A_k. Similarly, one can also attempt to write some operator $\hat{O}$ as a matrix product operator (MPO) [33,34,59–62]. Operators which act locally on the subsystems, such as Eqs. (3)–(5), can be efficiently described as a MPO. That is to say, we can represent them exactly in MPO form with only a small constant bond dimension. The diagrammatic representation for MPOs is

$\hat{O} =$  . (14)

MPSs allow for the easy and efficient implementation of the widely used density matrix renormalization group (DMRG) method [63,64], an algorithm designed to iteratively minimize the energy of a state E_Ψ with respect to some Hamiltonian $\hat{H}$. In the language of MPS [30], we start with some guess at some fixed bond dimension, and sweep through each tensor applying local optimizations with all other tensors fixed. This is done until we reach convergence, which is usually when the change in energy of the state per sweep is small. At the end of the routine, one can efficiently calculate the variance of the state with respect to the Hamiltonian

$$\delta E_\Psi{}^2 = \text{var}_{\hat{H}}(\Psi) = \langle \hat{H}^2\rangle_\Psi - \langle \hat{H}\rangle_\Psi^2\,, \tag{15}$$

where $\langle\cdot\rangle_\Psi = \langle\Psi|\cdot|\Psi\rangle$ denotes an expectation value. We check to see if it has fallen below some desired value ϵ; if not, we run the algorithm with an increased bond dimension, where we typically use the state from the previous run as an initial guess. For more details on the workings of variational MPS (vMPS) algorithms, see the reviews [30,65].

Many recent works have shown that vMPS algorithms are very effective for studying the LD statistics of classically constrained systems which obey detailed balance [40,41,43]. In particular, if we write the tilted generator in a way such that it is Hermitian, then the state we are searching for is the ground state. This guarantees each update is an improvement upon the last. For dynamics obeying detailed balance, the activity-tilted generator can be brought to a Hermitian form using a similarity transformation that is independent of s [5],

$$\mathbb{H}_s = -Q^{-1}\mathbb{W}_s Q. \tag{16}$$

For the case of the East/FA models [5], the diagonal operator Q is given by

$$Q^{\text{FA/East}} = [\sqrt{1-c}\,|0\rangle\,\langle 0| + \sqrt{c}\,|1\rangle\,\langle 1|]^{\otimes N}, \tag{17}$$

and for the SSEP by $Q^{\text{SSEP}} = \mathbb{I}$. The Hamiltonian $\mathbb{H}_s$ has the ground state $|\psi_s\rangle$ with energy $-\theta(s)$. The ground state is

related to the left and right eigenvectors of $\mathbb{W}_s$ in the following way [40],

$$|\psi_s\rangle = Q^{-1}\,|r_s\rangle\,, \tag{18}$$

$$\langle\psi_s| = \langle l_s|\,Q, \tag{19}$$

$$|\psi_s\rangle = \sum_x \sqrt{l_s(x)r_s(x)}\,|x\rangle\,, \tag{20}$$

where $l_s(x) = \langle l_s|x\rangle$ and $r_s(x) = \langle x|r_s\rangle$.

III. DOOB TRANSFORMATION AND OPTIMAL SAMPLING

We now define the so-called generalized Doob transformation [19,22–25,27,66], and show how one can use our MPS solution to Eq. (16) to construct a reference dynamics which closely resembles the true Doob dynamics. We then present a method to optimally sample the rare events of our toy models using these dynamics.

A. Generalized Doob dynamics

The goal is to find a proper stochastic generator which generates trajectories with the same probabilities as those in the tilted dynamics $\mathbb{W}_s$, cf. Eq. (11). This can be achieved using the (long-time) generalized Doob transformation [19,22–25,27,66], defined as

$$\mathbb{W}_s^{\text{Doob}} = \mathbb{L}[\mathbb{W}_s - \theta(s)\mathbb{I}]\mathbb{L}^{-1}, \tag{21}$$

where $\mathbb{L} = \text{diag}(\langle l_s|)$ is the left eigenvector $\langle l_s|$ as a diagonal matrix. It is easy to check that Eq. (21) is annihilated by the flat state $\langle -|$, which means that $\mathbb{W}_s^{\text{Doob}}$ is a stochastic operator. Its stationary state is

$$|\text{ss}\rangle_s^{\text{Doob}} = \sum_x \tilde{P}(x)\,|x\rangle = \sum_x l_s(x)r_s(x)\,|x\rangle\,. \tag{22}$$

The generator $\mathbb{W}_s^{\text{Doob}}$ can also be expressed as a sum of its diagonal and off-diagonal elements

$$\mathbb{W}_s^{\text{Doob}} = \sum_{x,x'\neq x} \frac{l_s(x')}{l_s(x)} e^{-s} w_{x\to x'}\,|x'\rangle\,\langle x| - \sum_x [R_x + \theta(s)]\,|x\rangle\,\langle x|\,. \tag{23}$$

Thus our dynamics has the transition rates and escape rates

$$\tilde{w}_{x\to x'} = \frac{l_s(x')}{l_s(x)} e^{-s} w_{x\to x'}, \tag{24}$$

$$\tilde{R}_x = R_x + \theta(s), \tag{25}$$

respectively. That is to say, the transition rates are reweighted by e^{-s} and by some ratio $l_s(x')/l_s(x)$ which depends on the structure of the configurations, and the escape rate is shifted by $\theta(s)$.

We now consider some general time-dependent observable $\hat{A}$, and ask what is the expectation value in the tilted dynamics,

$$\langle\hat{A}\rangle_s \equiv \frac{\langle\hat{A}e^{-sK}\rangle}{\langle e^{-sK}\rangle} = Z_t(s)^{-1}\sum_{\omega_t}\pi(\omega_t)\hat{A}(\omega_t)e^{-s\hat{K}(\omega_t)}. \tag{26}$$

One can now apply importance sampling to arrive at

$$\langle\hat{A}\rangle_s = Z_t(s)^{-1}\sum_{\omega_t}\tilde{\pi}(\omega_t)\frac{\pi(\omega_t)}{\tilde{\pi}(\omega_t)}\hat{A}(\omega_t)e^{-s\hat{K}(\omega_t)} = Z_t(s)^{-1}\left\langle\frac{\pi}{\tilde{\pi}}\hat{A}e^{-s\hat{K}}\right\rangle_{\text{Doob}}, \tag{27}$$

where $\tilde{\pi}(\omega_t)$ is the probability of observing ω_t in the dynamics generated by $\mathbb{W}_s^{\text{Doob}}$ and $\langle\cdot\rangle_{\text{Doob}}$ denotes an expectation value with respect to trajectories with probabilities from the Doob dynamics. At a first glance, it might appear that we have not gained much from expressing the expectation of A using the Doob generator $\mathbb{W}_s^{\text{Doob}}$. However, if one calculates the ratio of probabilities in Eq. (27), then the power of this expression becomes apparent.

Let us first consider the original dynamics described by Eq. (2). If we have some system in configuration x, then the probability it flips to some other state x' at the time Δt is

$$P_{x\to x'}(\Delta t) = w_{x\to x'}e^{-R_x\Delta t}. \tag{28}$$

It then follows that the trajectory ω_t occurs with probability

$$\pi(\omega_t) = P(x_0)e^{-R_{x_K}(t-t_{x_K})}\prod_{i=1}^{K} w_{x_{i-1}\to x_i}\,e^{-R_{x_{i-1}}(t_{x_i}-t_{x_{i-1}})}, \tag{29}$$

where we have also accounted for the fact that the system must remain in the same state after the final flip for the remainder of the time and the probability of the initial configuration $P(x_0)$ (where we assume it is in the steady state). The probability of the trajectory under the Doob dynamics has a similar form, with the substitutions $w_{x\to x'} \to \tilde{w}_{x\to x'}$, $R_x \to \tilde{R}_x$, and $P(x_0) \to \tilde{P}(x_0)$,

$$\tilde{\pi}(\omega_t) = \tilde{P}(x_0)e^{-sK}e^{-t\theta(s)}\frac{l_s(x_K)}{l_s(x_0)}e^{-R_{x_k}(t-t_k)} \times\prod_{i=1}^{K} w_{x_{i-1}\to x_i}\,e^{-R_{x_{i-1}}(t_{x_i}-t_{x_{i-1}})}, \tag{30}$$

where all but the endpoint factors of $l_s(x)$ cancel out telescopically. The ratio of probabilities then goes as

$$\frac{\pi(\omega_t)}{\tilde{\pi}(\omega_t)} = \frac{e^{sK}e^{t\theta(s)}}{l_s(x_0)\,l_s(x_K)}, \tag{31}$$

where we have used $\tilde{P}(x_0) = P(x_0)\,l_s(x_0)^2$. Substituting Eq. (31) back into Eq. (27) cancels out the exponential tilting e^{sK}. Furthermore, for large times, $Z_t(s)^{-1} \approx e^{-t\theta(s)}$, giving the final result

$$\langle\hat{A}\rangle_s = \left\langle\frac{1}{l_s(x_0)l_s(x_K)}\hat{A}\right\rangle_{\text{Doob}}. \tag{32}$$

And so it follows that one can exactly sample the expectation value of a trajectory observable in the tilted ensemble defined by the nonstochastic tilted generator, by sampling it directly from trajectories generated by the stochastic Doob dynamics Eq. (21), up to factors at the endpoints of each trajectory (which become negligible in the long-time limit if $\hat{A}$ is time extensive). We note that Eq. (32) can also be derived by means

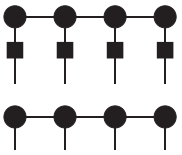
of linear algebra, using Eq. (21) and the ratio $P(x_0)/\tilde{P}(x_0)$ (see Ref. [67] for details).

B. Reference dynamics

While the above shows how to optimally sample if one has access to the Doob generator, which is obtained from the exact minimization of the tilted generator, we now consider how to approximate it efficiently.

Suppose we have an MPS approximation $|\psi_s^{\rm ref}\rangle$ to the ground state of the Hermitian operator $\mathbb{H}_s$, where our choice of bond dimension D controls the error. By applying the operator Q^{-1} to $|\psi_s^{\rm ref}\rangle$, as is done in Eq. (19), one can also retrieve an approximation to the left eigenvector. This is easily done as an MPS-MPO product,

$\langle l_s^{\rm ref}| = \langle \psi_s^{\rm ref}| Q^{-1} =$



$=$

. (33)

We then construct the generator of the so-called *reference dynamics*, which goes as Eq. (2) with the transition rates and escape rates given by

$$w^{\rm ref}_{x\to x'} = \frac{l_s^{\rm ref}(x')}{l_s^{\rm ref}(x)} e^{-s} w_{x\to x'}, \tag{34}$$

$$R^{\rm ref}_x = \sum_{x'\neq x} w^{\rm ref}_{x\to x'}, \tag{35}$$

respectively. Note that here we have not used Eq. (25) for the escape rates, as these reference dynamics only act as an approximation to the Doob dynamics, and thus would not give a true stochastic dynamics. In Appendix A, we show the steady-state solution to the reference dynamics is given by

$$|{\rm ss}\rangle_s^{\rm ref} = \sum_x \psi_s^{\rm ref}(x)^2 \, |x\rangle \, , \tag{36}$$

where $\psi_s^{\rm ref}(x) = \langle x|\psi_s^{\rm ref}\rangle$.

If we repeat the steps between Eqs. (27) and (32) but for the reference dynamics, the expectation Eq. (27) appears as

$$\langle \hat{A}\rangle_s = \left\langle \frac{1}{l_s^{\rm ref}(x_0) l_s^{\rm ref}(x_K)} e^{-t\theta(s)+\int dt \Delta\hat{R}} \hat{A} \right\rangle_{\rm ref}, \tag{37}$$

where $\int dt \Delta\hat{R}$ is the time integral of the difference of escape rates between the reference dynamics and the original dynamics, with $\Delta\hat{R}_x = R_x^{\rm ref} - R_x$. We can estimate a sampling error when using Eq. (37) in the following way [66]. First, let us assume the effects of the time-edge factors are negligible (as they are not exponential in time) and try to sample the quantity

$$\langle e^{-s\hat{K}}\rangle = \langle \mathcal{R} e^{-s\hat{K}}\rangle_{\rm ref} \approx \frac{1}{N_{\rm sp}} \sum_{\alpha=1}^{N_{\rm sp}} \mathcal{R}(\omega^\alpha) e^{-s\hat{K}(\omega^\alpha)}, \tag{38}$$

where $\mathcal{R}(\omega^\alpha) = e^{s\hat{K}(\omega^\alpha)+\int dt \Delta\hat{R}(\omega^\alpha)}$ is the umbrella which compensates for change in the sampling dynamics and we estimate for a fixed number of samples $N_{\rm sp}$. The variance of Eq. (38) gives a way to quantify the sampling error,

$$\begin{aligned}\epsilon^2_{\rm ref} &= \frac{{\rm Var}_{\rm ref}\left(\frac{1}{N_{\rm sp}}\sum_{\alpha=1}^{N_{\rm sp}} \mathcal{R}(\omega^\alpha) e^{-s\hat{K}(\omega^\alpha)}\right)}{\left\langle \frac{1}{N_{\rm sp}}\sum_{\alpha=1}^{N_{\rm sp}} \mathcal{R}(\omega^\alpha) e^{-s\hat{K}(\omega^\alpha)}\right\rangle^2_{\rm ref}} \\ &= \frac{1}{N_{\rm sp}}\left[\frac{\langle \mathcal{R}^2 e^{-2s\hat{K}}\rangle_{\rm ref}}{\langle \mathcal{R} e^{-s\hat{K}}\rangle^2_{\rm ref}} - 1\right] \\ &= \frac{1}{N_{\rm sp}}\left[\frac{\langle e^{2\int dt' \Delta\hat{R}}\rangle_{\rm ref}}{\langle e^{\int dt' \Delta\hat{R}}\rangle^2_{\rm ref}} - 1\right].\end{aligned} \tag{39}$$

In Appendix B we show

$$\epsilon^2_{\rm ref} \approx \frac{e^{t\delta E^2} - 1}{N_{\rm sp}} \approx \frac{t\delta E^2}{N_{\rm sp}}. \tag{40}$$

The last approximation holds for δE small enough ($t\delta E^2 \ll 1$). In Eq. (40), δE^2 is the calculated variance on our MPS approximation of the leading eigenvector [see Eq. (15)].

C. Simulating trajectories

We are now in a position to efficiently simulate trajectories from our reference dynamics. The sampling of trajectories from a classical generator is usually achieved using a continuous-time Monte Carlo [CTMC, otherwise known as the Bortz-Kalos-Lebowitz (BKL) algorithm] [68]. Given that our system is in some configuration x at time t', we need to calculate the next jump in the trajectory. That is, we need to decide the next configuration the system will move into, and the time it does so. Calculating this can be split into five separate steps:

(1) Find each configuration x' the system can move into from x.

(2) Calculate the transition rates $w_{x\to x'}$ for each x'.

(3) Calculate the escape rate R_x as the sum of all transition rates.

(4) Randomly choose one x', each with the probability $w_{x\to x'}/R_x$

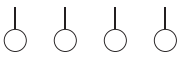
(5) Randomly choose the jump time Δt with probability $P(\Delta t) = R_x e^{-R_x \Delta t}$.

By starting at a configuration sampled from equilibrium (which in the case of the reference dynamics can be efficiently done using the MPS $|\psi_s^{\rm ref}\rangle$ [69,70]), or otherwise, one can simply repeat this procedure until some total time t has elapsed.

We can use this method for our reference dynamics, where the only step that needs slight adjustment is the second. While one must still calculate the transition rates of the original dynamics in the usual way, we must also calculate the left vector components $l_s^{\rm ref}(x)$ and $l_s^{\rm ref}(x')$. Let us assume the former is carried over from the previous jump in the algorithm. Then all one needs to do is calculate each $l_s^{\rm ref}(x')$. We start by noting that any configuration x can be written in MPS form with bond dimension 1,

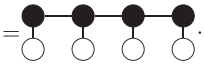
$|x\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \cdots \otimes |n_N\rangle$

$=$



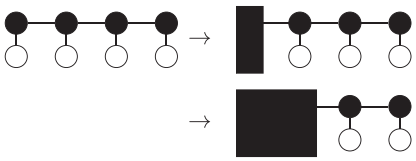
(41)

and then we can simply calculate the left component as a MPS-MPS contraction

$$l_s^{\mathrm{ref}}(x) = |\langle l_s^{\mathrm{ref}}|x\rangle| =$$  (42)

The transition rates for the reference dynamics are then calculated using Eq. (34), and the method proceeds as before. The total computational cost for calculating each $l_s^{\mathrm{ref}}(x)$ is $O(D^2N)$, and thus the total cost of each Monte Carlo (MC) step is $O(D^2NN_F)$, where N_F is the total number of configurations x' for a given step.

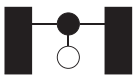
Let us now consider our KCMs where we have single-spin-flip dynamics. We first note that the number of possible configuration changes from x is bounded by the number of sites, that is, $1 \leqslant N_F \leqslant N$. Using the method described above, the computational cost for each step is, at worst, quadratic in the system size. However, by realizing that the tensor network contractions $\langle l_s^{\mathrm{ref}}|x\rangle$ and $\langle l_s^{\mathrm{ref}}|x'\rangle$ are identical apart from just one tensor (corresponding to the spin which would flip), we can reduce the computational cost by recycling partial contractions from the edges. We first need to identify the first and last sites on the lattice which are able to flip, which we label i_l and i_r, respectively. In a similar fashion to variational algorithms, we then contract from the left edge of the tensor network $\langle l_s^{\mathrm{ref}}|x\rangle$ up to $i_r - 1$, and saving each tensor block along the way,



We do the same but from the right and up to $i_l + 1$. This initialization of partial contractions has a one-time cost of

$$O(D^2(N + i_r - i_l - 2)) < O(2D^2N). \tag{43}$$

Calculating each $l_s^{\mathrm{ref}}(x')$ at site j is then easy. We just contract our remaining tensors at site j with the previously saved left and right blocks,

$$l_s^{\mathrm{ref}}(x') =$$ 

This is done for each possible site which can flip, and thus entails a computational cost $O(D^2N_F)$. Once a choice is made for which site to flip, which we will label i, we must update the blocks of partial contractions up to (the now possibly different) i_l and i_r. Note that this time we do not have to start from the edges of the MPS, but just from site i as the previous partial contractions that come before do not change. The total cost of updating the partial contractions is

$$O(D^2[(i_r - i) + (i - i_l)]) = O(D^2(i_r - i_l)). \tag{44}$$

The total computation cost for each MC step is the sum of the cost for calculating each $l_s^{\mathrm{ref}}(x)$ and updating the partial blocks after a choice is made,

$$O(D^2(N_F + i_r - i_l)) \leqslant O(2D^2N). \tag{45}$$

Consequentially, the cost of each MC step is reduced to one which is at most linear in system size.

IV. NUMERICAL RESULTS

A. Approximating the Doob dynamics

We put to the test the general method presented above by approximating the Doob dynamics of each model defined in Sec. II. We show that the Doob dynamics is well estimated using the MPS reference dynamics, and can even be well approximated with truncated MPS.

Each of the three models is known to exhibit a trajectory phase transition (when tilted against the activity) for long times and in the thermodynamic limit $N \to \infty$, manifested in the SCGF $\theta(s)$ at $s = 0$ with a discontinuous drop in the dynamical activity $\hat{K}(s) = -\theta'(s)/N$ [3,5,43,52,54]. We call the dynamical phase for $s < 0$ the *active phase*, and that for $s > 0$, the *inactive phase*. One is able to do a detailed investigation of this first-order phase transition by considering the finite-size scaling of the model [40,43,53,71,72]. We can estimate a critical point $s_c(N) \gtrsim 0$ by finding the peak of the dynamical susceptibility $\chi(s) = \theta''(s)$, which shows a drastic change in a small region around the transition point.

We start by taking the usual approach of approximating the ground states $|\psi_s\rangle$ using vMPS. That is, we run the algorithm allowing the bond dimension to increase until the variance of the energy (with respect to the Hamiltonian) falls sufficiently [cf. Eq. (15)]. The resultant MPS is then used to construct the reference dynamics, which approximates the Doob dynamics to a high accuracy, as explained in the previous section. Note that because the vMPS tries to keep entanglement as low as possible, for $s > s_c(N)$ the approximated ground state exhibits localization at just one edge of the system [40]. While for the East case this corresponds to the structure of the ground state in the sector with fixed occupation 1 in the leftmost site, the FA and SSEP models have reflection symmetry, spontaneously broken for $s > 0$ and large N. Thus, in order to maintain the symmetry in the latter two cases, we construct an MPS which is a superposition of the result from vMPS and its spatially reflected state to obtain our dynamics in the inactive phase.

1. Direct sampling with the reference dynamics but without reweighting

We first check that the CTMC algorithm with our MPS reference dynamics gives the expected results. We do this *without using* the trajectory reweighting [cf. Eq. (37)]. This amounts to only considering infinite-time dynamics, and assuming that our approximation is actually exact. Despite this strong assumption, we find that it produces excellent results as shown in Fig. 1. The expected dynamical activity (per unit site and time, dashed lines) can be calculated as a TN contraction over our MPS and MPO,

$$\tilde{k}(s) = \frac{1}{N}\langle\psi_s|\frac{d\mathbb{H}_s}{ds}|\psi_s\rangle. \tag{46}$$

The same quantity can be calculated on a trajectory level (symbols) by counting the total number of configuration changes, $\langle K\rangle$ and taking its time (and spatial) average,

$$\langle k\rangle_s = \frac{\langle K\rangle}{Nt}, \tag{47}$$

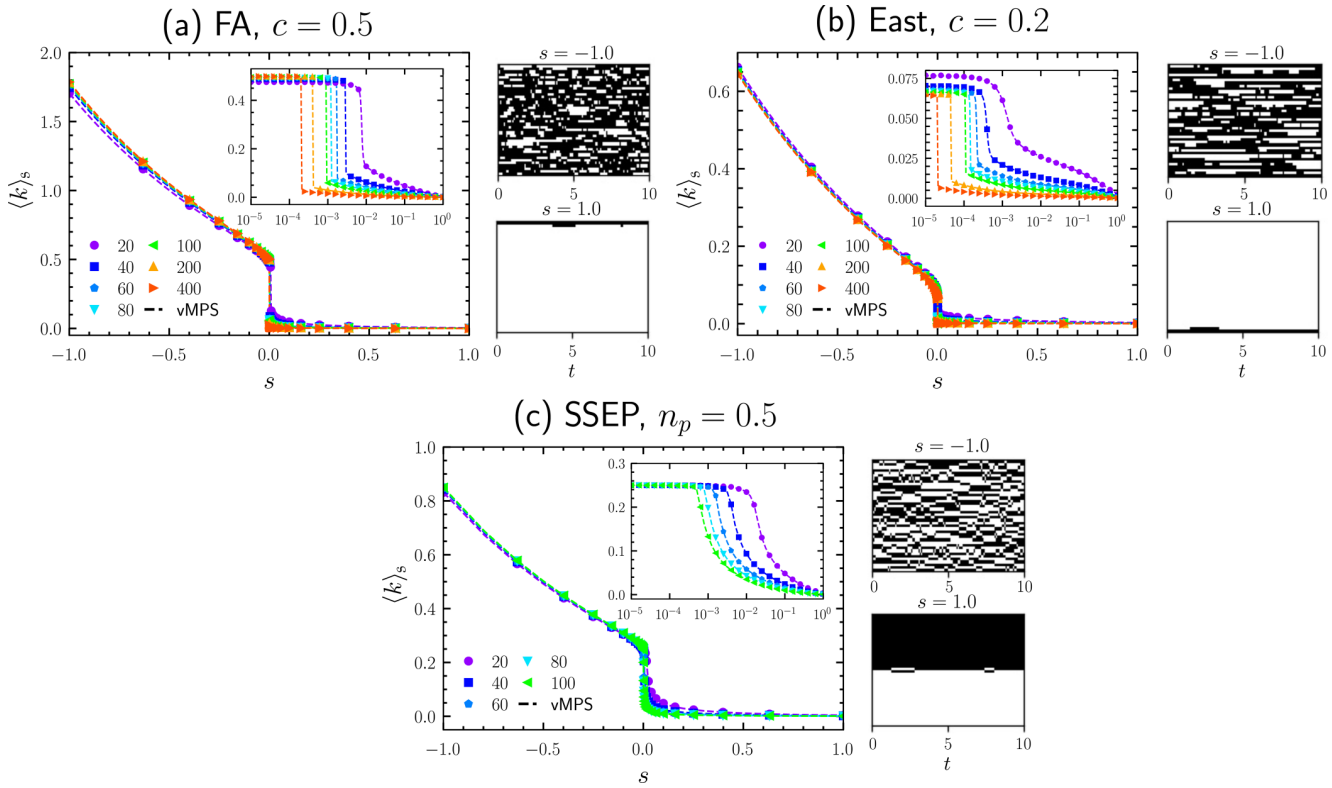


FIG. 1. The dynamical activity from brute-force Monte Carlo. We show the dynamical activity measured for (a) the FA model with $c = 0.5$, (b) the East model with $c = 0.2$, and (c) SSEP with $n_p = 0.5$. A variety of system sizes $N \in [20, 400]$ are shown for each. The dashed lines show expected activity calculated directly from the MPS $\tilde{k}(s)$, whereas the markers show the activity measured via CTMC $\langle k \rangle_s$ with a time $t = 100/\tilde{k}(s)$. The inactive phase is shown with a log s scale in the insets. We also show representative trajectories at $s = -1, 1$ for each model.

where t is the run time for each trajectory. We show results for each model, for a range of system sizes of $N \in [20, 400]$. The expected and measured results have excellent agreement. This simplified algorithm struggles most around the transition point $s_c(N)$ due to the required large bond dimension (see Refs. [40,43]).

We also show representative trajectories for the active ($s = -1$) and inactive phases ($s = 1$). Each model excellently demonstrates the difference in dynamics between the two phases. The active phase displays very rapid changes with structures that allow for unconstrained dynamics. For the FA and East models this means having a large number of excitations, while SSEP requires particles to be spaced apart. Conversely, this inactive phase has just few configuration changes with highly constrained dynamics. This means minimizing the number of excitations for the FA and East resulting in the dynamics responsible for the so-called "space-time bubble" in local regions of space [5,47,73], while for SSEP we restrict the activity by clustering the particles [54,74]. To our knowledge, direct dynamical sampling of trajectories for these system sizes and values of $s \neq 0$ is different for these three models.

2. *Reference dynamics with truncated bond dimensions*

While in the extreme active/inactive limits we can achieve a good MPS description with just a bond dimension of $O(10)$, one may need a bond dimension of $O(100)$ for the more difficult regions such as around $s = 0$ [40,43]. One reason for the necessity of this high bond dimension could be that the state has longer-ranged spatial correlations. Another could be that when one runs the vMPS, we run it against some constraint in the state space. For the FA model, this is the weak constraint that restricts to the connected component of all configurations but the one with $n_i = 0$ for all i. For the SSEP, we have the stronger constraint that we are within the state space with fixed N_p particles.

The goal is to look for a state with a smaller bond dimension than we currently have which still contains all the necessary interactions, but, if necessary, discards the information which enforces the constraint. Then, by starting our CTMC algorithm in a state which satisfies the constraint, we will automatically enforce it for the rest of the trajectory, as the dynamics keeps the system in the constrained subspace.

Approximating a TN by another one with a small bond dimension is known as truncation. For MPS as we use, this can be achieved via a singular value decomposition across each bond, where only the largest $D' < D$ singular values are kept. We show this truncation in Fig. 2(a) for SSEP (as this typically requires the largest bond dimension), where we run the vMPS to at least (but higher if required) $D = 50$ to find the state $|\psi_D\rangle$, and then truncate to $|\psi_{D'}\rangle$ with the bond dimension $D' < D$. We measure the truncation error $\varepsilon = 1 - |\langle \psi_D | \psi_{D'} \rangle|^2$ between the two states, where we assume both are normalized. We find that when far from the critical point, we can describe the original state to a high accuracy with bond dimensions as

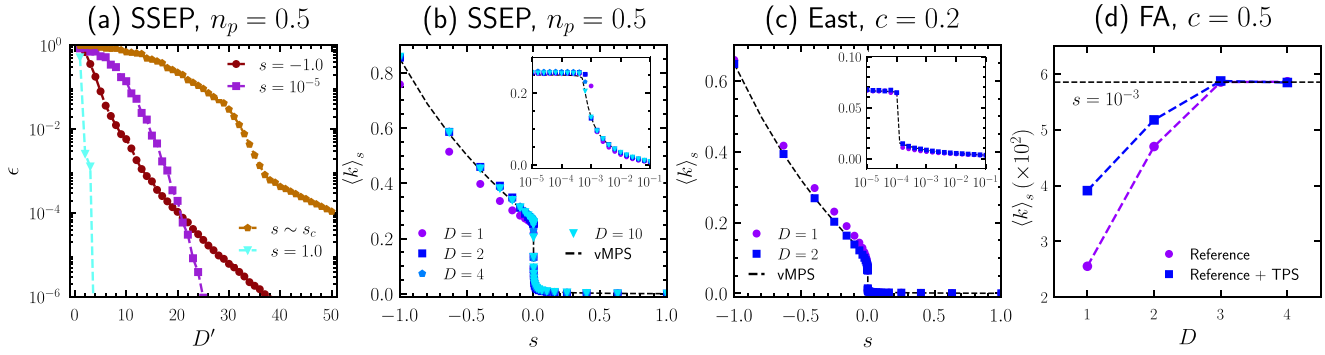


FIG. 2. Reference dynamics from a truncated MPS. All data are for the system size $N = 100$. (a) The truncation error $\varepsilon = 1 - |\langle \psi_D | \psi_{D'} \rangle|^2$ as a function of the truncated bond dimension D' for SSEP at various values of s. (b) The measured average dynamical activity $\langle k \rangle_s$ with a reference dynamics constructed with truncated MPS. The dashed line shows the expected value obtained through vMPS with a large ($D \geqslant 50$) bond dimension, and the inset shows the same but on a log scale around the critical point. (c) The same but for the East model. (d) The measured dynamical activity $\langle k \rangle_s$ as a function of D for $s = 10^{-3}$ close to the critical point s_c. The purple circles show the values measured using the reference dynamics alone, whereas the blue squares show values obtained using the reference dynamics and TPS to incorporate umbrella sampling. The dashed line shows the expected value obtained from vMPS. Each point is done for a trajectory time of $t = 100$ and $N_{sp} = 10^6$ trajectories.

small as $D' \sim 20$. Conversely, we cannot attain the same level of accuracy for $s \sim s_c$, where the state exhibits larger amounts of entanglement.

There are multiple reasons that one may want to find a state with a truncated bond dimension. The first is that our Monte Carlo algorithm scales quadratically with the bond dimension—this could hinder the convergence of time-dependent observables at large times, which can require a large sample size to be determined with sufficient accuracy. For such situations, reducing the scaling of the algorithm would be desired. Another reasoning could be that we want to investigate a system which requires a higher complexity of TN, such as a 2D system with projected entangled pair states (PEPS) [65,75]. Not only would the scaling of our CTMC algorithm increase, but so would the scaling of the variational algorithm used to find the reference dynamics. In this case, one may not be able to reach a bond dimension large enough to give a desirable variance.

We show the measured dynamical activity for SSEP and the East model (symbols), with a reference dynamics constructed from states with a truncated bond dimension in Figs. 2(b) and 2(c), and compare to the expected result from the nontruncated MPS (dashed line). Surprisingly, we find that for the most part, even for bond dimensions as small as $D = 2$, we can accurately reproduce the correct dynamical activity for each of the models. As expected, the truncation struggles mostly around the transition point. Nevertheless, we can achieve good results for the FA (not shown) and East with a truncated bond dimension of $D = 4$, and $D = 10$ for SSEP.

The calculations done thus far have been with a reference dynamics constructed using a truncated bond dimension without any trajectory reweighting. In principle, Eq. (37) is exact and thus allows for further improvements by using the umbrella

$$g(\omega) = e^{-t\theta(s)+\int dt\,\Delta\hat{R}(\omega)}. \tag{48}$$

We implement this reweighting via transition path sampling (TPS) with the *shifting* method (see Refs. [12,66] for further details). Figure 2(d) shows the results of this umbrella sampling for the FA model with an s value close to the critical point s_c. It is here the discrepancy is the largest, and we can do a more detailed analysis by looking at a larger range of bond dimensions. We see a significant improvement when using the reweighting factor Eq. (48). It might be that we could see further improvements with more TPS iterations.

The main point to take from this is that we are able to achieve accurate results for the dynamical activity (the observable we are tilting) and some local observables with a relatively small bond dimension. This of course comes at a cost, however, as when we truncate we discard some of the information that accounts for the long-ranged spatial correlations. For the case of SSEP, even though apparently we are discarding a large amount of information when truncating [cf. Fig. 2(a)], it seems that we keep the relevant information needed to reproduce the correct dynamics, but at the cost of not maintaining the conservation law. We note, however, that it is possible to explicitly implement the conservation laws in the MPS [76], but it is not clear how this will affect the quality of the reference dynamics in the CTMC algorithm.

B. Sampling rare events of finite times

For the previous results, we disregard finite-time effects by considering our sampled trajectories to be a "slice" of a larger infinite-time trajectory. We now look to incorporate these effects back into our sampling by considering the full reweighting factor

$$g(\omega) = \frac{e^{-t\theta(s)+\int dt\,\Delta\hat{R}(\omega)}}{l_s^{\mathrm{ref}}(x_0) l_s^{\mathrm{ref}}(x_K)}. \tag{49}$$

Note that previously, for a large bond dimension, the part of Eq. (49) which accounts for the difference in escape rate had a negligible effect, and could be ignored. This is not always

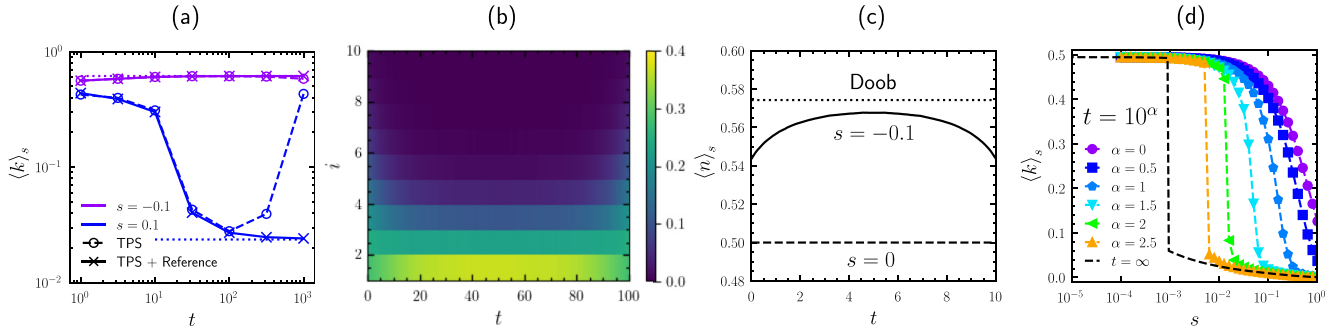


FIG. 3. Sampling finite-time trajectories. All results are done for the FA model with $c = 0.5$. (a) The measured dynamical activity $\langle k \rangle$ as a function of time t for $s = -0.1$ (top) and $s = 0.1$ (bottom) for $N_{\rm sp} = 10^6$ trajectories and system size $N = 40$. The circles show the values obtained via TPS with the normal dynamics, and crosses TPS with the reference dynamics. The dotted lines show the expected value at infinite times. (b) The local occupations $\langle n_i \rangle$ as a function of time in the inactive regime $s = 0.1$ and $N = 40$. (c) The average excitation density $\langle n \rangle$ as a function of time in the active regime, $s = -0.1$. The value approaches the expected value in the Doob dynamics (dotted line) in the bulk, but moves towards the equilibrium value ($s = 0$, dashed line) at the edges and $N = 40$. (d) The dynamical activity as a function of s and time t. The data for $t = \infty$ are obtained directly from the MPS, whereas finite t is obtained using TPS. Note that the sharp drop in activity shifts with time. The dynamics are run at the system size $N = 100$.

the case here, as the umbrella sampling at the time edges of the trajectory causes the system to visit configurations which are atypical in the Doob dynamics, and not well described by our MPS approximation.

As a proof of principle, we start by comparing results from TPS with the original dynamics against TPS with the reference dynamics for a small system size $N = 40$, and a variety of times, as is shown for the FA model at $s = \pm 0.1$ in Fig. 3(a). For small times, both show excellent agreement. For large times, however, the normal dynamics struggles to correctly account for the expected activity shown by the dotted lines, a result of the exponential time dependence in Eq. (26) (as $\hat{K}$ is time extensive). While sampling with our reference dynamics reduces the exponential cost in time, the time edges still suffer from an exponential sampling difficulty in the system size. This is most noticeable for the inactive phase, where each model exhibits an exponential localization [40,43,77] at the spatial edge(s) of the system. This causes the $l_s^{\rm ref}(x)$ values to exponentially vary. Nevertheless, it is still a significant improvement on the previously exponential cost in space, time, and s.

The average occupations $\langle n_i \rangle_s$ (at site i) for $s = 0.1$ and $t = 100$ is shown in Fig. 3(b), while Fig. 3(c) shows the average excitation density $\langle n \rangle_s = N^{-1} \sum_i \langle n_i \rangle_s$ for $s = -0.1$ and $t = 10$. It is here the time-edge effects become obvious; we start at a state which lies somewhere between the expected $s = 0$ (dashed line) dynamics and the expected long-time dynamics, which depends on the whole spectrum of $\mathbb{W}_s$, as well as the total trajectory time. The system quickly evolves and resembles the Doob dynamics. Note that at the end of trajectory, it is again described by the original probability vector, as is expected due to the time symmetry in Eq. (37).

Finally, Fig. 3(d) shows the average dynamical activity as a function of s and time t. We show the expected activity in the infinite time limit $t = \infty$ as a black dashed line, and the measured activity for finite times as symbols. Notice that as time decreases, the drop in activity becomes less sharp. Furthermore, the transition from the active to inactive phase happens at decreasing s. While the methods presented here could allow for a detailed investigation into the temporal scaling of the critical point [52,53,71,72], doing so for desirable system sizes would be at a large computational cost. We hope to investigate this more extensively using time-evolution methods (see, e.g., Refs. [30,65,78]).

V. CONCLUSIONS

We have expanded on previous applications of TNs to classical constrained models [40–43], using the MPS approximation of the leading eigenstates of a tilted stochastic generator in 1D to construct a reference dynamics which well approximates the exact Doob dynamics. This allows us to (nearly) optimally sample the rare events of 1D constrained systems with just a polynomial cost in both space and time, rather than the exponential cost of most sampling methods. We have demonstrated here the efficiency of this approach by generating tilted trajectory ensembles for the FA and East KCMs and the symmetric simple exclusion process. Our simulations are for sizes and times different for such large deviation studies.

Furthermore, our results show that it is possible to obtain an accurate dynamics away from the dynamical transitions of the models we studied with a truncated bond dimension, which enables close to optimal sampling simulations at little cost. Further extensions of our work include generalizing our methods to higher dimensions, for example, by using two-dimensional variational tensor network techniques, such as PEPS [65,75] to approximate the leading eigenvectors of 2D classical generators, as is done in Ref. [42]. From the associated leading eigenvectors, as we have shown here, we can in turn construct a reference dynamics which is nearly optimal for sampling rare trajectories. While PEPS algorithms do not currently allow for bond dimensions comparable to vMPS, they remain a fruitful area of research which is constantly being improved on [79–87]. Recent works [88] have shown the effectiveness of using recurrent neural networks (RNNs) to approximate the leading eigenstates of tilted generators in two dimensions. The methods presented here can be

generalized to RNN to allow for the efficient sampling of 2D rare events.

Another area that deserves exploration is to apply similar TN methods to systems which do not obey detailed balance, and for which their generators cannot be brought to a Hermitian form. While this would damper the effectiveness of variational algorithms, approaches based on time evolution may offer a promising solution (see, e.g., Refs. [30,65,78]). Such approaches could also offer further insights into intermediate-time rare events, where both usual sampling methods and large deviation approaches fall short. We hope to report on such studies in the near future.

ACKNOWLEDGMENTS

We acknowledge financial support from EPSRC Grant No. EP/R04421X/1 and the Leverhulme Trust Grant No. RPG-2018-181. M.C.B. acknowledges support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2111 – 390814868. J.P.G. is grateful to All Souls College, Oxford, for support through a Visiting Fellowship. We acknowledge access to the University of Nottingham Augusta HPC service. Much of the numerical data during the later stages of this work was acquired using the ITensor library [89].

APPENDIX A: STEADY STATE SOLUTION IN THE REFERENCE DYNAMICS

The generator of the reference dynamics defined by Eqs. (34) and (35) can be written as

$$\mathbb{W}_s^{\mathrm{ref}} = \sum_{x,x'\neq x} e^{-s} w_{x\to x'} \frac{l_s^{\mathrm{ref}}(x')}{l_s^{\mathrm{ref}}(x)} [|x'\rangle\langle x| - |x\rangle\langle x|]. \tag{A1}$$

By definition, the stationary state $|\mathrm{ss}\rangle_s^{\mathrm{ref}} = \sum_z P_s^{\mathrm{ref}}(z)\,|z\rangle$ is annihilated by Eq. (A1). It follows that

$$\begin{aligned}\mathbb{W}_s^{\mathrm{ref}}\,|\mathrm{ss}\rangle_s^{\mathrm{ref}} &= \sum_{x,x'\neq x} e^{-s} w_{x\to x'} \frac{l_s^{\mathrm{ref}}(x')}{l_s^{\mathrm{ref}}(x)} P_s^{\mathrm{ref}}(x)[|x'\rangle - |x\rangle] \\ &= \sum_{x,x'\neq x} e^{-s}\Bigg[w_{x\to x'} \frac{l_s^{\mathrm{ref}}(x')}{l_s^{\mathrm{ref}}(x)} P_s^{\mathrm{ref}}(x) \\ &\quad - w_{x'\to x} \frac{l_s^{\mathrm{ref}}(x)}{l_s^{\mathrm{ref}}(x')} P_s^{\mathrm{ref}}(x')\Bigg] |x'\rangle \\ &= 0, \end{aligned} \tag{A2}$$

where we have used a change of variables in the second and third line. Let us assume our original dynamics obeys detailed balance, and that the state space is connected (that is, the dynamics is irreducible). Then it follows that if $w_{x\to x'} = 0$, so does $w_{x'\to x}$, in which Eq. (A2) is satisfied. Otherwise, we must have that

$$\frac{P_s^{\mathrm{ref}}(x)}{P_s^{\mathrm{ref}}(x')} = \frac{w_{x'\to x}}{w_{x\to x'}} \frac{l_s^{\mathrm{ref}}(x)^2}{l_s^{\mathrm{ref}}(x')^2}. \tag{A3}$$

Given detailed balance we can use a similarity transformation to write the generator in a Hermitian form [cf. Eq. (16)]. In particular, let us define the diagonal transformation matrix as

$$Q = \sum_z Q(z)\,|z\rangle\langle z|\,. \tag{A4}$$

One can easily show that for $\mathbb{H}$ to be Hermitian, we must have

$$\frac{Q(x)^2}{Q(x')^2} = \frac{w_{x'\to x}}{w_{x\to x'}}. \tag{A5}$$

Substituting this back into Eq. (A3), we find

$$\frac{P_s^{\mathrm{ref}}(x)}{P_s^{\mathrm{ref}}(x')} = \frac{Q(x)^2\, l_s^{\mathrm{ref}}(x)^2}{Q(x')^2\, l_s^{\mathrm{ref}}(x')^2}, \tag{A6}$$

and it follows the stationary state is given by

$$|\mathrm{ss}\rangle_s^{\mathrm{ref}} = \sum_x l_s^{\mathrm{ref}}(x)^2 Q(x)^2\,|x\rangle\,. \tag{A7}$$

For the case of our MPS dynamics, we defined $\langle l_s| = \langle\psi_s|\,Q^{-1}$ of our solution $\langle\psi_s|$. It follows that Eq. (A7) can be written as $|\mathrm{ss}\rangle_s^{\mathrm{ref}} = \sum_x \psi_s^{\mathrm{ref}}(x)^2\,|x\rangle$, where $\psi_s^{\mathrm{ref}}(x) = \langle x|\psi_s\rangle$.

APPENDIX B: SAMPLING VARIANCE IN THE REFERENCE DYNAMICS

We start by assuming that we are always at the stationary state of the dynamics. This allows us to calculate the trajectory ensemble average of some observable (per unit time) as the average over all configurations with respect to the stationary state,

$$\langle\hat{O}\rangle_{\mathrm{ref}} \equiv \frac{1}{t}\left\langle \int_0^t dt'\hat{O}(t')\right\rangle_{\mathrm{ref}} = \left\langle l_s^{\mathrm{ref}}\middle|\hat{O}\middle|r_s^{\mathrm{ref}}\right\rangle. \tag{B1}$$

The aim is to calculate the expectation value and the variance of the time-integrated difference in escape rates [cf. Eq. (39)]. Using Eq. (B1), we can write

$$\begin{aligned}\langle\Delta\hat{R}\rangle_{\mathrm{ref}} &= \sum_{x,y\neq x} l_x\left(\frac{l_y}{l_x}e^{-s}\omega_{x\to y} - \omega_{x\to y}\right) r_x \\ &= \sum_{x,y\neq x} l_y e^{-s}\omega_{x\to y} r_x - l_x\omega_{x\to y} r_x \\ &= \left\langle l_s^{\mathrm{ref}}\middle|\,\mathbb{W}_s\middle|r_s^{\mathrm{ref}}\right\rangle = \theta^{\mathrm{ref}}(s), \end{aligned} \tag{B2}$$

where we have written $l_x \equiv l_s^{\mathrm{ref}}(x)$ and $r_x \equiv r_s^{\mathrm{ref}}(x)$ for brevity. Performing the same calculation for $\Delta\hat{R}^2$, we find

$$\langle\Delta\hat{R}^2\rangle_{\mathrm{ref}} = \left\langle l_s^{\mathrm{ref}}\middle|\,\mathbb{W}_s^{\,2}\middle|r_s^{\mathrm{ref}}\right\rangle, \tag{B3}$$

giving the variance

$$\mathrm{Var}_{\mathrm{ref}}\Delta\hat{R} \equiv \langle\Delta\hat{R}^2\rangle_{\mathrm{ref}} - \langle\Delta\hat{R}\rangle_{\mathrm{ref}}^2 = \delta E^2, \tag{B4}$$

where δE^2 is the measured variance of the MPS used to construct the reference dynamics with respect to the tilted generator (or tilted Hamiltonian).

We are now in a position to estimate the sampling error Eq. (39). From Eqs. (B2) and (B4) we have that the integrated difference in escape rate,

$$\int_0^t dt'\,\Delta\hat{R}(t'), \tag{B5}$$

has the average

$$t\,\langle \Delta\hat{R}\rangle = t\theta^{\mathrm{ref}}(s), \tag{B6}$$

and variance

$$t\delta E^2. \tag{B7}$$

If we also assume this integrated difference to be normally distributed, then we get Eq. (39),

$$\epsilon^2_{\mathrm{ref}} = \frac{1}{N_{\mathrm{sp}}}\left[\frac{\left\langle e^{2\int dt'\Delta\hat{R}}\right\rangle_{\mathrm{ref}}}{\left\langle e^{\int dt'\Delta\hat{R}}\right\rangle^2_{\mathrm{ref}}} - 1\right] \approx \frac{e^{t\delta E^2}-1}{N_{\mathrm{sp}}}. \tag{B8}$$

[1] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, Oxford, U.K., 1987).
[2] H. Touchette, Phys. Rep. **478**, 1 (2009).
[3] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[4] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
[5] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).
[6] J. P. Garrahan, Physica A **504**, 130 (2018).
[7] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).
[8] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[9] V. Lecomte and J. Tailleur, J. Stat. Mech. (2007) P03004.
[10] F. Cérou, A. Guyader, and M. Rousset, Chaos **29**, 043108 (2019).
[11] C. Pérez-Espigares and P. I. Hurtado, Chaos **29**, 083106 (2019).
[12] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[13] U. Ray, G. K.-L. Chan, and D. T. Limmer, J. Chem. Phys. **148**, 124120 (2018).
[14] U. Ray, G. K.-L. Chan, and D. T. Limmer, Phys. Rev. Lett. **120**, 210602 (2018).
[15] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitelam, Phys. Rev. E **97**, 032123 (2018).
[16] A. Guyader and H. Touchette, J. Stat. Phys. **181**, 551 (2020).
[17] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Phys. Rev. E **93**, 062123 (2016).
[18] G. Ferré and H. Touchette, J. Stat. Phys. **172**, 1525 (2018).
[19] T. H. E. Oakes, A. Moss, and J. P. Garrahan, Mach. Learn.: Sci. Technol. **1**, 035004 (2020).
[20] D. C. Rose, J. F. Mair, and J. P. Garrahan, New J. Phys. **23**, 013013 (2021).
[21] E. Gillman, D. C. Rose, and J. P. Garrahan, arXiv:2002.05185.
[22] D. Simon, J. Stat. Mech.: Theory Exp. (2009) P07017.
[23] V. Popkov, G. M. Schütz, and D. Simon, J. Stat. Mech.: Theory Exp. (2010) P10007.
[24] R. L. Jack and P. Sollich, Prog. Theor. Phys. Suppl. **184**, 304 (2010).
[25] R. Chetrite and H. Touchette, Ann. Henri Poincaré **16**, 2005 (2015).
[26] J. P. Garrahan, J. Stat. Mech.: Theory Exp. (2016) 073208.
[27] F. Carollo, J. P. Garrahan, I. Lesanovsky, and C. Pérez-Espigares, Phys. Rev. A **98**, 010103(R) (2018).
[28] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).
[29] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett. **93**, 227205 (2004).
[30] U. Schollwöck, Ann. Phys. **326**, 96 (2011).
[31] R. Orús, Nat. Rev. Phys. **1**, 538 (2019).
[32] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **100**, 030504 (2008).
[33] I. P. McCulloch, J. Stat. Mech. (2007) P10014.
[34] B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, New J. Phys. **12**, 025012 (2010).
[35] A. Honecker and I. Peschel, J. Stat. Phys. **88**, 319 (1997).
[36] Y. Hieida, J. Phys. Soc. Jpn. **67**, 369 (1998).
[37] A. Kemper, A. Gendiar, T. Nishino, A. Schadschneider, and J. Zittartz, J. Phys. A: Math. Gen. **36**, 29 (2002).
[38] K. Ueda, R. Otani, Y. Nishio, A. Gendiar, and T. Nishino, J. Phys. Soc. Jpn. **74**, 111 (2005).
[39] M. Gorissen, J. Hooyberghs, and C. Vanderzande, Phys. Rev. E **79**, 020101(R) (2009).
[40] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).
[41] P. Helms, U. Ray, and G. K.-L. Chan, Phys. Rev. E **100**, 022101 (2019).
[42] P. Helms and G. K.-L. Chan, Phys. Rev. Lett. **125**, 140601 (2020).
[43] L. Causer, I. Lesanovsky, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **102**, 052132 (2020).
[44] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[45] J. P. Garrahan, P. Sollich, and C. Toninelli, in *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media*, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos, International Series of Monographs on Physics (Oxford University Press, Oxford, UK, 2011).
[46] J. Jäckle and S. Eisinger, Z. Phys. B **84**, 115 (1991).
[47] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[48] G. Biroli and J. P. Garrahan, J. Chem. Phys. **138**, 12A301 (2013).
[49] K. Mallick, Physica A **418**, 17 (2015).
[50] R. A. Blythe and M. R. Evans, J. Phys. A **40**, R333 (2007).
[51] T. Bodineau and B. Derrida, C. R. Acad. Sci. **8**, 540 (2007).
[52] C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van Wijland, Phys. Rev. E **78**, 021122 (2008).
[53] T. Bodineau, V. Lecomte, and C. Toninelli, J. Stat. Phys. **147**, 1 (2012).
[54] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).
[55] G. H. Fredrickson and H. C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984).
[56] C. Maes, Phys. Rep. **850**, 1 (2020).
[57] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).

[58] S. Östlund Stellanstlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
[59] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[60] M. Zwolak and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
[61] C. Hubig, I. P. McCulloch, and U. Schollwöck, Phys. Rev. B **95**, 035129 (2017).
[62] D. E. Parker, X. Cao, and M. P. Zaletel, Phys. Rev. B **102**, 035147 (2020).
[63] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[64] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
[65] F. Verstraete, V. Murg, and J. I. Cirac, Adv. Phys. **57**, 143 (2008).
[66] T. Oakes, S. Powell, C. Castelnovo, A. Lamacraft, and J. P. Garrahan, Phys. Rev. B **98**, 064302 (2018).
[67] N. Tizón-Escamilla, V. Lecomte, and E. Bertin, J. Stat. Mech.: Theory Exp. (2019) 013201.
[68] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).
[69] A. J. Ferris and G. Vidal, Phys. Rev. B **85**, 165146 (2012).
[70] S. Iblisdir, New J. Phys. **16**, 103022 (2014).
[71] T. Bodineau and C. Toninelli, Commun. Math. Phys. **311**, 357 (2012).
[72] T. Nemoto, R. L. Jack, and V. Lecomte, Phys. Rev. Lett. **118**, 115702 (2017).
[73] J. P. Garrahan and D. Chandler, Phys. Rev. Lett. **89**, 035704 (2002).
[74] V. Lecomte, J. P. Garrahan, and F. van Wijland, J. Phys. A **45**, 175001 (2012).
[75] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orús, Phys. Rev. B **92**, 035142 (2015).
[76] C. Hubig, SciPost Phys. **5**, 47 (2018).
[77] N. Pancotti, G. Giudice, J. I. Cirac, J. P. Garrahan, and M. C. Bañuls, Phys. Rev. X **10**, 021051 (2020).
[78] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Ann. Phys. **411**, 167998 (2019).
[79] P. Corboz, Phys. Rev. B **93**, 045116 (2016).
[80] P. Corboz, Phys. Rev. B **94**, 035133 (2016).
[81] L. Vanderstraeten, J. Haegeman, P. Corboz, and F. Verstraete, Phys. Rev. B **94**, 155123 (2016).
[82] P. Corboz, P. Czarnik, G. Kapteijns, and L. Tagliacozzo, Phys. Rev. X **8**, 031031 (2018).
[83] M. Rader and A. M. Läuchli, Phys. Rev. X **8**, 031030 (2018).
[84] L. Vanderstraeten, J. Haegeman, and F. Verstraete, Phys. Rev. B **99**, 165121 (2019).
[85] P. Czarnik, J. Dziarmaga, and P. Corboz, Phys. Rev. B **99**, 035115 (2019).
[86] R. Haghshenas, M. J. O'Rourke, and G. K.-L. Chan, Phys. Rev. B **100**, 054404 (2019).
[87] M. J. O'Rourke and G. K.-L. Chan, Phys. Rev. B **101**, 205142 (2020).
[88] C. Casert, T. Vieijra, S. Whitelam, and I. Tamblyn, arXiv:2011.08657.
[89] M. Fishman, S. R. White, and E. M. Stoudenmire, arXiv:2007.14822.

Chapter 7

Finite time large deviations via matrix product states

The following work is from the publication *"Finite time large deviations via matrix product states"* by Luke Causer, Mari Carmen Bañuls and Juan P. Garrahan, in Physical Review Letters 128 (9), 090605.

The previous efforts to investigate biased finite time trajectory ensembles included using the long time Doob dynamics alongside TPS to sample the finite time dynamics. Here, the finite time trajectory ensembles are directly targeted by simulating the time evolution of a probability vector under the biased master equation using MPS. This allows for a systematic way to study dynamical phase transitions at arbitrary times. It is also shown how the resulting MPS can be used to construct a dynamics which directly targets the finite time trajectory ensembles, thus generalizing the methods of the previous section to finite times.

The work shown here demonstrates how TNs can be used beyond ground state approximation methods to study stochastic dynamics, while also investigating the finite time properties of dynamical phase transitions. The methods are applied to the FA and East KCMs, and also the SSEP. For each model, the dynamical phase diagram is uncovered, allowing the determination of the temporal scaling exponent. Previous works [201, 202] have shown that the s-ensemble can be used to prepare glassy states similar to those found in the hierarchical relaxation of glasses. The methods introduced in this paper are used to demonstrate the trajectory times one would need to simulate in the s-ensemble to prepare such states.

Further details of the numerical methods and an error analysis can be found in Appendix A.

Finite Time Large Deviations via Matrix Product States

Luke Causer,[1,2] Mari Carmen Bañuls,[3,4] and Juan P. Garrahan[1,2]
[1]School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom
[2]Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham NG7 2RD, United Kingdom
[3]Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany
[4]Munich Center for Quantum Science and Technology (MCQST), Schellingstrasse 4, D-80799 München, Germany

(Received 25 October 2021; accepted 15 February 2022; published 4 March 2022)

Recent work has shown the effectiveness of tensor network methods for computing large deviation functions in constrained stochastic models in the infinite time limit. Here we show that these methods can also be used to study the statistics of dynamical observables at arbitrary finite time. This is a harder problem because, in contrast to the infinite time case, where only the extremal eigenstate of a tilted Markov generator is relevant, for finite time the whole spectrum plays a role. We show that finite time dynamical partition sums can be computed efficiently and accurately in one dimension using matrix product states and describe how to use such results to generate rare event trajectories on demand. We apply our methods to the Fredrickson-Andersen and East kinetically constrained models and to the symmetric simple exclusion process, unveiling dynamical phase diagrams in terms of counting field and trajectory time. We also discuss extensions of this method to higher dimensions.

DOI: 10.1103/PhysRevLett.128.090605

Introduction.—Large deviation (LD) theory provides a powerful framework to investigate the statistical fluctuations of time-averaged observables in stochastic systems (for reviews, see, e.g., Refs. [1–4]). At long times (assuming finite correlation times) the probabilities of such observables obey a LD principle, and the corresponding scaled cumulant generating function (see below) can be retrieved from the leading eigenvalue of the "tilted" (or deformed or biased) generator [1]. For large systems, estimating this eigenvalue is difficult, so one resorts to sampling the corresponding biased trajectory ensemble via numerical methods, such as trajectory importance sampling [5–8], population dynamics [9–11], optimal control [12–18], or machine learning approaches [19–24]. For lattice models, recent work has focused on the use of tensor network (TN) techniques to approximate the leading eigenvector of the tilted generator through variational means [25–27] or power methods [28].

A harder problem is that of computing the statistics of time-averaged observables for *finite* time [29–31]. The reason is that away from the long time limit the corresponding dynamical partition sums (i.e., moment generating functions) do not obey a LD principle in time—only obeying a LD principle in space for large sizes—and, as a consequence, they are not determined only by the leading eigenvalue of the tilted generator, but by their whole spectrum. If time is very short, one can get away with direct sampling, but for intermediate times the usual sampling approaches fall short [32]. Here we develop a scheme to study these rare events by implementing well-developed TN techniques to simulate time evolution. This allows us to calculate dynamical partition functions for trajectories of arbitrary time extent. Furthermore, we show how to use the results here to directly simulate stochastic trajectories in finite time tilted ensembles at small computational cost, thus generalizing the method of Ref. [32].

We focus for concreteness on one-dimensional kinetically constrained models (KCMs)—often used in the modeling of structural glasses [2,33–35]—specifically the Fredrickson-Andersen (FA) [36] and the East [37] models, as well as on the symmetric simple exclusion process (SSEP). Both KCMs and SEPs display phase transitions in their dynamical LDs in the long time limit [38–44]. With the methods developed here, we are able to construct the dynamical phase diagram both as a function of counting field and of trajectory time, determining finite time scaling of active-inactive phase transitions in these models and uncovering the emergence with time of the correlated structure of the active phase in the East model and the SSEP.

Models.—The three models we consider live in a one-dimensional lattice of N sites, with binary variables $n_j = 0, 1$ for each $j = 1, \ldots, N$, evolving under continuous-time Markov dynamics with local transitions. The probability for each configuration $|x\rangle = |n_1, \ldots, n_N\rangle$ at time t, encoded in a vector $|P(t)\rangle = \sum_x P(t,x)|x\rangle$, evolves deterministically via a master equation, $\partial_t|P(t)\rangle = \mathbb{W}|P(t)\rangle$, where $\mathbb{W}$ is the Markov generator. Being a stochastic operator, $\mathbb{W}$ has a structure $\mathbb{W} = \mathbb{K} - \mathbb{R}$, with an off-diagonal matrix of transition rates $\mathbb{K}$ and a diagonal matrix of positive escape rates $\mathbb{R}$.

For the KCMs the generator reads

$$\mathbb{W}^{\mathrm{KCM}} = \sum_i f_i [c\sigma_i^+ + (1-c)\sigma_i^- - c(1-n_i) - (1-c)n_i], \quad (1)$$

where $c \in (0, 1/2]$ defines the site occupation at equilibrium, and $\sigma_i^{\pm}$ are the Pauli raising and lowering operators at site i. Spin flips are only permitted if the kinetic constraint f_i is satisfied. We consider two paradigmatic KCMs, the Fredrickson-Andersen [36] model and the East [37] model. They are defined by the respective constraint functions

$$f_i^{\mathrm{FA}} = n_{i-1} + n_{i+1}, \qquad f_i^{\mathrm{East}} = n_{i-1}. \quad (2)$$

We consider lattices with open boundary conditions (OBCs) to allow for efficient tensor network contractions. For numerical convenience, we choose the fixed boundaries $n_1 = n_N = 1$ for the FA [45] model and $n_1 = 1$ for the East model. The corresponding stationary states (SS) are product states,

$$|\mathrm{SS}^{\mathrm{FA}}\rangle = |1\rangle \otimes [(1-c)|0\rangle + c|1\rangle]^{\otimes N-2} \otimes |1\rangle, \quad (3)$$

$$|\mathrm{SS}^{\mathrm{East}}\rangle = |1\rangle \otimes [(1-c)|0\rangle + c|1\rangle]^{\otimes N-1}. \quad (4)$$

The third model we consider is the symmetric simple exclusion process whose generator reads

$$\mathbb{W}^{\mathrm{SSEP}} = \frac{1}{2}\sum_i [\sigma_i^+\sigma_{i+1}^- + \sigma_i^-\sigma_{i+1}^+ - (n_i + n_{i+1}) + 2n_i n_{i+1}]. \quad (5)$$

For the SSEP we consider OBCs such that particles can enter and leave at the boundaries with rate $1/4$. The stationary state is $|\mathrm{SS}^{\mathrm{SSEP}}\rangle = 2^{-N}|-\rangle = 2^{-N}\sum_x |x\rangle$, with the "flat" state $\langle -|$ being the leading left eigenvector of each generator above.

Dynamical rare events and LDs.—We now consider the ensemble of all possible trajectories $\{\omega_\alpha\}$ with trajectory time t, where $\omega_\alpha = \{x_0 \to x_{t_1} \to \cdots \to x_t\}$ defines jumps to configurations x_{t_k} at times t_k. The probability of observing the value $K(\omega_\alpha) = K$ of some time-integrated observable K is

$$P_t(K) = \sum_\alpha \pi(\omega_\alpha)\delta[K(\omega_\alpha) - K], \quad (6)$$

where $\pi(\omega_\alpha)$ defines the probability of observing the trajectory. The corresponding moment generating function (or trajectory partition sum) is

$$Z_t(s) = \sum_\alpha \pi(\omega_\alpha) e^{-sK(\omega_\alpha)} \quad (7)$$

where the "counting field" s is conjugate to the observable.

For large times, both Eqs. (6) and (7) take a large deviation form in time [1,38,41,46], $P_t(K) \asymp e^{-t\varphi(K/t)}$ and $Z_t(s) \asymp e^{t\theta(s)}$. The LD rate function $\varphi(K/t)$ and the scaled cumulant generating function $\theta(s)$ play the roles of a trajectory entropy density and a free-energy density, respectively, and are related through the Legendre transform $\theta(s) = -\min_k[sk + \varphi(k)]$.

The partition sum Eq. (7) can be written as

$$Z_t(s) = \langle -|e^{t\mathbb{W}_s}|\mathrm{ss}\rangle \quad (8)$$

in terms of the tilted generator $\mathbb{W}_s$ [1,38,41,46]. In what follows we focus on the *dynamical* activity [38,41,46,47]), that is, the total number of spin flips, as a trajectory observable. In this case, $\mathbb{W}_s = e^{-s}\mathbb{K} - \mathbb{R}$. While for large times all that is needed to determine Eq. (8) is the dominant eigenstate of $\mathbb{W}_s$, for finite times the whole spectrum of $\mathbb{W}_s$ is required.

Finite time statistics from MPS.—The models we consider obey detailed balance. This allows us to write $\mathbb{W}_s$ in a Hermitian form through a similarity transformation independent of s [41], $\mathbb{H}_s = \mathbb{P}^{-1/2}\mathbb{W}_s\mathbb{P}^{1/2}$, where $\mathbb{P}^{1/2}$ is a diagonal matrix of probability amplitudes at equilibrium (for the SSEP, $\mathbb{P}$ is the identity). As a consequence, the leading eigenvalue of $\mathbb{H}$ obeys a Rayleigh-Ritz variational principle, allowing the application of variational methods such as the density matrix renormalization group [48]. We then write Eq. (8) as

$$Z_t(s) = \langle \psi_0|e^{t\mathbb{H}_s}|\psi_0\rangle, \quad (9)$$

where $|\psi_0\rangle = \mathbb{P}^{-1/2}|\mathrm{ss}\rangle = [\langle -|\mathbb{P}^{1/2}]^\dagger$. It is useful to define the time-evolved vector $|\psi_\tau\rangle = e^{\tau\mathbb{H}_s}|\psi_0\rangle$ $(\tau \le t)$. The partition function can then be written as $Z_t(s) = \langle \psi_{t-\tau}|\psi_\tau\rangle$ and, in particular, can be determined by only evolving the vector by $\tau = t/2$.

The average dynamical activity (per unit time and site) of the biased ensemble of trajectories follows from the partition sum,

$$k(s) = -\frac{1}{Nt}\frac{d}{ds}\log[Z_t(s)]. \quad (10)$$

We can also calculate time-dependent configurational observables for any $0 \le \tau \le t$,

$$\langle O(\tau)\rangle_s = Z_t(s)^{-1}\langle \psi_0|e^{(t-\tau)\mathbb{H}_s}Oe^{\tau\mathbb{H}_s}|\psi_0\rangle = Z_t(s)^{-1}\langle \psi_{t-\tau}|O|\psi_\tau\rangle. \quad (11)$$

In order to compute the time-evolved state $|\psi_t\rangle$, we use methods from quantum many-body physics, in particular,

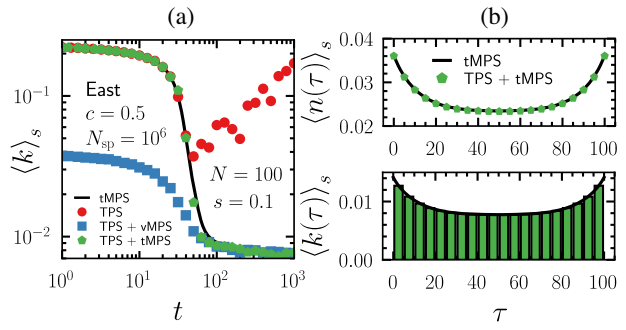


FIG. 1. Demonstration of the methods. East model at $c = 0.5$, $N = 100$, and $s = 0.1$. (a) Dynamical activity $\langle k \rangle$ from tMPS (black line), TPS with no auxiliary dynamics (red circles), TPS with the LD eigenvector auxiliary dynamics via vMPS (blue squares), and TPS with a tMPS reference dynamics (green pentagons). (b) Time-dependent occupations (top) and instantaneous activity (bottom) from MPS time evolution (black line) from direct sampling with a tMPS auxiliary dynamics (green pentagons and bars).

matrix product states (MPS) (for reviews, see Refs. [49,50]) [51]. Here we use both variational optimization of MPS (vMPS, e.g., [25,49]) and time-evolved MPS (tMPS, e.g., [53]). Notice that, for long times, $|\psi_\tau\rangle$ becomes close to the leading eigenvector of $\mathbb{W}_s$. We exploit this fact to simulate evolution for long times with higher precision; see Supplemental Material [54] for details.

In Ref. [32], we used the MPS approximation (from vMPS) to the leading eigenstate of $\mathbb{H}_s$ to construct a near-optimal dynamics, which when supplemented with trajectory importance sampling [specifically, transition path sampling [5] (TPS)] allowed us to efficiently simulate trajectories in large time tilted ensembles. Here we apply the same scheme, but instead use the time-evolved state $|\psi_{t/2}\rangle$. We construct a time-independent dynamics that approximates the optimal (or Doob) dynamics at the center of finite time trajectories under tilting [57].

Figure 1(a) compares various sampling methods in the East model at $s > 0$. The dynamics is active at short times (due to initial conditions) and inactive at large times [25,38,41]. We show the activity from the partition sum calculated via MPS time evolution (black line) as a function of trajectory length. We also show sampling with TPS with the original dynamics (red circles); this method only accounts for the dynamical activity $\langle k \rangle$ at short times and fails at long times. The methods introduced in Ref. [32] construct the long time optimal (Doob) dynamics with the approximate leading eigenstate from vMPS. We then apply TPS with this dynamics to sample trajectories for arbitrary time. This accounts for $\langle k \rangle$ at long times [32], but fails at short times. If we adopt this method, but replace the MPS used in the auxiliary dynamics with the time-evolved state (green pentagons), we get accurate results for the activity for all trajectory lengths. Despite the fact that the exact Doob dynamics for finite time is, in general, time-dependent [58], this latter approach with a time-independent dynamics for each trajectory length t is efficient enough for TPS to converge to the actual finite time tilted ensemble, thus correcting any discrepancies. In the Supplemental Material, we provide a detailed comparison [54]. Figure 1(b) shows the averaged time-dependant occupations (top) $\langle n(\tau) \rangle_s$ and instantaneous activity (bottom) $\langle k(\tau) \rangle_s$ for some fixed trajectory time $t = 100$, generated from the s ensemble at $s = 0.1$ for both tMPS and tMPS + TPS.

Finite time scaling of active-inactive transition.—The three models we study here display a LD phase transition [38,40,41,43] in the long time and large size limit between a dynamical phase where activity is extensive in space and one where activity is subextensive. The finite size scaling analysis of this transition in the long time limit has been studied theoretically [40,42,59–61] and numerically [25,60]: for finite size, the active-inactive transition is smoothed into a sharp crossover located at $s_c(N, t = \infty) > 0$, which decreases as an inverse power of the system size. In general, however, the location of the transition point depends both on time and size, $s_c(N, t)$, but a detailed numerical analysis of the finite time scaling has not been possible to date due to the difficulty of simulating efficiently rare trajectories at intermediate times [32]. With the approach presented above we can now investigate this issue in detail.

Figure 2(a) shows the dynamical activity $k(s)$ as a function of s and inverse time t^{-1} (East model, top row; FA model, middle; SSEP, bottom). There is a transition from a high activity (light) to low activity (dark) as s is increased, which becomes sharper and moves to smaller s with increasing time. The point $s_c(N, t)$ (shown by the red dashed line) is that of the peak in the dynamical susceptibility $\chi(s, t) = dk(s)/ds$. These dynamical phase diagrams are reminiscent of those of (first-order) quantum phase transitions [62], with s as an applied field and the inverse time as temperature.

The scaling of the transition point is shown as a function of (inverse) time for multiple system sizes $N \in [20, 200]$ in Fig. 2(b). For small times the transition point scales approximately as $s_c \sim t^{-1}$ for the three models. When time becomes large enough, finite size effects start to dominate. For simplicity, we use the approximate form

$$s_c(N, t) \approx s_c(N) + s_c(t), \tag{12}$$

where $s_c(N) \sim N^{-\alpha}$ can be extracted from vMPS [25,27]. For the FA and East models, the exponent $\alpha > 1$ [25], while for the SSEP we find the expected $\alpha \approx 2$ [40]. In Fig. 2(c), we show how the $s_c(N, t)$ curves can be collapsed, allowing us to estimate $s_c(t) \sim t^{-\beta}$. We find $\beta \approx 1$ for all models.

Also important to the rare event statistics is the probability distribution of the dynamical activity $P_t(K)$. While for finite times $Z_t(s)$ and $P_t(K)$ do not obey a LD principle

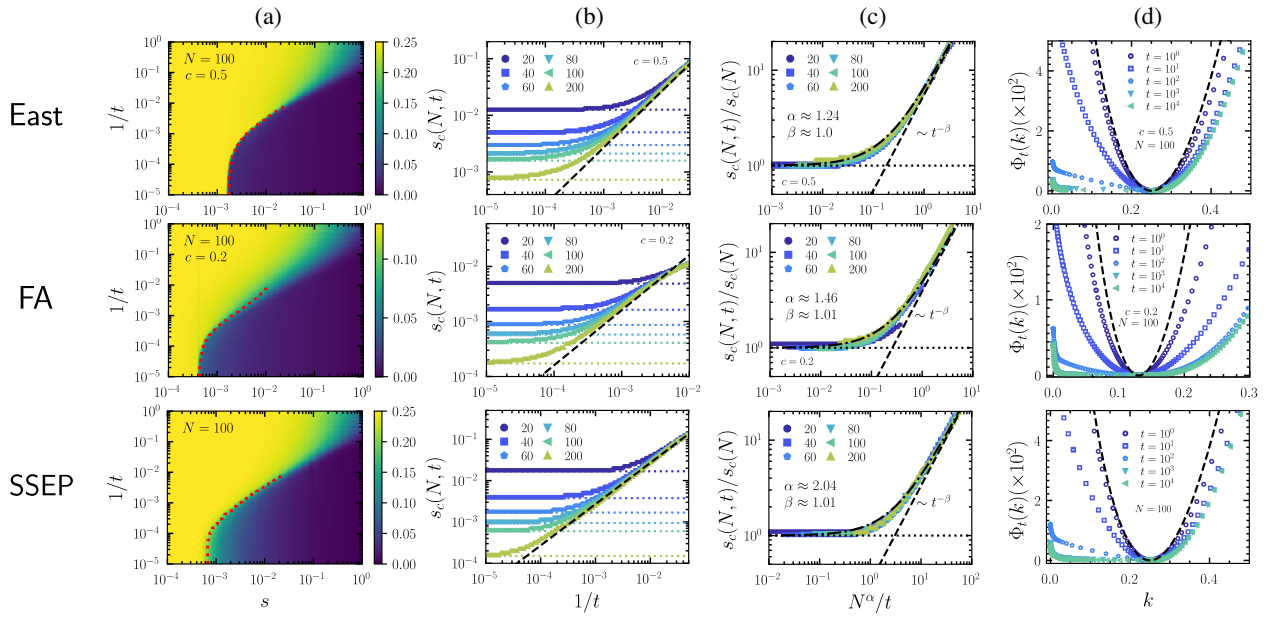


FIG. 2. Rare event statistics. The rare event statistics for the East model with (top) $c = 0.5$ and the FA model with (middle) $c = 0.2$ and (bottom) SSEP. (a) The dynamical activity $k(s,t)$ as a function of s and inverse time $1/t$ for $N = 100$. The red dotted line indicates our estimate of the transition point. (b) The transition point for various system sizes $N \in [10, 200]$. The dotted lines indicate the infinite time value (see Refs. [25,27]), and the dashed line shows $s_c(N,t) \sim t^{-1}$. (c) The same data are shown in (b) but with $s_c(N,t)$ scaled by $s_c(N)$ and time scaled by $N^{-\alpha}$, where α is the critical exponent extracted from $s_c(N)$. The dotted line shows where the y axis is one, and the dashed line shows $t^{-\beta}$. The sum of both lines is given by the dash-dotted line. (d) The estimate of the rate function $\Phi_t(k)$ defined in Eq. (13). The dashed line shows a Poisson distribution with the equilibrium average as its mean. All of the data are calculated using the dynamical partition sum $Z_t(s)$ from tMPS.

in time, for large sizes they still obey one in system size, $Z_t(s) \asymp e^{N\Theta(s,t)}$ and $P_t(K) \asymp e^{-N\Phi_t(K/N)}$. We can therefore obtain the time-dependent rate function $\Phi_t(K)$ through the Legendre transform

$$\Phi_t(k) = -\max_s[\Theta(s,t) + sk], \tag{13}$$

where $\Theta(s,t) = N^{-1} \log Z_t(s)$. From the numerical estimate of $Z_t(s)$ we can therefore estimate $\Phi_t(K/N)$ for all times. Figure 2(d) shows the corresponding rate functions for system size $N = 100$. For small times, the distribution of the activity is close to Poissonian (dashed line), in agreement with the absence of a transition. As time increases, the rate function broadens into the characteristic shape of a first-order phase transition [25,38].

Structure of the active phase.—Long time trajectories with an atypically large activity are known to display an interesting structure in two of the models we consider here [25,43,63]. Our finite time method allows us to study how such structure depends on the trajectory length.

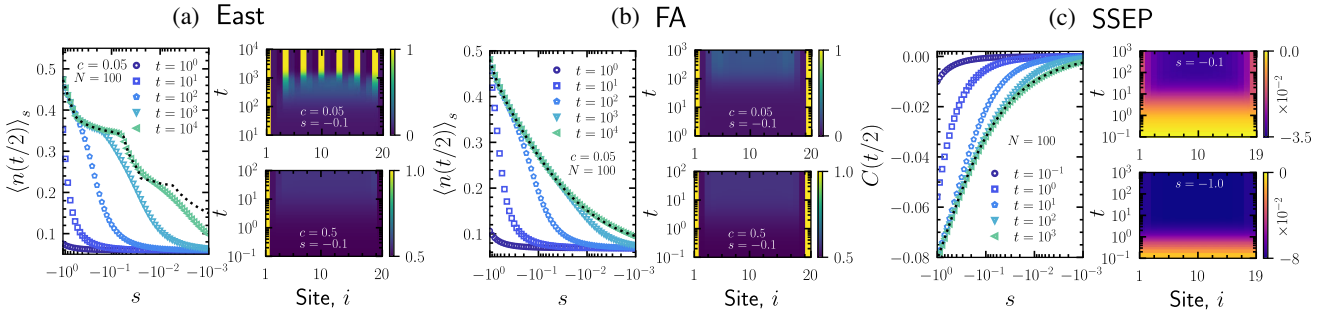


FIG. 3. Structures in the active phase. We show the average occupations at the center of the trajectory $\langle n_i(t/2)\rangle_s$ for $s = -0.1$ for the (a) East model and (b) FA models. The left panels of each show the lattice average for a range of s and t with $c = 0.05$, while the right panels show the occupations at each site for (top) $c = 0.05$ and (bottom) $c = 0.5$, with $s = -0.1$. We show the same for the SSEP in (c) but with the nearest-neighbor correlations $C_i(t/2)$. The right panels are for (top) $s = -0.1$ and (bottom) $s = -1.0$. Dotted lines show the expected value at infinite times. All observables are calculated from the time-evolved MPS $|\psi_{t/2}\rangle$.

In Fig. 3(a), we show the average occupation at the midpoint of the dynamics. The left panel shows the lattice-averaged occupations $\langle n(t/2)\rangle_s$ at time $\tau = t/2$ for ensembles of trajectories with total time t, as a function of s for various t, at $c = 0.05$. The panels on the right show the average spatial profile $\langle n_i(t/2)\rangle_s$ at $s = -0.1$ for (top) $c = 0.05$ and (bottom) $c = 0.5$. In both cases, the average density is spatially featureless at short times, but arranges to maximize activity at long times. For $c = 0.05$ it does so by forming an anticorrelated structure, while these anticorrelations are absent for $c = 0.5$ (cf. the long time case [63]). Figure 3(b) shows the same for the FA model, where there is no appreciable structure forming for small c. Notice also from the left panels the longer times needed to reach the LD behavior in the East compared to the FA model.

In Fig. 3(c) we quantify the local structure of the SSEP in terms of the nearest-neighbor correlations

$$C_i(t) = \langle n_i n_{i+1}(t)\rangle_s - \langle n_i(t)\rangle_s \langle n_{i+1}(t)\rangle_s, \qquad (14)$$

and the lattice average $C(t) = (N-1)^{-1}\sum_{i=1}^{N-1} C_i(t)$. The right panels show a growth of anticorrelated order with increasing trajectory length toward the "hyperuniform" arrangement at long times, cf. Ref. [43].

Conclusions.—We have implemented a time evolution scheme using MPS to study the rare events of one-dimensional KCMs in finite time trajectories. In this way we have extended recent efforts on the long time LD statistics via TNs to the arguably harder problem of the LDs away from the long time limit. We showed how to directly compute dynamical partition sums, and we derived an efficient sampling scheme for finite time rare trajectories. Understanding the finite time behavior of dynamical systems is significant, as the times required to observe long time behavior can be too large to implement experimentally. A next step would be to extend these ideas to dimensions larger than 1. A possibility could be to implement sampling through two-dimensional TNs, such as projected entangled-pair states (e.g., [64,65]), which have already proven useful in studying the LDs in the long time limit of two-dimensional exclusion processes [28]. While bond dimensions will be limited in this case, using a time evolution scheme like we presented here one could approximate the reference dynamics for the center of trajectories (i.e., evolve by $e^{t\mathbb{W}_s/2}$) alongside a scheme such as TPS to obtain reliable results. Another direction would be to apply the methods demonstrated here to driven problems, such as currents in exclusion processes. Here we cannot exploit Hermiticity and would have to compute the time-evolved left and right eigenvectors. We hope to report on such studies in the near future.

We acknowledge financial support from EPSRC Grant No. EP/R04421X/1 and the Leverhulme Trust Grant No. RPG-2018-181. M. C. B. acknowledges support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2111–390814868. We acknowledge access to the University of Nottingham Augusta HPC service.

[1] H. Touchette, Phys. Rep. **478**, 1 (2009).
[2] J. P. Garrahan, Physica (Amsterdam) **504A**, 130 (2018).
[3] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).
[4] D. T. Limmer, C. Y. Gao, and A. R. Poggioli, Eur. Phys. J. B **94**, 145 (2021).
[5] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[6] U. Ray, G. K.-L. Chan, and D. T. Limmer, J. Chem. Phys. **148**, 124120 (2018).
[7] K. Klymko, P. L. Geissler, J. P. Garrahan, and S. Whitelam, Phys. Rev. E **97**, 032123 (2018).
[8] A. Guyader and H. Touchette, J. Stat. Phys. **181**, 551 (2020).
[9] V. S. Borkar, S. Juneja, and A. A. Kherani, Commun. Inf. Syst. **3**, 259 (2003).
[10] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[11] V. Lecomte and J. Tailleur, J. Stat. Mech. (2007) P03004.
[12] R. L. Jack and P. Sollich, Prog. Theor. Phys. Suppl. **184**, 304 (2010).
[13] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Phys. Rev. E **93**, 062123 (2016).
[14] G. Ferré and H. Touchette, J. Stat. Phys. **172**, 1525 (2018).
[15] U. Ray, G. K.-L. Chan, and D. T. Limmer, Phys. Rev. Lett. **120**, 210602 (2018).
[16] D. Jacobson and S. Whitelam, Phys. Rev. E **100**, 052139 (2019).
[17] A. Das and D. T. Limmer, J. Chem. Phys. **151**, 244123 (2019).
[18] U. Ray and G. K.-L. Chan, J. Chem. Phys. **152**, 104107 (2020).
[19] T. H. E. Oakes, A. Moss, and J. P. Garrahan, Mach. Learn. **1**, 035004 (2020).
[20] D. C. Rose, J. F. Mair, and J. P. Garrahan, New J. Phys. **23**, 013013 (2021).
[21] C. Casert, T. Vieijra, S. Whitelam, and I. Tamblyn, Phys. Rev. Lett. **127**, 120602 (2021).
[22] S. Whitelam, D. Jacobson, and I. Tamblyn, J. Chem. Phys **153**, 44113 (2020).
[23] A. Das, D. C. Rose, J. P. Garrahan, and D. T. Limmer, J. Chem. Phys **155**, 134105 (2021).
[24] J. Yan, H. Touchette, and G. M. Rotskoff, Phys. Rev. E **105**, 024115 (2022).
[25] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).
[26] P. Helms, U. Ray, and G. K.-L. Chan, Phys. Rev. E **100**, 022101 (2019).
[27] L. Causer, I. Lesanovsky, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **102**, 052132 (2020).
[28] P. Helms and G. K.-L. Chan, Phys. Rev. Lett. **125**, 140601 (2020).
[29] T. Nemoto, E. Guevara Hidalgo, and V. Lecomte, Phys. Rev. E **95**, 012102 (2017).

[30] E. Guevara Hidalgo, T. Nemoto, and V. Lecomte, Phys. Rev. E **95**, 062134 (2017).

[31] Y. Baek, Y. Kafri, and V. Lecomte, J. Stat. Mech. (2019) 103202.

[32] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **103**, 062144 (2021).

[33] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).

[34] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).

[35] J. P. Garrahan, P. Sollich, and C. Toninelli, in *Dynamical Heterogeneities in Glasses, Colloids, and Granular Media*, edited by L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, and W. van Saarloos, International Series of Monographs on Physics (Oxford University Press, Oxford, UK, 2011).

[36] G. H. Fredrickson and H. C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984).

[37] J. Jäckle and S. Eisinger, Z. Phys. B **84**, 115 (1991).

[38] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).

[39] T. Bodineau and B. Derrida, C. R. Acad. Sci. **8**, 540 (2007).

[40] C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van Wijland, Phys. Rev. E **78**, 021122 (2008).

[41] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).

[42] T. Bodineau, V. Lecomte, and C. Toninelli, J. Stat. Phys. **147**, 1 (2012).

[43] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).

[44] D. Karevski and G. Schutz, Phys. Rev. Lett. **118**, 030601 (2017).

[45] This choice of boundaries has no spontaneous breaking of symmetry at $s \gg 0$.

[46] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).

[47] C. Maes, Phys. Rep. **850**, 1 (2020).

[48] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[49] U. Schollwck, Ann. Phys. (Amsterdam) **326**, 96 (2011).

[50] R. Orús, Nat. Rev. Phys. **1**, 538 (2019).

[51] Tensor network calculations were implemented using the ITensor library [52].

[52] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor Software Library for tensor network calculations, 10.48550/arXiv.2007.14822.

[53] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).

[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.128.090605 for a description of the numerical methods, which includes [55,56].

[55] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).

[56] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Ann. Phys. (Amsterdam) **411**, 167998 (2019).

[57] It is possible also to approximate dynamics at any point in the trajectory. For large times, the state evolved to time $\tau = t/2$ is, however, the most efficient.

[58] R. Chetrite and H. Touchette, Ann. Henri Poincaré **16**, 2005 (2014).

[59] T. Bodineau and C. Toninelli, Commun. Math. Phys. **311**, 357 (2012).

[60] T. Nemoto, R. L. Jack, and V. Lecomte, Phys. Rev. Lett. **118**, 115702 (2017).

[61] R. L. Jack, T. Nemoto, and V. Lecomte, J. Stat. Mech. (2020) 053204.

[62] M. Vojta, Rep. Prog. Phys. **66**, 2069 (2003).

[63] R. L. Jack and P. Sollich, J. Phys. A **47**, 015003 (2014).

[64] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.

[65] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys. **57**, 143 (2008).

Chapter 8

Optimal sampling of dynamical large deviations in two dimensions via tensor networks

The following work is from the arXiv pre-print (arXiv:2209.11788) *"Optimal sampling of dynamical large deviations in two dimensions via tensor networks"* by Luke Causer, Mari Carmen Bañuls and Juan P. Garrahan.

This final chapter generalizes the previous work (including that shown in this thesis) involving TNs to study stochastic dynamics to two-dimensions. Studying dynamical LDs of models in dimensions larger than one is difficult, with current methods limited to approaches from machine learning [115, 117] and TNs [45], which are expanded on here by applying PEPS to determine the dynamical LDs, as was done in 1D with MPS. It is also shown how the results of PEPS can be used to implement a trajectory sampling algorithm, generalizing the results of Chapter 5 to 2D.

The methods are applied to the 2D East model, and the 2D SSEP. Evidence is provided for first- and second- order dynamical phase transitions with respect to the activity for each model. These works help further our understanding of KCMs in higher dimensions, and introduce advanced methods for trajectory sampling in higher dimensions using TNs.

Further details of the numerical methods can be found in Appendix B. Following the submission of this thesis, a modified version of this preprint was accepted for publication in the journal "Physical Review Letters".

Corrections to the manuscript:

1. The Legendre transform defined in the "Dynamical LDs" section should read $\varphi(k) = -\max_s [\theta(s) + sk]$.

Optimal sampling of dynamical large deviations in two dimensions via tensor networks

Luke Causer,[1,2] Mari Carmen Bañuls,[3,4] and Juan P. Garrahan[1,2]

[1] *School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK*
[2] *Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham, NG7 2RD, UK*
[3] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*
[4] *Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, D-80799 München*

We use projected entangled-pair states (PEPS) to calculate the large deviations (LD) statistics of the dynamical activity of the two dimensional East model, and the two dimensional symmetric simple exclusion process (SSEP) with open boundaries, in lattices of up to 40×40 sites. We show that at long-times both models have phase transitions between active and inactive dynamical phases. For the 2D East model we find that this trajectory transition is of the first-order, while for the SSEP we find indications of a second order transition. We then show how the PEPS can be used to implement a trajectory sampling scheme capable of directly accessing rare trajectories. We also discuss how the methods described here can be extended to study rare events at finite times.

Introduction.- Over the last few years we have seen progress in the application of numerical tensor network (TN) techniques to compute statistical properties of dynamical trajectories in classical stochastic systems. The first such application was to long time statistics—the dynamical large deviation (LD) regime—of one-dimensional lattice systems using variational algorithms (such as density matrix renormalization group [1], or DMRG) to approximate the leading eigenvectors of tilted Markov generators by matrix product states (MPS, e.g. Ref. [2]) [3–10]. Building on these results, we introduced a sampling method which exploited such MPS to efficiently sample rare trajectories, and then presented a method based on MPS time-evolution to precisely compute trajectory statistics at finite times [11]. For more than one spatial dimension, a more suitable variational class is that of projected-entangled pair states (PEPS) [12], which fulfills an entanglement area law [13] and was recently applied to the classical asymmetric exclusion process in two dimensions in Ref. [14]. A computationally cheaper alternative, without an area law, but accommodating more entanglement than MPS, is that of tree tensor networks (TTN) [15], used for example in Refs. [16, 17] , in combination with a time-dependent variational principle [18] to study driven problems.

Here we use PEPS to study the large deviations of the *dynamical activity* in two paradigmatic two-dimensional models, the 2D East model (also known as North-or-East model) [19–22], and the 2D symmetric simple exclusion process (SSEP) with open boundaries where particles can be injected and removed [23]. We are able to accurately estimate the leading eigenvector of the tilted generator, and thus the LDs, of these models using the *simple update* (SU) algorithm for PEPS, see e.g. [24], and verify that further improvements can come from more complex update schemes, such as *full update* (FU), see e.g. [25, 26]. We then use the approximate leading eigenvector to construct an auxiliary dynamics which can directly sample the corresponding rare trajectories. Such an algorithm requires efficient sampling from the PEPS, and we show how to do this in the context of trajectory sampling. We benchmark our methods, showing how the bond dimension of the PEPS allows for a controlled accuracy of optimal dynamics. We demonstrate that both models have a phase transition between active and inactive dynamical phases, a first-order transition for the 2D East and a second-order transition for the 2D SSEP.

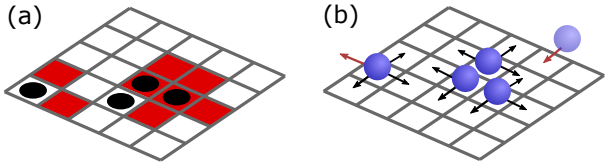


FIG. 1. **Models.** (a) The 2D East model. An occupation, shown by the black circles, can facilitate flips marked by the red-coloured cells at the neighbouring sites, but only in two directions. (b) The 2D SSEP. The sites on the lattice can be occupied by particles which can hop in any direction, as long as the target site is not occupied. Particles can enter or leave at the boundaries, as shown by the red arrows.

Models.- The models we study here live in a two-dimensional square lattice of size $N = L \times L$, with each site being occupied by a binary variable $n_{\boldsymbol{k}} = 0$ or 1, where $\boldsymbol{k} = (k_x, k_y)$ denotes the position of the site for $k_x, k_y = 1 \cdots L$. Their continuous-time dynamics is defined by a Markov generator (e.g. see Refs. [27, 28]),

$$\mathbb{W} = \sum_{x, y \neq x} w_{x \to y} \, |y\rangle \langle x| - \sum_x R_x \, |x\rangle \langle x| \,, \qquad (1)$$

where $|x\rangle$ and $|y\rangle$ are configurations on the lattice, $w_{x\to y}$ the transition rate from x to y, and $R_x = \sum_{y\neq x} w_{x\to y}$ the escape rate out of x. We can write this as $\mathbb{W} = \mathbb{K} - \mathbb{R}$, where $\mathbb{K}$ contains the off-diagonal transition rates, and $\mathbb{R}$ the diagonal escape rates.

The first model we consider is the 2D East model [19–22], often studied in the context of the glass transition.

This is a kinetically constrained model (KCM) such that an excited site $n_{\boldsymbol{k}} = 1$, allows ("facilitates") a site to its North or East to flip stochastically, see Fig. 1(a). The generator of the 2D East model reads

$$\mathbb{W}^{\text{East}} = \sum_{\boldsymbol{k}} P_{\boldsymbol{k}} \Big[c \left(\sigma^+_{\boldsymbol{k}} - (1 - n_{\boldsymbol{k}}) \right) + (1 - c) \left(\sigma^-_{\boldsymbol{k}} - n_{\boldsymbol{k}} \right) \Big], \quad (2)$$

where $c \in (0, 1/2]$ controls the average occupation density, and the kinetic constraint is $P_{(k_x,k_y)} = n_{(k_x-1,\,k_y)} + n_{(k_x,\,k_y-1)}$. In addition, we choose open boundary conditions (OBC) with $n_{(1,1)} = 1$ fixed. This ensures the entire state space remains dynamically connected [20].

The second model is the 2D SSEP. This describes particles hopping to neighbouring sites on a 2D lattice with unit rate, but only if the target site is not already occupied by a particle. We also allow particles to be injected or removed at the boundaries of the lattice with rate 1/2, see Fig. 1(b). The generator for the SSEP is

$$\mathbb{W}_{\text{SSEP}} = \sum_{\langle \boldsymbol{k}, \boldsymbol{l} \rangle} \Big[\sigma^+_{\boldsymbol{k}} \sigma^-_{l} - (1 - n_k) n_{\boldsymbol{l}} + \sigma^-_{\boldsymbol{k}} \sigma^+_{\boldsymbol{l}} - n_{\boldsymbol{k}} (1 - n_{\boldsymbol{l}}) \Big] + \frac{1}{2} \sum_{\boldsymbol{k} \in \partial} \Big[\sigma^x_{\boldsymbol{k}} - 1 \Big], \quad (3)$$

where $\langle \boldsymbol{k}, \boldsymbol{l} \rangle$ denotes a pair of nearest neighbours, and ∂ the boundary of the lattice.

Dynamical LDs.- We consider the statistics of some dynamical observable $\hat{K}$ through its probability distribution $P_t(K) = \sum_{\omega_t} \pi(\omega_t) \delta[\hat{K}(\omega_t) - K]$, where ω_t denotes a stochastic trajectory and $\pi(\omega_t)$ is the probability it occurs under the stochastic generator $\mathbb{W}$. Essentially the same information is encoded in the moment generating function (MGF), $Z_t(s) = \sum_{\omega_t} \pi(\omega_t) e^{-s\hat{K}(\omega_t)}$, where we have introduced the counting field s. In the $t \to \infty$ limit, the two obey LD principles $P_t(K) \asymp e^{-t\varphi(K/t)}$ and $Z_t(s) \asymp e^{t\theta(s)}$, with the rate function $\varphi(K/t)$ and scaled cumulant generating function (SCGF) $\theta(s)$ being time-independent. The LD functions are related through a Legendre transform, $\varphi(k) = -\min_s \left[\theta(s) + sk(s) \right]$, for $k = K/t$. For reviews, see Refs. [28–31].

A convenient way to determine the SCGF is to construct a biased or *tilted* generator [29, 32–34], a (non-stochastic) deformation of the Markov generator $\mathbb{W}$ such that the associated trajectories are exponentially biased by $e^{-s\hat{K}(\omega_t)}$. We consider as an observable the *dynamical activity* [32, 35], which counts the number of jumps in a stochastic trajectory and thus quantifies the overall level of motion. The corresponding tilted generator then takes the form $\mathbb{W}_s = e^{-s}\mathbb{K} - \mathbb{R}$. The SCGF can be retrieved by calculating the largest eigenvalue and eigenvector(s),

$$\mathbb{W}_s \left| r_s \right\rangle = \theta(s) \left| r_s \right\rangle, \quad \left\langle l_s \right| \mathbb{W}_s = \theta(s) \left\langle l_s \right|. \quad (4)$$

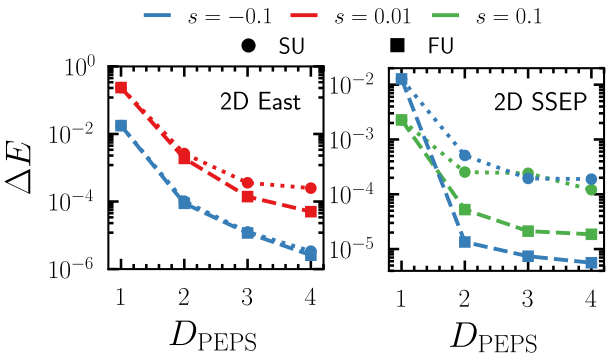


FIG. 2. **Optimization of PEPS.** The error in measured energy for the SU and the FU compared to the high accuracy 2D DMRG (with a MPS bond dimension up to $D_{\text{MPS}} = 1024$) for various values of s and a 10×10 lattice. The left panel shows the 2D East model with $c = 0.5$, and the right panel shows the 2D SSEP. The PEPS environment in the FU uses a boundary dimension $\chi_B = 4D^2$ for the East and $\chi_B = 6D^2$ for the SSEP.

For the models and dynamical observable considered here, we can introduce a similarity transformation independent of s, $\mathbb{H}_s = -\mathbb{P}^{-1}\mathbb{W}_s\mathbb{P}$, where $\mathbb{P}$ is a diagonal matrix with the square roots of the steady state probabilities [34]. This results in a Hermitian matrix with minimal eigenvalue and associated eigenvector

$$\mathbb{H}_s \left| \psi_s \right\rangle = -\theta(s) \left| \psi_s \right\rangle, \quad (5)$$

where $|\psi_s\rangle$ is related to the original eigenvectors by $|r_s\rangle = \mathbb{P}|\psi_s\rangle$ and $\langle l_s| = \langle \psi_s|\mathbb{P}^{-1}$. This representation is convenient as the minimum eigenvalue is bounded by the Rayleigh-Ritz principle.

PEPS.- Determining the minimal eigenvalue and eigenvector of $\mathbb{H}_s$ boils down to an optimization problem. We approach this using TN methods. A natural ansatz choice in this case is PEPS, the direct two-dimensional generalization of MPS [12]. PEPS are known to obey the area law in 2D [36], with the amount of entanglement controlled by its virtual bond dimension D_{PEPS}. To calculate observables we also need a scheme to contract the TN. For the case of PEPS, this cannot be done efficiently and thus we have to use an approximate scheme. We use the boundary MPS scheme [12, 37], where we contract from the edge of a PEPS network with an MPS with some boundary dimension χ_B which controls the accuracy of contraction. A common heuristic choice for local problems is $\chi_B \sim O(D^2_{\text{PEPS}})$ (see e.g. Ref. [38]).

The final step is to choose an update scheme to estimate the wavefunction $|\psi_s\rangle$. Broadly speaking, there are three popular approaches. The computationally cheapest but least precise is the SU scheme [24], which we use for the most part here. SU makes use of imaginary time evolution, with updates which only consider the local environment. It is not optimal but only entails compu-

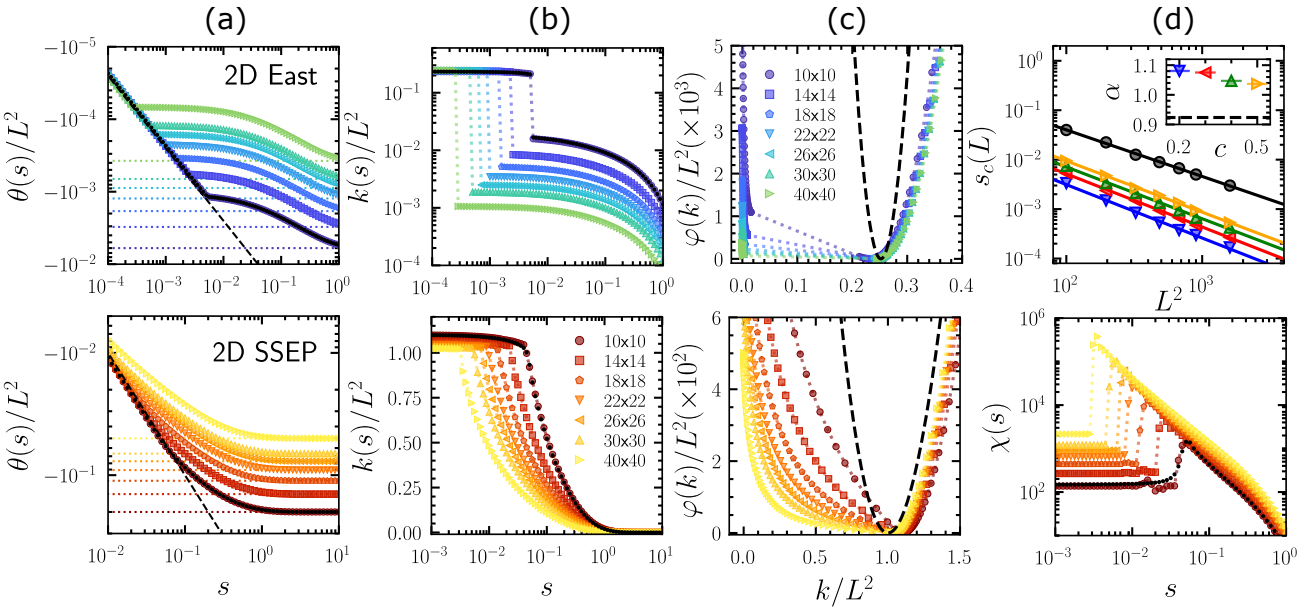


FIG. 3. **Dynamical large deviations and active-inactive transitions from PEPS.** (a) The SCGF $\theta(s)/L^2$ for the 2D East with $c = 0.3$ (top) and the SSEP (bottom) for system sizes $N \in [10^2, 40^2]$. The black dashed line shows the linear response for small s, and the colour dotted lines show the value for $s \to \infty$. (b) The dynamical activity $k(s)/L^2$ for the systems in (a). The East is on a log-log scale, and the SSEP a log-linear scale. (c) The rate function $\varphi(k)/L^2$ as a function of activity k/L^2 for the systems in (a). The dashed line shows the Poisson distribution with mean $k(s=0)/L^2$. (d) The transition points $s_c(L)$ for the 2D SSEP (black circles) and the 2D East for $c \in [0.2, 0.5]$. The solid lines show the fitted power-law curves $s_c(L) \sim L^{-2\alpha}$, with the exponents shown in the inset. The black dashed line is the exponent for the SSEP, and the symbols are for the East. The symbols can be used to read the value of c in the main figure. The bottom panel shows the dynamical susceptibility $\chi(s) = \theta''(s)$ for the 2D SSEP. All the data was acquired using the SU except for the black markers, which show 2D DMRG data for a $N = 10$ lattice for comparison.

tation cost of $O(D^5_{\text{PEPS}})$ [39]. A more efficient update which also relies on imaginary time evolution is the FU [25, 26]: here we have to contract the whole TN (except for the tensors which are being updated). While this costs $O(\chi_B^3 D^4_{\text{PEPS}} + \chi_B^2 D^6_{\text{PEPS}})$, with a good approximate environment it ensures the update is optimal. The final class of updates are Variational Updates (VU) [40, 41] which we do not consider here. For details see e.g. Ref. [42].

Figure 2 compares the SU and FU schemes [43] for both models against 2D DMRG [44] for small 10×10 lattices, where almost exact results can be determined with DMRG [45]. We show the relative difference in energy $\Delta E = (E_{\text{PEPS}} - E_{\text{DMRG}})/E_{\text{DMRG}}$. We find that the SU is able to achieve accuracy $\delta E < 10^{-3}$, which is enough for our study. Even though the FU could improve the results, we thus proceed with SU with a maximal bond dimension $D_{\text{PEPS}} = 4$, which allows us reaching large sizes at low computational cost.

Large deviations from PEPS.- The East and SSEP in 1D are known to have LD transitions in terms of the activity or other dynamical observables [6, 8, 32, 46–52]. In two-dimensions, the SSEP has a transition in the LDs of the current [14]. We now provide evidence by means of PEPS for both the 2D East and 2D SSEP having active-inactive phase transitions. Figure 3(a-c) shows the LD statistics for both the 2D East model (top) and the 2D SSEP (bottom). For the East model, we see from Fig. 3(a) that the SCGF follows linear response, $\theta(s) \approx sk(0)$, for small s, but at $s_c(L)$ it sharply changes to another branch. This point corresponds to a sudden drop in activity, $k(s) = -\theta'(s)$, which becomes discontinuous in the limit $N \to \infty$, see Fig. 3(b). Having access to both the SCGF and the dynamical activity allows us to estimate the rate function $\varphi(k)$, shown in Fig. 3(c). We see broadening of the rate function around the mean, indicating the coexistence of active and inactive dynamics. All this behaviour is characteristic of a first-order phase transition.

For the SSEP we see something different: Fig. 3(a) shows no sharp change in $\theta(s)$, and the activity in Fig. 3(b) has no discontinuity. This is indicative of a second-order transition, with the rate function showing critical broadening, see Fig. 3(c), and a divergence in the susceptibility $\chi(s) = \theta''(s)$, see Fig. 3(d). Note that this is different from the 1D SSEP with open boundaries [8] in which this transition is first-order.

For both models we can extract a transition point from

the drop in either first or second cumulant. The top panel of Fig. 3(d) shows how the transition point scales with L for both models (for a range of c for the 2D East). We are able to fit the data with the power laws $s_c(L) \sim L^{-2\alpha}$, as shown by the solid lines. We find the exponents $\alpha \gtrsim 1$ for the 2D East and $\alpha \lesssim 1$ for the SSEP, see inset to the top panel of Fig. 3(d).

Optimal sampling of rare trajectories from PEPS.- Sampling trajectories corresponding to the $s \neq 0$ phases is difficult as they are exponentially rare in system size and time. The optimal sampling dynamics at long times is given by the so-called *generalized Doob transform* [53–57], which maps the tilted generator into a true stochastic generator for the rare trajectories, $\mathbb{W}_s^{\mathrm{Doob}} = \mathbb{L}\,[\mathbb{W}_s - \theta(s)\mathbb{I}]\,\mathbb{L}^{-1}$, where $\mathbb{L}$ is the leading left eigenvector of $\mathbb{W}_s$ as a diagonal matrix. This gives a new dynamics with the transition rates

$$\tilde{w}_{x\to y} = \frac{l_s(y)}{l_s(x)} e^{-s} w_{x\to y}, \tag{6}$$

with $l_s(x) = \langle l_s|x\rangle$. In $\mathbb{W}_s^{\mathrm{Doob}}$ the counting field s appears as a physical control parameter, and running dynamics with rates (6) gives trajectories at $s \neq 0$ on demand. While optimal, $\mathbb{W}_s^{\mathrm{Doob}}$ is difficult to construct in general as one needs the exact left leading eigenvector. However, we can exploit our PEPS approximation to estimate the rates Eq. (6), similar to Ref. [58] for 1D and MPS.

To obtain Eq. (6) for the transitions out of a state x we calculate $l_s(y)$ from the PEPS using a boundary dimension $\chi_B = D_{\mathrm{PEPS}}$ [41, 59–61], thus entailing a maximum cost $O(ND_{\mathrm{PEPS}}^6)$. If we neglect the time edges of trajectories, we can estimate an time-extensive observable by importance sampling

$$\langle O\rangle_s \approx \frac{\sum_{\alpha_t} O(\alpha_t) g(\alpha_t)}{\sum_{\alpha_t} g(\alpha_t)}, \tag{7}$$

where α_t denotes a trajectory generated with (6) (the *reference dynamics*), and $O(\alpha_t)$ is the trajectory observable. The re-weighting factor $g(\alpha_t)$ is

$$g(\omega_t) = e^{-\int_0^t dt' R(t') - \tilde{R}(t')}, \tag{8}$$

where $R(t')$ and $\tilde{R}(t')$ are the escape rates of the system at time t' in the original dynamics and the approximate Doob dynamics, respectively. Notice that with a large enough number of trajectories, Eq. (7) can be used to correct on the imperfections in the reference dynamics due to an imperfect PEPS approximation.

Figures 4 show results from our sampling algorithm for the 2D East with $c = 0.5$ and the 2D SSEP, both for system sizes $N = 22 \times 22$. The average dynamical activity measured in trajectories (symbols) [with umbrella sampling (7,8)] coincides with that obtained directly from the PEPS (solid line), except for $D_{\mathrm{PEPS}} = 1$ for the East model. The accuracy of our dynamics is quantified by the

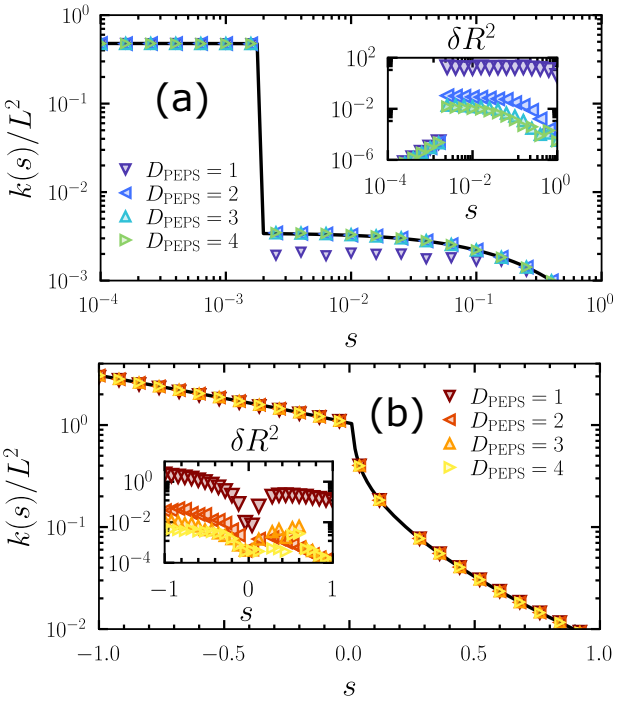


FIG. 4. **Optimal sampling of trajectories.** The average dynamical activity from CTMC with importance sampling (symbols) for the (a) 2D East model with $c = 0.5$ and (b) the 2D SSEP respectively on a 22×22 lattice. We show data over a range of s, and $D_{\mathrm{PEPS}} \in [1, 4]$. The trajectory times are chosen such that on average we expect 100 transitions per trajectory. The solid black line shows the activity measured directly from the PEPS with $D = 4$ for comparison. The insets show the variance in the time-integrated difference of escape rates, δR^2 (see main text). For the East model, we show results over the dynamical phase transition $s > 0$, while the SSEP shows results for over negative and positive s. Each data point is calculated from $N_{\mathrm{sp}} \in [10^3, 10^4]$ trajectories. For visualisations of representative trajectories see Ref. [62].

variance of the time integrated difference in escape rates, cf. Eq. (8), which vanishes for the exact Doob rates. We show this for each D in the insets of Figs. 4: increasing the D_{PEPS} consistently reduces the variance, indicating a better sampling dynamics and less need for importance sampling.

Conclusions.- We have shown that the dynamical large deviations of two-dimensional stochastic models can be studied efficiently with PEPS, including the quasi-optimal sampling of rare trajectories. We showed here that both the 2D East model and the 2D SSEP have active-inactive trajectory transitions, of the first-order and second-order, respectively, the latter in contrast to the case of the 1D SSEP. Our work adds to the continuously expanding application [3–11, 14, 16, 17, 63] of tensor network methods to study the dynamical fluctuations in classical stochastic systems.

There are several interesting avenues to pursue building on this work. One is to integrate 2D trajectory sampling via tensor networks with a method such as transition path sampling (TPS) [64] for investigating statistics of fluctuations at *finite times*, cf. [11, 65]. While the current implementations with PEPS are too demanding to reasonably incorporate TPS, tree tensor networks (TTNs) [15] are a promising alternative that could allow to reliably investigate finite time scaling. We hope to report on this is the near future.

Acknowledgements.- We acknowledge financial support from EPSRC Grant no. EP/R04421X/1 and the Leverhulme Trust Grant No. RPG-2018-181. M.C.B. acknowledges support from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2111 – 390814868. Calculations were performed using the Sulis Tier 2 HPC platform hosted by the Scientific Computing Research Technology Platform at the University of Warwick. Sulis is funded by EPSRC Grant EP/T022108/1 and the HPC Midlands+ consortium. We acknowledge access to the University of Nottingham Augusta HPC service. Example code used to generate the data can be found at `https://github.com/lcauser/2d-optimal-sampling`. Data is available at from Ref. [62].

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[2] U. Schollwöck, Ann. Phys **326**, 96 (2011).
[3] M. Gorissen, J. Hooyberghs, and C. Vanderzande, Phys. Rev. E **79**, 020101 (2009).
[4] M. Gorissen and C. Vanderzande, Phys. Rev. E **86**, 051114 (2012).
[5] M. Gorissen, A. Lazarescu, K. Mallick, and C. Vanderzande, Phys. Rev. Lett. **109**, 170601 (2012).
[6] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).
[7] P. Helms, U. Ray, and G. K.-L. Chan, Phys. Rev. E **100**, 022101 (2019).
[8] L. Causer, I. Lesanovsky, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **102**, 052132 (2020).
[9] L. Causer, J. P. Garrahan, and A. Lamacraft, Phys. Rev. E **106**, 014128 (2022).
[10] J. Gu and F. Zhang, (2022), arXiv:2206.05322.
[11] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. Lett. **128**, 090605 (2022).
[12] F. Verstraete and J. I. Cirac, (2004), arXiv:cond-mat/0407066.
[13] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[14] P. Helms and G. K.-L. Chan, Phys. Rev. Lett. **125**, 140601 (2020).
[15] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).
[16] N. E. Strand, H. Vroylandt, and T. R. Gingrich, J. Chem. Phys **156**, 221103 (2022).
[17] N. E. Strand, H. Vroylandt, and T. R. Gingrich, (2022), arXiv:2201.04107.
[18] D. Bauernfeind and M. Aichhorn, SciPost Phys. **8**, 024 (2020).
[19] J. Jäckle and S. Eisinger, Z. Phys, B **84**, 115 (1991).
[20] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[21] L. Berthier and J. Garrahan, J. Phys. Chem. B **109**, 3578 (2005).
[22] C. Casert, T. Vieijra, S. Whitelam, and I. Tamblyn, Phys. Rev. Lett. **127**, 120602 (2021).
[23] K. Mallick, Physica A **418**, 17 (2015).
[24] H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **101**, 090603 (2008).
[25] P. Corboz, R. Orús, B. Bauer, and G. Vidal, Phys. Rev. B **81**, 165104 (2010).
[26] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orús, Phys. Rev. B **92**, 035142 (2015).
[27] C. Gardiner, *Handbook of stochastic methods* (Berlin: Springer, 2004).
[28] J. P. Garrahan, Physica A **504**, 130 (2018).
[29] H. Touchette, Phys. Rep. **478**, 1 (2009).
[30] H. Touchette, Physica A **504**, 5 (2018).
[31] R. L. Jack, Eur. Phys. J. B **93**, 74 (2020).
[32] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[33] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).
[34] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, J. Phys. A **42**, 075007 (2009).
[35] C. Maes, Phys. Rep. **850**, 1 (2020).
[36] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).
[37] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys **57**, 143 (2008).
[38] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, New J. Phys. **16**, 033014 (2014).
[39] This is the case when we use the so-called "reduced tensor" scheme, but otherwise scales as $O(D^6_{\text{PEPS}})$.
[40] K. Hyatt and E. M. Stoudenmire, (2019), 1908.08833.
[41] T. Vieijra, J. Haegeman, F. Verstraete, and L. Vanderstraeten, Phys. Rev. B **104**, 235141 (2021).
[42] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, Phys. Rev. B **90**, 064425 (2014).
[43] Both are used with a decreasing update time-step $\delta t \in [10^{-1}, 10^{-3}]$, and the SU uses a minimum truncation cut-off error 10^{-12}. The SU is used to serve an initial guess for the FU.
[44] E. Stoudenmire and S. R. White, Annu. Rev. Condens. Matter Phys **3**, 111 (2012).
[45] We "snake" the MPS across the 2D lattice diagonally, running the DMRG algorithm with a fixed bond dimension. The algorithm is allowed enough sweeps until complete convergence in energy. The bond dimension is then routinely doubled, and the DMRG algorithm is run again. After each convergence, we record the energy and local observables. We check for convergence in the MPS bond dimension. If convergence is met (differences in energy and observables are smaller than 10^{-5}), we terminate the algorithm. We find a maximal bond dimension of $D_{\text{MPS}} = 1024$ is sufficient to meet these criteria.
[46] B. Derrida, J. Stat. Mech. **2007**, P07023 (2007).
[47] T. Bodineau and B. Derrida, C. R. Acad. Sci. **8**, 540 (2007).
[48] C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van

Wijland, Phys. Rev. E **78**, 021122 (2008).
[49] V. Lecomte, J. P. Garrahan, and F. van Wijland, J. Phys. A **45**, 175001 (2012).
[50] R. L. Jack, I. R. Thompson, and P. Sollich, Phys. Rev. Lett. **114**, 060601 (2015).
[51] T. Nemoto, R. L. Jack, and V. Lecomte, Phys. Rev. Lett. **118**, 115702 (2017).
[52] D. Karevski and G. M. Schütz, Phys. Rev. Lett. **118**, 30601 (2017).
[53] V. S. Borkar, S. Juneja, and A. A. Kherani, Commun. Inf. Syst. **3**, 259 (2003).
[54] D. Simon, J. Stat. Mech.: Theory Exp **2009**, P07017 (2009).
[55] R. L. Jack and P. Sollich, Prog. Theor. Phys. Supp. **184**, 304 (2010).
[56] R. Chetrite and H. Touchette, Ann. Henri Poincaré **16**, 2005 (2015).
[57] J. P. Garrahan, J. Stat. Mech.: Theory Exp **2016**, 073208 (2016).
[58] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **103**, 062144 (2021).
[59] K. Ueda, R. Otani, Y. Nishio, A. Gendiar, and T. Nishino, Journal of the Physical Society of Japan **74**, 111 (2005), https://doi.org/10.1143/JPSJS.74S.111.
[60] M. Frías-Pérez, M. Mariën, D. Pérez-García, M. C. Bañuls, and S. Iblisdir, (2021), arXiv:2104.13264.
[61] M. M. Rams, M. Mohseni, D. Eppens, K. Jałowiecki, and B. Gardas, Phys. Rev. E **104**, 025308 (2021).
[62] L. Causer, M. C. Bañuls, and J. P. Garrahan, "Data for 'optimal sampling of dynamical large deviations in two dimensions via tensor networks'," (2022).
[63] J. P. Garrahan and F. Pollmann, (2022), arXiv:2203.08200.
[64] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Ann. Rev. Phys. Chem **53**, 291 (2002).
[65] Y. Tang, J. Liu, J. Zhang, and P. Zhang, (2022), arXiv:2208.08266.

Chapter 9

Conclusions & outlook

This thesis has furthered the development of the application of TNs to stochastic dynamics, with a strong focus on dynamical LDs. It began by introducing two new classes of KCMs. The first was the XORFA model, which has experimental relevance through its connection to Rydberg atoms in their "anti-blockade" regime. The slow relaxation properties of the model can be attributed to the fact its dynamical degrees of freedom are limited to the diffusion of its DWs, allowing a SEP description of the model. The second model was the stochastic generalization of the Fredkin model. This model behaves like an ASEP with additional kinetic constraints on the neighbouring sites, and can be considered a generalization of the East model to a SEP. For each model, evidence was provided of first-order dynamical phase transitions, highlighting the predominance of dynamical heterogeneity in the stationary dynamics. These works have introduced further examples of stochastic dynamics with dynamical phase coexistence, with links to well studied models with quantum dynamics. Furthermore, they demonstrate how TN methods can be used within stochastic dynamics, with local and global dynamical constraints which must be adhered to. An important discovery not yet fully explored is that the inactive phase of each of the models exhibits localization. This could motivate the future study of their tilted stochastic generators under closed unitary evolution, as was done for the "quantum East model" [146]. This exponential localization, along with the strong dynamical constraints, are known mechanisms for anomalous non-thermal eigenstates: eigenstates with properties which are inconsistent with the canonical ensemble at the energy of the eigenstate [146, 203]. These often include eigenstates with low amounts of entanglement throughout the energy spectrum, resembling those of "quantum scars" [204–207].

The remainder of the thesis was devoted to the algorithmic developments of the application of TNs to dynamical LDs. It was first shown how the variational MPS results can be used to construct an auxiliary dynamics capable of optimally and almost exactly sampling the Doob dynamics. This is useful for a few reasons. One application was to use the approximate variational MPS results with trajectory path sampling methods to correct on any errors. This was used to incorporate efficient path sampling with limited computational resources. However, it could also be used for systems where MPS are not able to estimate the leading eigenvectors to a high precision. On the other hand, the trajectories provide much more information than the time-average MPS, including that of time-correlations. This approach was used to study finite-time trajectory ensembles. Previously, finite-times were mostly limited to small times, due to the exponential cost of rare trajectory sampling in time. Indeed, very long times are made possible by the framework of LDs. However, the intermediate times where transitions between fast and

slow dynamics occur were previously inaccessible; investigating these times helps to paint a complete picture of the full phase diagram.

The study of finite time ensembles then inspired the application of MPS time-evolution methods to better study finite time LDs. Through a hybrid of variational and time-evolution methods, it was possible to precisely determine the time-evolution of probability vectors under the stochastic evolution of tilted Markov generators, and in turn determine the finite time dynamical fluctuations. This approach was superior to the previous, as the time-evolution approach considers the evolution of the entire ensemble of trajectories (whereas the previous was a sampling based approach). However, the results of this method could be easily integrated into the sampling approach to more effectively perform trajectory sampling at finite times. This allows for the benefits of the time-evolution approach, and also to use sampling to correct on systematic errors introduced through the approximate time-evolution MPS methods. The new methods to investigate finite time ensembles enabled a more detailed characterization of the first-order dynamical phase transitions for various lattice models, namely the East, FA and SSEP.

The final part of this work addressed the generalization of the MPS methods to 2D systems. This problem was tackled through numerical PEPSs, the natural 2D generalization of MPSs. The dynamical LDs for the East and SSEP models were estimated using well studied optimization methods of PEPSs, revealing first- and second- order dynamical phase transitions in activity for each respectively. As was done in 1D, the PEPSs results were incorporated with trajectory sampling methods, generalizing the methods from MPSs. While the trajectory sampling approach was efficient at sampling the dynamical large deviations at large times, it was not powerful enough to deal with finite times. This problem was recently approached using recurrent neural networks for small system sizes [208]. An interesting avenue for exploration would be to study how PEPS, or other TN architectures, such as TTNs, compare to these methods.

The models and methods considered here have furthered the understanding of dynamical phase transitions in KCMs. In particular, they allowed detailed finite-size scaling analyses of numerous KCMs in one- and two- spatial dimensions, enabling the determination of the order of their respective transitions, and their scaling exponents. Many theoretical works have demonstrated that prototypical spin facilitated models of glass formers, such as the East and FA model, have first-order phase transitions in their activity in 1D [52, 145]. Ref. [43] made use of the MPSs to confirm this to a high accuracy, finding the point of transition from active-to-inactive dynamics scales as $s_c(N) \sim N^{-\alpha}$ with $\alpha \gtrsim 1$. This was also argued to be the case for the FA model in 2D [117], and for the North-or-East model in Chapter 8 of this thesis. On the contrary, the results for the XORFA model in Chapter 4 demonstrated a first-order transition but with a scaling exponent $\alpha \gtrsim 2$. This difference in this exponent might be explained through a geometric picture (e.g. see Ref. [209]). While the first-order transition is indicative of space-time bubbles in the East, FA and the XORFA, the difference is in the way the boundaries of the bubbles evolve in each model. For the FA (and East) model, the boundaries grow linearly in time; that is, for example, the bubbles of slow domains are penetrated by fast domains with linear slopes [209]. Conversely, the dynamics of the XORFA is attributed to the diffusion of DWs, much like in the SSEPs (but with particles). DWs from the active regions only diffusively penetrate inactive regions which are *empty* of DWs (and similarly inactive regions which are rich in DWs penetrate active regions in the same way). Thus, one might expect the slow domains to persist for a longer time. This perspective is supported by the larger exponent: the position of the transition point is indicative of how

"slow" the dynamics is. The closer the transition point is to $s = 0$, the slower one can expect the dynamics to be. [1] Hence, a larger exponent indicates a slower penetration of domains.

Similarly, the Fredkin model with $c = 0.5$ also has a first-order transition, with an exponent of $\alpha \sim 2.5$. While it is not entirely clear if this would hold in the large size limit, it was shown to be the case for the range of system sizes available by means of MPS. Indeed, this larger exponent might not be too surprising. The model behaves diffusively like the SSEP, and as such, one might expect the exponent $\alpha \sim 2$. However, here are additional kinetic constraints, making the dynamics even slower, and thus yielding a larger exponent. Nevertheless, it is still an open question to explain the exact nature of the exponent. For $c < 0.5$, it was found the systems behave similarly to the East and FA models, with exponents $\alpha \gtrsim 1$. One plausible explanation for this is that inactive regions penetrate the neighbouring active regions linearly due to the asymmetry of the hopping. Furthermore, the stationary state has much smaller fluctuations for $c < 0.5$ due to the choice of boundary conditions. It would be interesting to see how (if at all) the dynamical fluctuations are affected for a choice of boundaries which have greater amounts of stationary fluctuations.

Interestingly, the work presented in Chapter 8 showed that the SSEP in 2D had a second-order phase transition in the activity, despite the fact the SSEP in 1D appears to be first-order. While this still indicates large amounts of dynamical fluctuations, the physical intuition is not as clear. The continuous nature of the transition suggests that there are algebraically decaying correlations in the space-time dynamics of the model (in analogy to criticality in a static classical 2D system). It might be instructive to investigate the two-point correlation functions of the dynamics in space-time (auto-correlation functions over a region of space) to better understand the significance of the transition. The difference in the nature of the transition is likely a consequence of the dimensionality of the problem. Generally speaking, consider a cluster of jammed particles. Particles at the middle of the cluster must first wait for particles at the edge to disperse (which happens diffusively) before they can become dynamically active. In 2D, there are more directions for this to happen, and thus one might expect the dispersion to be quicker than the 1D case. One possible route to investigate this premise could be to consider the SSEP on various 2D geometries which allows for a range of coordination numbers. The exponent, $\alpha \lesssim 1$, is also smaller than that found in the 1D SSEP. Consider the inactive phase in 2D (which is devoid of particles), with the boundary conditions in Chapter 8. Particles can enter at the boundaries of the system, and then diffuse throughout. To relax to equilibrium, the particles only need to diffuse a distance $l \sim L$, i.e. the linear dimension of the lattice. As such, it could be reasonable to suggest the scaling $s_c(N) \sim L^{-2} = N^{-1}$ (where the exponent -2 comes from the fact the particles are diffusive).

Additionally, the methods introduced here allow for the determination of the scaling of the transition point with respect to time, $s_c(t) \sim t^{-\beta}$. For each of the 1D models investigated in Chapter 7, it was shown that the exponent goes as $\beta \sim 1$. This is surprising given the differences in the value of α for each model. More work is required to investigate if this exponent is universal to all models with a dynamical phase transition in activity (and extensions to other standard non-equilibrium observables, such as currents).

[1] This argument can be made more concrete by considering the position of the Lee-Yang zeroes of the biased partition sum in the complex space, e.g. Ref. [160]. The cumulants of the dynamics can be determined exactly by the Lee-Yang zeroes, with the zeros closest to $s = 0$ giving the dominant contributions.

The methods developed here enabled the investigation of dynamical fluctuations of stochastic lattice models. The focus was on the dynamical activity due to its significance in glass forming systems and because of the numerical simplifications, namely that the dynamics under the biasing of dynamical activity will still respected detailed balance (assuming the original system obeyed detailed balance). However, it is possible to adapt the methods presented here to non-equilibrium (driven) systems, e.g. see Refs. [44, 45]. One future application of this to KCMs could be to consider initial probability distributions which are out-of-equilibrium, such as quenches in temperature. Previous works have shown that the s-ensemble can be used to probe dynamics of intermediate timescales [202]. Chapter 7 also demonstrated this, and also gave indications to the trajectory times required to probe these dynamics using the s-ensemble. It would be interesting to see if TNs could provide a more straightforward way to study the metastable dynamics at times which are inaccessible, e.g. not knowing the expected behaviour of the dynamics at some time a priori.

It is important to note that these approaches have been restricted to models with a discrete number of (spatial) degrees of freedom. An important future direction would be to adapt these approaches for a broader class of systems, such as those which are continuous in space (e.g. simulations of liquids). While one should not expect the same level of accuracy as is seen for lattice models, TNs might still provide benefits to trajectory sampling through approaches similar to Chapters 6 and 8. Furthermore, the focus of this thesis was on the *canonical* ensemble – the true dynamical fluctuations are described by the PDF of a given dynamical observable. While the two are related through a Laplace transform, they are in practice difficult to disentangle. One possible solution is to sample the microcanonical ensemble through TNs. In particular, MPS can represent a probability distribution over configurations. By performing a Dyson series expansion, we can represent an entire ensemble of trajectories with *fixed transition times* by a 2D TN. This approximately considers all spatial possibilities, but restricts the temporal degrees of freedom. To account for time, one could design a Monte Carlo algorithm similar to TPS, which adjusts the transition times, allowing us to sample the full microcanonical ensemble. Alternatively, one could attempt to target the full ensemble by designing an approximate ensemble. This could be of the form of a Gaussian ensemble, whereby the probabilities of trajectories with values different to the target value decay as a Gaussian distribution. Formulating the problem in this way could allow the application of methods used in closed quantum systems to target the microcanonical properties of quantum Hamiltonians [5].

Another possibility for future investigation is the study of dynamical LDs in open quantum systems. Many efforts have gone towards generalizing the framework of LDs to open quantum systems, including: (i) finding the "tilted Lindbladians" whose leading eigenvalues correspond to the LDs [55], (ii) the formulation of a "quantum Doob dynamics" [55, 210–212] which exploits the leading eigenmatrix of the tilted Lindbladian to construct a dynamics whose statistics are the same as the those of interest, and (iii) constructing an auxiliary dynamics (Doob dynamics) capable of sampling the *true* rare quantum trajectories [134, 211]. These methods have been applied to various systems of interest for small system sizes to uncover dynamical phase transitions [213–216]. Indeed, one could try to apply MPS methods for larger system sizes, as has done for classical stochastic systems in this thesis. However, there are many difficulties in doing so. The tilted generator is non-Hermitian, and thus are not bound by a Raleigh-Ritz variational principle. This could hinder the performance of variational algorithms [217, 218]. One

could try to use time-evolution methods [219–223], however, the intermediate states before reaching the long-time limit could have too much entanglement to be faithfully captured by an MPS, and thus the truncation of the MPS could destroy the positivity requirements of the density matrix. Future works would require the development of methods which enforce positivity [224], but are also expressive enough to well describe the true density matrix.

Appendix A

Supplemental material for "Finite time large deviations via matrix product states"

The following document is the "Supplemental material" for the publication *"Finite time large deviations via matrix product states"* by Luke Causer, Mari Carmen Bañuls and Juan P. Garrahan, in Physical Review Letters 128 (9), 090605. It includes some extra details on how to perform time-evolution using MPS in the context of rare-event sampling, and quantifies the errors.

Supplemental Material: Finite time large deviations via matrix product states

Luke Causer,[1,2] Mari Carmen Bañuls,[3,4] and Juan P. Garrahan[1,2]

[1] *School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK*
[2] *Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham, NG7 2RD, UK*
[3] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*
[4] *Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, D-80799 München*

This document provides details about the numerical methods and their associated errors.

METHODS

The vector $|\psi_t\rangle$ is given as a matrix product state (MPS) ansatz,

$$|\Psi\rangle = \sum_{i_1 \cdots i_N}^{d} \mathrm{Tr}(A_1^{i_1} A_2^{i_2} \cdots A_N^{i_N})\, |i_1\, i_2\, \cdots\, i_N\rangle \tag{S1}$$

where each A_j^k is a rank-3 tensor with dimensions (d, D, D), with the variational parameter D (known as the *bond dimension*) and $k = 1 \ldots d$. We implement time evolution using a hybrid approach. We start with the equilibrium steady state, $|\psi_0\rangle$, which can be written as a product state (an MPS with $D = 1$),

$$|\psi_0^{\mathrm{FA}}\rangle = |1\rangle \otimes \left[\sqrt{1-c}\,|0\rangle + \sqrt{c}\,|1\rangle\right]^{\otimes N-2} \otimes |1\rangle\,, \tag{S2}$$

$$|\psi_0^{\mathrm{East}}\rangle = |1\rangle \otimes \left[\sqrt{1-c}\,|0\rangle + \sqrt{c}\,|1\rangle\right]^{\otimes N-1}. \tag{S3}$$

The first step is to find the leading eigenvector $|\psi^{\mathrm{LD}}\rangle$ of $\mathbb{H}_s$. This can be achieved by employing variational MPS (vMPS, see e.g. Ref. [1] for details). We then project the initial state $|\psi_0\rangle$ onto the (unnormalized) LD vector and its orthogonal complement,

$$|\psi_0^{\mathrm{LD}}\rangle = \mathcal{P}\,|\psi_0\rangle\,, \tag{S4}$$

$$|\psi_0^{\mathrm{rem}}\rangle = (1-\mathcal{P})\,|\psi_0\rangle\,, \tag{S5}$$

where $\mathcal{P} = |\psi^{\mathrm{LD}}\rangle\langle\psi^{\mathrm{LD}}|$. The two states are then evolved separately. The first is an eigenstate of the evolution operator (up to some small error given by the variance from vMPS), $|\psi_t^{\mathrm{LD}}\rangle = e^{t\theta(s)}\,|\psi_0^{\mathrm{LD}}\rangle$. The remaining state is a mixture of all other eigenstates in the spectrum and cannot be easily evolved in the same way. Fortunately, a range of techniques have been developed for MPS to allow for time-evolution (see Ref. [2] for comparisons). We will focus on the method introduced in [3], with details below.

The time-evolution of a MPS (tMPS) can be achieved by sequentially applying the time evolution operator $U(\delta) = e^{\delta \mathbb{H}_s}$ to the MPS. This is approximated using (second order) Trotter-Suzuki decomposition [4] with small times $\delta \ll 1$. We find a Trotter step of $\delta \in [0.01, 0.1]$ to be sufficient. To avoid exponential growth of the bond dimension, we apply a full truncation scheme through singular value decomposition and variational sweeps to minimize the distance between the truncated and un-truncated MPS, keeping only a maximum of $D = 400$ states (although in practice we never reach this) and with a truncation error $\epsilon = 10^{-12}$ [5]. Note that when we perform the variational truncation, we must also project out the leading eigenvector again, as it may be re-introduced through truncation. The full time-evolved state is then brought back together by summing the two separate states. The method is outlined in Fig. S1.

We now offer a few comments to the effectiveness of this method. Firstly, the time evolution is in *real time* (by this convention, quantum evolution is imaginary time). In the large time limit, the leading eigenvector will dominate as the other vectors will be exponentially dampened. The methods here allow us to accurately unravel the contribution of the leading eigenvector up until the times it becomes dominant. Furthermore, this approach often allows us to simulate large times by only having small simulation times in tMPS. That is, the remaining state can quickly converge onto the second leading eigenvector (which we determine through the change in norm), allowing us to stop the simulation early and extrapolate to large times. This is particularly useful when determining the transition from active-to-inactive dynamics for small $s > 0$, where timescales diverge but all but the leading two eigenvectors are exponentially dampened. Notice that one can adapt this method to allow for multiple leading eigenvectors of the generator, which can be determined variationally. Indeed this could provide a more precise determination of the partition sums, but we find this not to be necessary here.

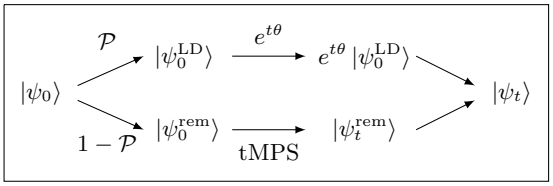


FIG. S1. A schematic drawing of the time evolution methods used here. We project the initial state onto the LD vector and the remainder. The former can be evolved exactly as itself up to the exponential pre-factor $e^{t\theta(s)}$, whereas the remainder must be evolved approximately using tMPS. We then add the two back together to give the overall state.

Errors

Due to the numerical nature of the methods used here, errors are unavoidable. The first error is introduced due to the approximation of the leading eigenvector. In particular, $|\psi^{\rm LD}\rangle$ has an error which can be measured through the variance with respect to the generator [1]. In practice, for the times considered here, this error is small and can be considered negligible. The dominant sources of error come from our approximation to the remainder, $|\psi_t^{\rm rem}\rangle$. This is calculated by evolving the initial state forward in time using a Trotter-decomposed MPO approximation to the evolution operator $U(\delta) = e^{\delta \mathbb{H}_s}$. Here we use a second order Trotter decomposition which entails an error $O\left(N\delta t^3\right)$ per time step (and system size N), resulting in the accumulated error $O\left(t\delta t^2\right)$ for $M = t/\delta t$ time steps. Furthermore, after each time step we then truncate the MPS to an upper-bounded bond dimension. In practice, we measure these truncation errors to be very small. Figure S2 compares the measured (log) partition sum for against numerically exact results for small system sizes (left panel), with the inset showing the error

$$\delta Z = \left| \frac{\ln(Z^{\rm exact}) - \ln(Z^{\rm MPS})}{\ln(Z^{\rm exact})} \right|. \tag{S6}$$

We observe the largest discrepancy around the time where the $|\psi_t^{\rm rem}\rangle$ becomes less dominant than $|\psi_t^{\rm LD}\rangle$. At large times of course, the leading eigenvector exponentially dominates and thus the error drops.

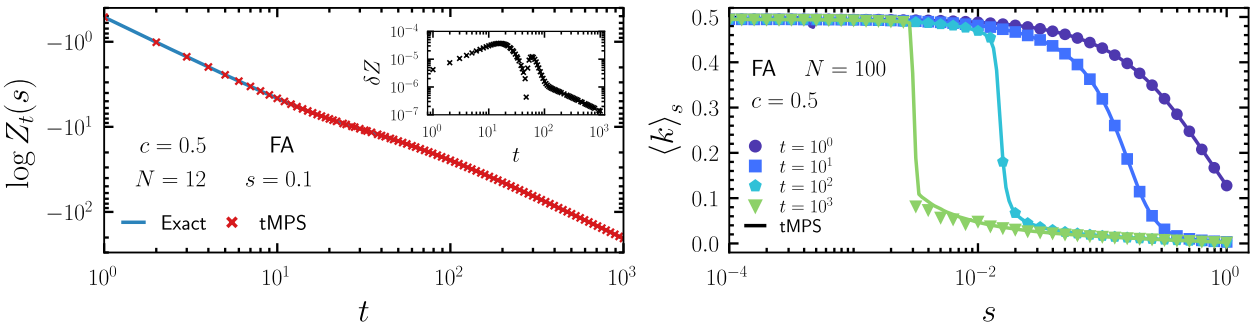


FIG. S2. Demonstration of the errors accumulated using the time evolution scheme here. The left panel compares results from exact numerics (line) and MPS (crosses) for a small system size. The inset shows the error (S6). The right panel compares the measured activity through determination of the partition sum (line), and sampling using the MPS reference dynamics (symbol) for various times and a large system size $N = 100$.

As discussed in the main text, an attempt can be made to correct on some of these errors by using the MPS retrieved after time evolution of *half* the trajectory time as a reference dynamics for umbrella sampling, see Ref. [6]. Granted enough simulations, if the reference dynamics well approximates the true dynamics, then we could see slight improvements on the measured dynamical activity - if the expected activity from the partition sum largely differs from this result however, it could indicate substantial errors. The right panel of the figure below shows this comparison for the FA model with $N = 100$. Notice the overwhelming agreement between results, with only small errors around the transition point at large times, although it still correctly predicts the location of the transition point.

[1] M. C. Bañuls and J. P. Garrahan, Phys. Rev. Lett. **123**, 200601 (2019).
[2] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, Ann. Phys. **411**, 167998 (2019).
[3] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
[4] M. Suzuki, Commun. Math. Phys **51**, 183 (1976).
[5] U. Schollwöck, Ann. Phys. **326**, 96 (2011).
[6] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **103**, 062144 (2021).

Appendix B

Supplemental material for "Optimal sampling of dynamical large deviations in two dimensions via tensor networks"

The following document is the "Supplemental material" to the arXiv pre-print (arXiv: 2209.11788) *"Optimal sampling of dynamical large deviations in two dimensions via tensor networks"* by Luke Causer, Mari Carmen Bañuls and Juan P. Garrahan. It includes some basic details on defining a stochastic dynamics on a 2D space, and a simple update scheme for optimizing PEPS.

Supplemental Material: Optimal sampling of dynamical large deviations in two dimensions via tensor networks

Luke Causer,[1,2] Mari Carmen Bañuls,[3,4] and Juan P. Garrahan[1,2]

[1] *School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK*
[2] *Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham, NG7 2RD, UK*
[3] *Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*
[4] *Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, D-80799 München*

STOCHASTIC DYNAMICS

We consider systems that live on a two-dimensional square lattice of size $N = L \times L$, where each site can take the binary values $n_{\boldsymbol{k}} = 0$ or 1, and $\boldsymbol{k} = (k_x, k_y)$ denotes the position of the lattice sites, $k_x, k_y = 1 \cdots L$. The system evolves under continuous-time Markov dynamics, defined by the transition rates $w_{x\to y}$ from configuration x to y. The average dynamics of the system can be encoded by a probability distribution $P_x(t)$, which describes the probability for the system to be in some configuration x at time t. This can be compactly written as a vector of probabilities, $|P(t)\rangle = \sum_x P_x(t)\,|x\rangle$, where $\sum_x P_x(t) = 1$. The evolution of the probability distribution is given by the master equation,

$$\frac{d}{dt}P_x(t) = \sum_{y\neq x} w_{y\to x}P_t(y) - R_x P_t(x), \tag{S1}$$

where $R_x = \sum_{y\neq x} w_{x\to y}$ is the escape rate out of the configuration x. It is convenient to write the master equation in terms of a Markov generator,

$$\mathbb{W} = \sum_{x,y\neq x} w_{x\to y}\,|y\rangle\langle x| - \sum_x R_x\,|x\rangle\langle x|\,, \tag{S2}$$

which yields $\partial_t\,|P(t)\rangle = \mathbb{W}\,|P(t)\rangle$. The Markov generator $\mathbb{W}$ conserves probability, $\langle -|\,\mathbb{W} = 0$, with the *flat state* $\langle -| = \sum_x \langle x|$, and has maximal eigenvalue zero.

For the models considered here, the Markov generator can easily be written in terms of local operators. The first model we consider is the 2D East model, with the Markov generator

$$\mathbb{W}^{\text{East}} = \sum_{\boldsymbol{k}} P_{\boldsymbol{k}}\Big[c\left(\sigma^+_{\boldsymbol{k}} - (1-n_{\boldsymbol{k}})\right) + (1-c)\left(\sigma^-_{\boldsymbol{k}} - n_{\boldsymbol{k}}\right)\Big], \tag{S3}$$

where $\sigma^\pm_{\boldsymbol{k}}$ are the Pauli raising/lowering operators at site $\boldsymbol{k}$, $c \in (0, 1/2]$ parameterizes the transition rates, and the kinetic constraint is $P_{(k_x,k_y)} = n_{(k_x-1,\,k_y)} + n_{(k_x,\,k_y-1)}$. The second model we consider is the 2D symmetric simple exclusion process (SSEP),

$$\mathbb{W}^{\text{SSEP}} = \sum_{\langle \boldsymbol{k},\boldsymbol{l}\rangle}\Big[\sigma^+_{\boldsymbol{k}}\sigma^-_{\boldsymbol{l}} - (1-n_{\boldsymbol{k}})n_{\boldsymbol{l}} + \sigma^-_{\boldsymbol{k}}\sigma^+_{\boldsymbol{l}} - n_{\boldsymbol{k}}(1-n_{\boldsymbol{l}})\Big] + \frac{1}{2}\sum_{\boldsymbol{k}\in\partial}\Big[\sigma^x_{\boldsymbol{k}} - 1\Big], \tag{S4}$$

where $\langle \boldsymbol{k},\boldsymbol{l}\rangle$ denotes a pair of nearest neighbours, $\sigma^x_i = \sigma^+_i + \sigma^-_i$, and ∂ denotes the boundary of the lattice. See the main text for a description of each.

Mapping the tilted generator onto a Hermitian operator

The long time statistics for the models are encoded in the leading eigenvalue $\theta(s)$ and eigenvectors of the tilted Markov generator, $\mathbb{W}_s$, retrieved by multiplying the off-diagonal components of Eqs. (S3-S4) by e^{-s},

$$\langle l_s|\,\mathbb{W}_s = \theta(s)\,\langle l_s|\,\mathbb{W}_s, \tag{S5}$$

$$\mathbb{W}_s\,|r_s\rangle = \theta(s)\mathbb{W}_s\,|r_s\rangle\,. \tag{S6}$$

The probability distribution over the configuration space under these biased dynamics then behaves as

$$P(x) \asymp \frac{\langle l_s|x\rangle \langle x|r_s\rangle}{\langle l_s|r_s\rangle} \tag{S7}$$

for sufficiently large times. The tilted Markov generator provides an efficient way to estimate the time-averaged dynamical properties by directly targeting its leading eigenvectors. However, we can exploit the fact that each of the models considered here obeys detailed balance. We define $\mathbb{P}$ as the diagonal matrix whose elements are the square roots of the steady state probability of $\mathbb{W}$. Then we can use the similarity transformation $\mathbb{H}_s = \mathbb{P}^{-1}\mathbb{W}_s\mathbb{P}$, to define the Hermitian matrix $\mathbb{H}_s$, with maximal eigenvalue and associated eigenvector

$$\mathbb{H}_s |\psi_s\rangle = \theta(s) |\psi_s\rangle , \tag{S8}$$

where $|\psi_s\rangle$ is related to the original eigenvectors by $|r_s\rangle = \mathbb{P}|\psi_s\rangle$ and $\langle l_s| = \langle \psi_s|\mathbb{P}^{-1}$. The models considered here allow for a simple representation as a Hermitian matrix,

$$\mathbb{H}_s^{\text{East}} = \sum_{\boldsymbol{k}} P_{\boldsymbol{k}} \Big[e^{-s}\sqrt{c(1-c)}\sigma_i^x - c(1-n_{\boldsymbol{k}}) - (1-c)n_{\boldsymbol{k}} \Big], \tag{S9}$$

$$\mathbb{H}_s^{\text{SSEP}} = \sum_{\langle \boldsymbol{k},\boldsymbol{l}\rangle} \Big[e^{-s}\sigma_{\boldsymbol{k}}^+\sigma_{\boldsymbol{l}}^- - (1-n_{\boldsymbol{k}})n_{\boldsymbol{l}} + e^{-s}\sigma_{\boldsymbol{k}}^-\sigma_{\boldsymbol{l}}^+ - n_{\boldsymbol{k}}(1-n_{\boldsymbol{l}}) \Big] + \frac{1}{2}\sum_{\boldsymbol{k}\in\partial} \Big[e^{-s}\sigma_{\boldsymbol{k}}^x - 1 \Big]. \tag{S10}$$

This representation is convenient due to the fact its expectation with any vector $|\psi\rangle$ is bounded by the maximal eigenvalue through the Rayleigh-Ritz variational principle,

$$\frac{\langle\psi|\mathbb{H}_s|\psi\rangle}{\langle\psi|\psi\rangle} \leq \theta(s). \tag{S11}$$

Furthermore, using this Hermitian matrix means that we only need to determine one eigenvector. Notice that in this representation, the leading eigenvector encodes the *probability amplitudes* of each configuration, $|\psi_s\rangle = \sum_x \psi(x)|x\rangle$, with $|\psi(x)|^2 = P(x)$.

PROJECTED ENTANGLED-PAIR STATES

The long-time dynamics of deformed Markov generators can be encoded by the probability amplitude, $\psi(\boldsymbol{n})$, with configurations $\boldsymbol{n} = (n_{\boldsymbol{k}_1}, n_{\boldsymbol{k}_2}, \cdots, n_{\boldsymbol{k}_N})$. This can be written more compactly as a vector of probability amplitudes, $|\psi\rangle = \sum_{\boldsymbol{n}} \psi(\boldsymbol{n})|\boldsymbol{n}\rangle$. This vector has the size 2^N for N lattice sites, and quickly becomes intractable to store. However, it is often the case that the components $\psi(\boldsymbol{n})$ are not independent of each other, such that it is possible to efficiently approximate the vector by an object with a smaller dimensionality (described by a number of parameters less than d^N). This realisation is at the heart of tensor network (TN) approximations. A TN representation of $\psi(\boldsymbol{n})$ amounts to decomposing the N-dimensional object into a network of many smaller tensors, connected along additional virtual dimensions, which are contracted (i.e. multiplied and summed over) to retrieve the original global tensor.

For the case of a 2D square lattice with $N = L \times L$ sites, the most natural TN ansatz is the *projected entangled-pair state* (PEPS). In this ansatz, each system (lattice site) is assigned its own rank-5 tensor, $A_{\boldsymbol{k}_j}^{d_j}$, where $\boldsymbol{k}_j$ denotes the position of system j, and d_j the state of the system. One of the dimensions corresponds to the physical dimension of the subsystem with size d, and the other four dimensions are virtual ones which connect the tensor to the tensors of the four neighbouring lattice sites. These virtual dimensions are of size D_{PEPS} (often referred to as the *bond dimension*), which controls the amount of mutual information shared between the lattice sites. It follows that each tensor is parameterized by dD_{PEPS}^4 parameters, with a maximum of $N_p = NdD_{\text{PEPS}}^4$ parameters for the whole PEPS. By specifying the local configuration of each site, $n_{\boldsymbol{k}}$, and contracting over the network, one is able to determine $\psi(\boldsymbol{n})$,

$$\psi(\boldsymbol{n}) = \mathcal{F}\left(A_{\boldsymbol{k}_1}^{n_{\boldsymbol{k}_1}} A_{\boldsymbol{k}_2}^{n_{\boldsymbol{k}_2}} \cdots A_{\boldsymbol{k}_N}^{n_{\boldsymbol{k}_N}} \right), \tag{S12}$$

where $\mathcal{F}$ is a function which represents the contraction over all virtual bond dimensions. It is convenient to represent the TN pictorially, as illustrated for $|\psi\rangle$ in Fig. S1(a). Notice that we refer to this as $|\psi\rangle$, as none of the physical dimensions are specified. The green cubes correspond to the tensors, the grey lines represent the virtual bond

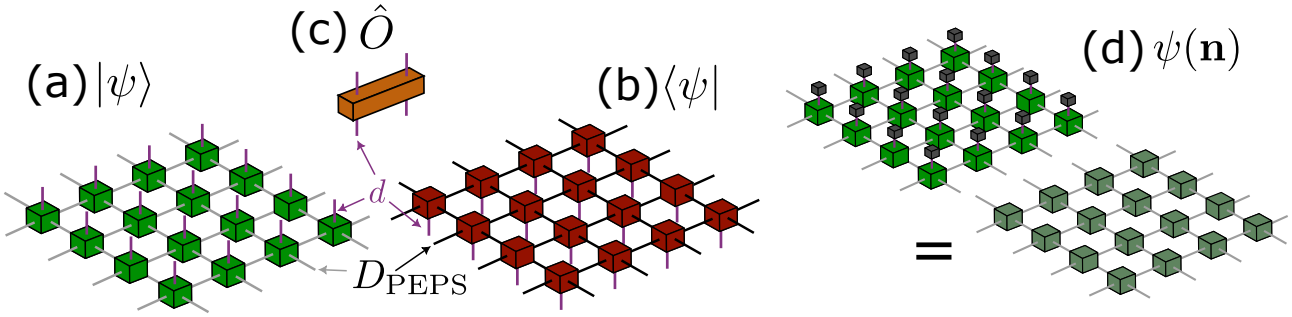


FIG. S1. **PEPS.** (a) The vector $|\psi\rangle$ as a PEPS. The green cubes represent tensors within the PEPS, the (open) purple lines are their physical dimensions, and the (closed) grey lines are the virtual dimensions between tensors. (b) The same for its conjugate, $\langle\psi|$. The conjugate tensors are represented by red cubes and black lines. (c) The two-body operator, $\hat{O}$, as a tensor. (d) The probability amplitude $\psi(\boldsymbol{n})$ can be retrieved by specifying the local dimension of each tensor, as shown by the grey cubes. This results in a network of rank-4 tensors, which give $\psi(\boldsymbol{n})$ when contracted.

dimensions to be contracted over, and the pink lines represent the physical dimensions. Similarly, we show the conjugate, $\langle\psi|$, in Fig. S1(b). It is also possible to represent local operators as a tensor, such as the two-body operators in the tilted generators. Local two-body operators can be represented by a rank-4 tensor, as shown in Fig. S1(c). Figure S1(d) shows how the probability amplitude $\psi(\boldsymbol{n})$ can be retrieved from the PEPS by specifying the local dimension for each system. This reduces the PEPS to a network of rank-4 tensors, which give $\psi(\boldsymbol{n})$ when contracted.

Contracting PEPS

PEPS allow for an efficient representation of probability amplitude vectors for large system sizes. However, to perform any tractable calculations, such as calculating the expectation of observables, it is necessary to have a way to efficiently contract the networks. Figures S2(a, b) show the networks which must be contracted to determine the inner products $\langle\psi|\psi\rangle$ and $\langle\psi|\hat{O}|\psi\rangle$ respectively, for some two-body operator $\hat{O}$. In general, contracting exactly such a two-dimensional TN is an intractable problem [1], as any exact contraction strategy has a cost that scales at least exponentially in L. For the ansatz to be of practical use, one needs ways to approximately (but precisely) contract the TN with a tractable cost.

A popular approach, and the one we will take here, is the boundary matrix product state (MPS) scheme [2, 3]. This involves contracting from the edge of the TN, one row (column) at a time, and approximating the result by an MPS (in this case, a collection of rank-4 tensors), whose bond dimension χ_{B} controls the accuracy of the approximation. We demonstrate this process in Figs. S3(a-d). The tensors at the boundary can be contracted to form the first boundary MPS, see Fig. S3(a). This boundary MPS is then contracted with the subsequent layer, and approximated by another boundary MPS, again with bond dimension χ_{B}, see Fig. S3(a). Applying the procedure from two opposing edges of the network up to the rows (or columns) of lattice sites neighbouring the two-body operator, one is able to approximate the expectation value depicted in Fig. S2(b) by the one shown in Fig. S3(e). This network can then be contracted from the edges exactly, resulting in Fig. S3(f), which can be easily and exactly contracted to determine the approximation of $\langle\psi|\hat{O}|\psi\rangle$. Contracting the networks in Fig. S3 scales as $O(\chi_B^3 D_{\mathrm{PEPS}}^4 + \chi_B^2 D_{\mathrm{PEPS}}^6)$.

A heuristic choice for χ_B is $\chi_B \sim O(D^2)$, in which case gives the total complexity $O(D^{10})$. Figure S4 demonstrates the role of the bond dimension of the boundary MPS, χ_B. We optimize the MPS using the simple update described below, and calculate its expectation value with respect to $\mathbb{H}_s$, $E(\chi_B) = \langle\psi_s|\mathbb{H}_s|\psi_s\rangle$, measured with the bond dimension χ_B. We compare the measurement to that using $\chi = 128$, which can be considered to be quasi-exact,

$$\Delta E(\chi_B) = \left| \frac{E(\chi_B) - E(128)}{E(128)} \right|. \tag{S13}$$

Indeed, the results confirm that $\chi_B \sim O(D^2)$ is a reasonable choice for approximating the environment. We use the bond dimension $\chi = 50$ when checking for convergence in observables during the optimization procedure, and $\chi = 200$ when taking the final measurements to ensure convergence in the boundary approximation.

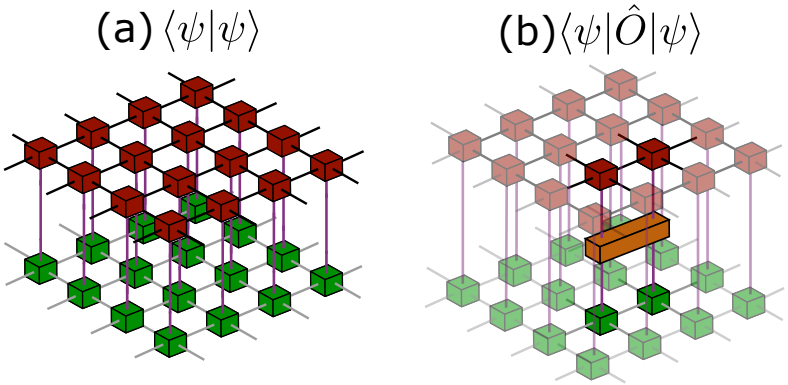


FIG. S2. **PEPS networks.** (a) The inner product $\langle\psi|\psi\rangle$ and (b) the expectation value $\langle\psi|\hat{O}|\psi\rangle$ as TNs. For visual convenience, the operator and the tensors which it is connected to are show in full colour, while the remaining tensors are semitransparent.

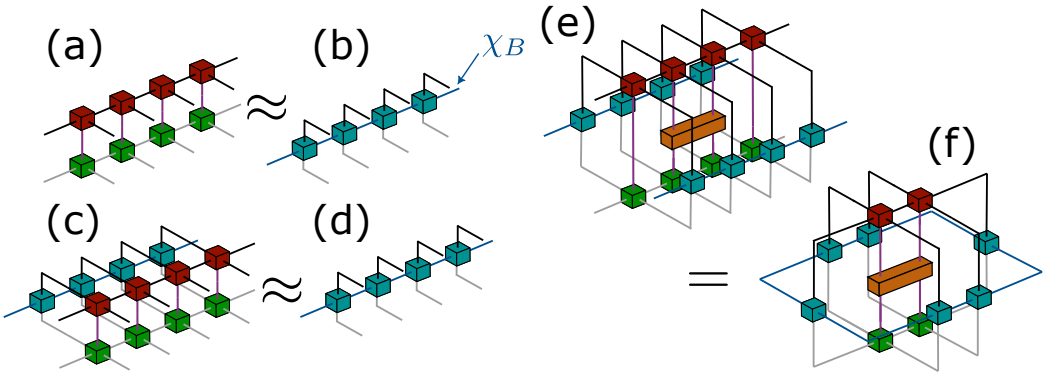


FIG. S3. **Contracting PEPS networks.** (a) The edge of a PEPS network can be approximated by (b) a boundary MPS with the auxiliary bond dimension χ_B. (c) The boundary MPS can be contracted with subsequent layers within the PEPS network, and approximated by (d) another boundary MPS also with the auxiliary bond dimension χ_B. By contracting from two opposite edges of the PEPS network, we have the reduced network shown in (e). One can then exactly contract from both edges of the reduced network to receive the network shown in (f). This can be exactly contracted to give the approximation of $\langle\psi|\hat{O}|\psi\rangle$.

Time evolution

To find a PEPS approximation for the leading eigenvector of the tilted generator, one needs to find a suitable optimization procedure. There are many approaches to optimizing TNs, but the one taken here will be to employ *time evolution* (often referred to imaginary time evolution for quantum many-body systems). The main idea is to project some probability amplitude vector onto the leading eigenvector of the tilted generator $\mathbb{H}_s$ by applying the time propagator operator, $U(\delta) = e^{\delta\mathbb{H}_s}$, to $|\psi\rangle$ until convergence is met, $|\psi_s\rangle = \lim_{t\to\infty} U(t)\,|\psi\rangle$ (up to normalization). In practice, the complete time propagation operator is difficult to compute, as it requires the matrix exponentiation of the tilted generator on the complete state space. However, for small $\delta \ll 1$, we can approximate the time propagation by a sequence of *Trotter gates*,

$$U(\delta) \approx \prod_{\langle \boldsymbol{k},\boldsymbol{l}\rangle} U_{\boldsymbol{k},\boldsymbol{l}}(\delta), \tag{S14}$$

where $U_{\boldsymbol{k},\boldsymbol{l}}(\delta) = e^{\delta H_{\boldsymbol{k},\boldsymbol{l}}}$, and $H_{\boldsymbol{k},\boldsymbol{l}}$ are the terms in $\mathbb{H}_s$ which act only on the lattice sites $\boldsymbol{k}$ and $\boldsymbol{l}$. This approximation is often referred to as a first-order Trotter decomposition, with each set of gates having an error of $O(\delta^2)$. This is the approach we take. However, it is possible to improve the accuracy by using higher order Trotter decompositions [4].

Each gate can be individually applied to the PEPS, see Fig. S5(a). The goal is to approximate the application of the

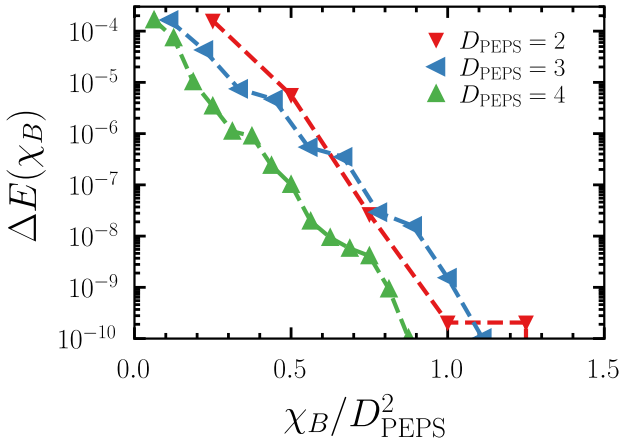


FIG. S4. **Accuracy of the boundary approximation.** We show the relative difference in the expectation value, $\Delta E(\chi_B)$, when a boundary MPS bond dimension χ_B is used, compared to $\chi_B = 128$. Results are for the 2D East model with $N = 10 \times 10$, $c = 0.5$ and $s = -1.0$.

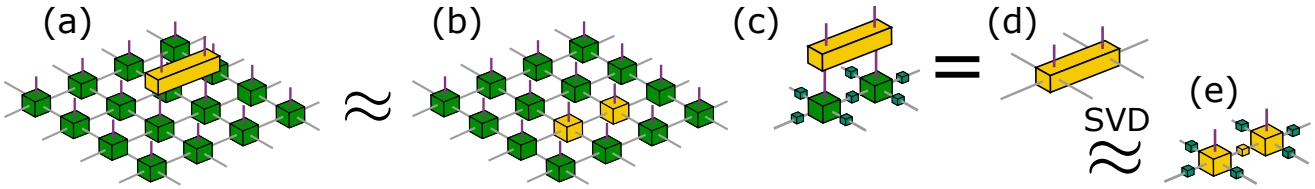


FIG. S5. **Time evolution.** (a) The PEPS $|\psi\rangle$ can be updated at a local neighbouring pair of sites through the application of a Trotter gate (shown by the yellow cuboid). (b) This can be approximated by a PEPS with same bond dimension. (c) The updated tensors can be retrieved using the SU. The small turqouise cubes are the "λ-matrices" retrieved from SVDs of the neighbouring tensors (see e.g. Ref. [5]). (d) First, we contract the lattice site tensors and the surrounding λ-matrices with the Trotter gate. (e) Then, through an SVD, we can restore the PEPS manifold. The λ-matrices outside of the pair of tensors are restored to their original values, while the λ-matrix between the two tensors is updated.

gate to PEPS by another PEPS with the same bond dimension, see Fig. S5(b). Naively contracting the gate moves the PEPS away from the PEPS manifold, as shown in Figs. S5(c, d). In order to restore the PEPS, we need an update scheme which restores the form of the original two tensors from the PEPS used in the contraction. Two popular approaches to achieving this task are the Simple Update (SU) [5] and the Full Update (FU) [6–8]. In general, to find the optimal truncation, which optimizes the overlap of the new PEPS with the untruncated TN, the environment of the pair of tensors (i.e. the contraction of the remaining PEPS tensors) needs to be taken into account. This is however a costly computation, and the SU includes only a primitive, but computationally inexpensive approximation of it as a product [9], and performs a truncated singular value decomposition (SVD) to split the resulting tensor, see Figs. S5(d, e). This results in a simple algorithm, with computational cost scaling as $O(D^5_{\mathrm{PEPS}})$. [10].

In contrast, the FU takes into account the full environment, but requires the approximate contraction of the complete TN, as illustrated in Figs. S2 and S3. This update has a greater accuracy, but with a much steeper scaling of $O(\chi_B^3 D^4_{\mathrm{PEPS}} + \chi_B^2 D^6_{\mathrm{PEPS}})$.

In what follows, we demonstrate the effectiveness of both methods for the system sizes $N = 10 \times 10$. For both approaches, we use an update schedule which reduces the time step from in the range $\delta \in [10^{-3}, 10^{-1}]$. After many iterations of time evolution, we estimate the expectation value of the state, $E = \langle\psi|\mathbb{H}_s|\psi\rangle$. This process is repeated until we find convergence. Figures S6(a, b) show the difference in the expectation value, $\Delta E = (E - E_{\mathrm{DMRG}})/E_{\mathrm{DMRG}}$, with respect to the quasi-exact results of 2D density matrix renormalization group (DMRG) for the 2D East and 2D SSEP respectively, and various values of s. The circles show the results of the SU, and the squares show the results of the FU. While it is clear in most instances the FU can provide significant improvements on the SU, it is worth noting that even the SU provides precise results for bond dimension $D = 4$, with errors $\delta E \lesssim O(10^{-3})$. We find these errors to be sufficiently small, and thus proceed using only the SU to allow us to reach large system sizes.

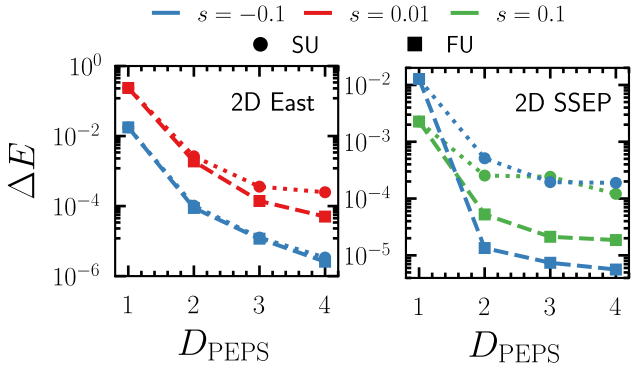


FIG. S6. **Optimization of PEPS.** The error in the measured expectation value for the SU and the FU compared to the high accuracy 2D DMRG (with an MPS bond dimension up to $D_{\mathrm{MPS}} = 1024$) for various values of s and a 10×10 lattice. The left panel shows the 2D East model with $c = 0.5$, and the right panel shows the 2D SSEP. The PEPS environment in the FU uses a boundary dimension $\chi_B = 4D^2$ for the East and $\chi_B = 6D^2$ for the SSEP.

OPTIMAL SAMPLING OF DYNAMICS

The long time statistics of dynamical observables are encoded in the deformed Markov generator $\mathbb{W}_s$. While in principle this object can be used to generate the trajectories which correspond to the statistics, it is difficult to do in practice due to the unnormalized nature of $\mathbb{W}_s$. That is, $\langle -| \mathbb{W}_s \neq 0$, and in general, the leading eigenvalue $\theta(s) \neq 0$. In the long time limit, we are able to overcome this difficulty using the so-called *Doob dynamics*, which maps the tilted generator onto a proper stochastic dynamics through the transformation $\mathbb{W}_s^{\mathrm{Doob}} = \mathbb{L}\left[\mathbb{W}_s - \theta(s)\mathbb{I}\right]\mathbb{L}^{-1}$, where $\mathbb{L} = \sum_x l_s(x) |x\rangle \langle x|$ is the left eigenvector as a diagonal matrix, $l_s(x) = \langle l_s | x \rangle$. It is simple to check that the flat state $\langle -|$ is an eigenvector of $\mathbb{W}_s^{\mathrm{Doob}}$ with the maximal eigenvalue zero. The Doob dynamics has the transition rates

$$\tilde{w}_{x\to y} = \frac{l_s(y)}{l_s(x)} e^{-s} w_{x\to y}, \tag{S15}$$

and escape rates $\tilde{R}_x = R_x + \theta(s)$.

While the Doob dynamics provides an efficient way to simulate the biased dynamics at long times, it is dependent on the fact that one has access to the leading eigenvector of the tilted generator. Using the PEPS optimization methods described above, we are able to estimate the left eigenvector $\langle l_s | \approx \langle \psi_s | \mathbb{P}^{-1}$, where $\langle \psi_s |$ is our approximation to the leading eigenvector of $\mathbb{H}_s$. Note that retrieving $\langle l_s |$ as a PEPS is simple due to the fact that $\mathbb{P}$ acts locally. This allows us to implement an efficient sampling algorithm which can estimate the dynamics Eq. (S15) by sampling $l_s(x)$ directly form our PEPS.

Extracting $l_s(x)$ from the PEPS is done similarly to estimating the contraction of the PEPS networks in Fig. S2. The first thing to notice is that we can reduce the PEPS to a network of rank-4 tensor by specifying the value of the local index of each tensor, which is defined by the configuration x (that is, the configuration x specifies each local $n_{\boldsymbol{k}}$), see Figs. S7(a, b). The values $l_s(x)$ are then retrieved by contracting the network. As was done for the networks in Fig. S3, we can estimate the exact contraction of the network using the boundary MPS method, with some bond dimension χ_B. However, this time the boundary MPS are composed of rank-3 tensors, and a heuristic choice for the bond dimension is $\chi_B \sim O(D_{\mathrm{PEPS}})$, see Fig. S7(c). By contracting from two opposing edges of the network, we can then estimate $l_s(x)$ through the exactly contractable MPS-MPS product, see Fig. S3(d).

Unlike the networks in Fig. S2—which are composed of two PEPS layers—the sampling of $l_s(x)$ only requires us to contract over a single PEPS. This leads to a significant reduction in computational cost, with each calculation of $l_s(x)$ only costing $O(ND_{\mathrm{PEPS}}^6)$. At each Monte Carlo step in the stochastic simulation algorithm, we need to calculate $l_s(x)$ for a maximum of N configurations, and thus the cost is $O(N^2 D_{\mathrm{PEPS}}^6)$. However, by recycling our partial contractions when calculating each $l_s(x)$, it is possible to reduce this to a cost of $O(ND_{\mathrm{PEPS}}^6)$ for each Monte Carlo step, see Refs. [8, 11]. While our approach provides a way to nearly optimally sample the Doob dynamics, the PEPS used in the sampling is only approximate. These errors must be accounted for through the use of umbrella sampling, see the main text and Ref. [11] for further details.

Figures S8, S9 demonstrate trajectories sampled for the 2D East and the 2D SSEP respectively, using the approximate Doob dynamics. The vertical plot in the left shows a light yellow line each time a transition occurs, and the

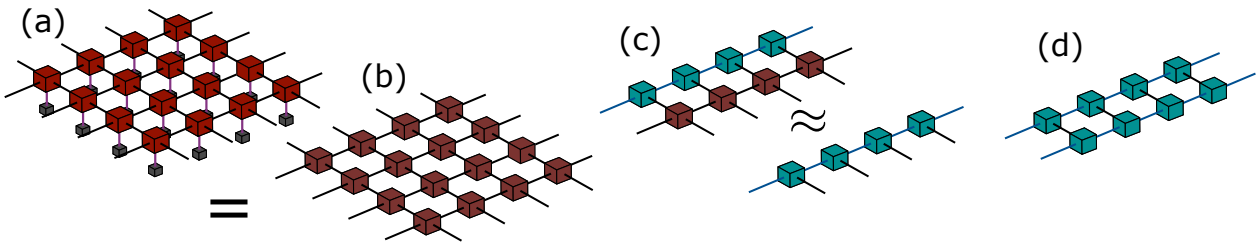


FIG. S7. **Optimal sampling from PEPS.** (a) The component $l_s(x) = \langle l_s|x\rangle$ can be extracted from the PEPS. (b) The physical dimensions can be contracted to give a PEPS composed of rank-4 tensors. (c) As is the case in Fig. S3, the contraction of the PEPS can be estimated through a boundary MPS, this time composed of rank-3 tensors and bond dimension χ_B. (d) By contracting from two opposing edges of the network, we can reduce approximate the contraction as an MPS-MPS product.

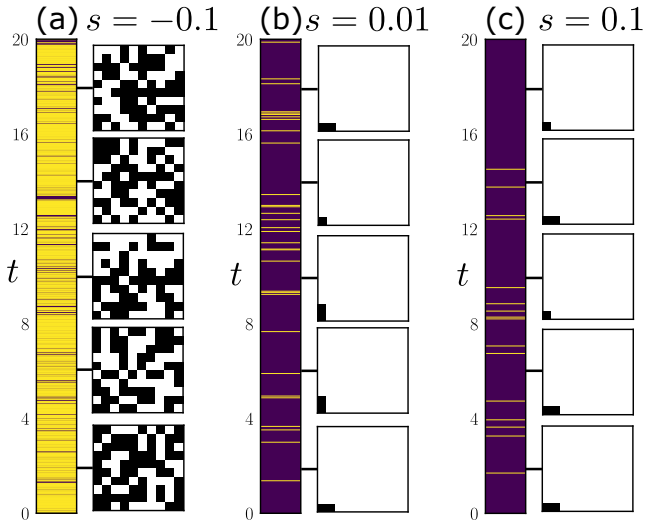


FIG. S8. **Representative trajectories for the 2D East.** Trajectories sampled from the approximate Doob dynamics for $L = 10$, $c = 0.5$, and (a) $s = -0.1$, (b) $s = 0.01$ and (c) $s = 0.1$. The vertical bars show the times when jumps occur (yellow/bright lines). The snapshots show the configurations at the marked times (black/white indicates a occupied/unoccupied).

right panels demonstrate a configuration snapshot at some points in time, as marked in the figures. We show (a) an active trajectory, (b) a trajectory close to the transition point s_c and (c) an inactive trajectory.

COMPARISON TO OTHER METHODS

A full detailed comparison to the other methods via their numerical implementation is beyond the scope of this work. However, we will provide a brief discussion explaining how this approach compares to other approaches. One popular approach to estimating large deviations is cloning methods. This approach has an exponential cost in system size, and is known to suffer from bias. This is most apparent around the transition point, meaning that while the method allows for a way to probe dynamical phase transitions, it is not reliable for an accurate finite size scaling analysis as performed here. Another popular approach is trajectory sampling methods. While this approach is unbiased, it suffers an exponential cost in both time and space. Methods such as transition path sampling [12] can be used to hinder the cost, reducing the exponent in the exponential. Nevertheless, the cost is still exponential, and can be problematic where large times are required.

Tensor network approaches estimate the long time statistics of the dynamics by directly targetting the maximal eigenvalue of a deformed Markov generator. Each iteration of the optimization methods scales only polynomially in system size for a fixed bond dimension. In the case of time evolution, the number of iterations required to reach convergence is expected to scale as the inverse gap between the two leading eigenvalues of the deformed Markov

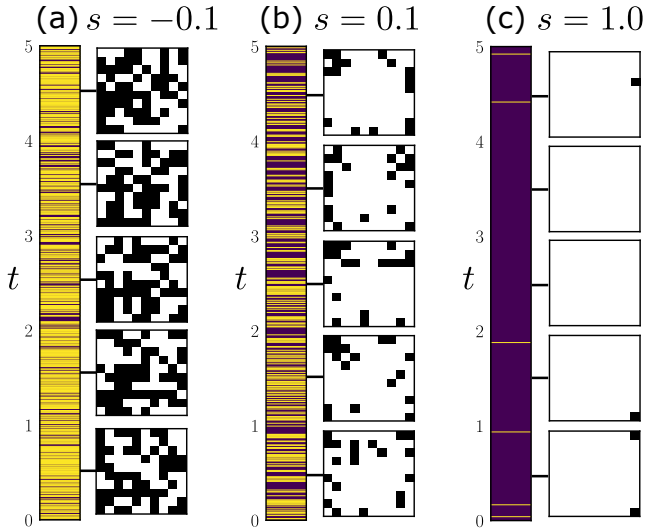


FIG. S9. **Representative trajectories for the 2D SSEP.** Trajectories sampled from the approximate Doob dynamics for $L = 10$ and (a) $s = -0.1$, (b) $s = 0.1$ and (c) $s = 1.0$. The vertical bars show the times when particle hops occur (yellow/bright lines). The snapshots show the configurations at the marked times (black/white indicates a particle/hole).

generator. Furthermore, whereas the required bond dimension for a fixed precision is not known a priori, PEPS allow for a controlled way to systematically increase the accuracy of the method by increasing the bond dimension at a cost which is only polynomial in bond dimension. As shown in these works, the PEPS can be combined with trajectory sampling algorithms. This allows us to approximate the most optimal sampling dynamics at a cost which again scales only polynomially in both space and PEPS bond dimension. In practice, the dynamics is only approximate, and errors will still exponentially propagate in time. However, the prefactor is hugely reduced from the usual setting.

[1] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **98**, 140506 (2007).
[2] F. Verstraete and J. I. Cirac, (2004), arXiv:cond-mat/0407066.
[3] F. Verstraete, V. Murg, and J. Cirac, Adv. Phys **57**, 143 (2008).
[4] M. Suzuki, J. Math. Phys. **26**, 601 (1985).
[5] H. C. Jiang, Z. Y. Weng, and T. Xiang, Phys. Rev. Lett. **101**, 090603 (2008).
[6] P. Corboz, R. Orús, B. Bauer, and G. Vidal, Phys. Rev. B **81**, 165104 (2010).
[7] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, Phys. Rev. B **90**, 064425 (2014).
[8] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orús, Phys. Rev. B **92**, 035142 (2015).
[9] M. Lubasch, J. I. Cirac, and M.-C. Bañuls, New J. Phys. **16**, 033014 (2014).
[10] As presented here, the complexity is $O(D^6_{\mathrm{PEPS}})$, but can be made cheaper by considering the reduced tensor method [7].
[11] L. Causer, M. C. Bañuls, and J. P. Garrahan, Phys. Rev. E **103**, 062144 (2021).
[12] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Ann. Rev. Phys. Chem **53**, 291 (2002).

Bibliography

[1]M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley, Sept. 2008).

[2]D. C. Handscomb, "The Monte Carlo method in quantum statistical mechanics", Math. Proc. Camb. Philos. Soc **58**, 594–598 (1962).

[3]A. W. Sandvik, "Stochastic Series Expansion Methods", arXiv preprint (2019).

[4]S. Lu, M. C. Bañuls, and J. I. Cirac, "Algorithms for Quantum Simulation at Finite Energies", PRX Quantum **2**, 020321 (2021).

[5]Y. Yang, J. I. Cirac, and M. C. Bañuls, "Classical algorithms for many-body quantum systems at finite energies", Phys. Rev. B **106**, 024307 (2022).

[6]A. J. Ferris, "Unbiased Monte Carlo for the age of tensor networks", arXiv preprint (2015).

[7]M. Frías-Pérez, M. Mariën, D. Pérez-García, M. C. Bañuls, and S. Iblisdir, "Collective Monte Carlo updates through tensor network renormalization", arXiv preprint (2021).

[8]D. Layden, G. Mazzola, R. V. Mishmash, M. Motta, P. Wocjan, J.-S. Kim, and S. Sheldon, "Quantum-enhanced Markov chain Monte Carlo", arXiv preprint (2022).

[9]U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states", Ann. Phys **326**, 96–192 (2011).

[10]R. Orús, "A practical introduction to tensor networks: matrix product states and projected entangled pair states", Ann. Phys **349**, 117–158 (2014).

[11]R. Orús, "Tensor networks for complex quantum systems", Nat. Rev. Phys **1**, 538–550 (2019).

[12]J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, "Matrix product states and projected entangled pair states: Concepts, symmetries, theorems", Rev. Mod. Phys. **93**, 045003 (2021).

[13]M. C. Bañuls, "Tensor Network Algorithms: a Route Map", arXiv preprint (2022).

[14]S. R. White, "Density matrix formulation for quantum renormalization groups", Phys. Rev. Lett. **69**, 2863–2866 (1992).

[15]S. Östlund and S. Rommer, "Thermodynamic limit of density matrix renormalization", Phys. Rev. Lett. **75**, 3537–3540 (1995).

[16]J. Dukelsky, M. A. Martín-Delgado, T. Nishino, and G. Sierra, "Equivalence of the variational matrix product method and the density matrix renormalization group applied to spin chains", EPL **43**, 457 (1998).

[17]G. Vidal, "Efficient Simulation of One-Dimensional Quantum Many-Body Systems", Phys. Rev. Lett. **93**, 040502 (2004).

[18]S. R. White, "Minimally Entangled Typical Quantum States at Finite Temperature", Phys. Rev. Lett. **102**, 190601 (2009).

[19]E. M. Stoudenmire and S. R. White, "Minimally entangled typical thermal state algorithms", New J. Phys **12**, 055026 (2010).

[20]Á. M. Alhambra and J. I. Cirac, "Locally Accurate Tensor Networks for Thermal States and Time Evolution", PRX Quantum **2**, 040331 (2021).

[21]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).

[22]G. Evenbly and G. Vidal, "Class of Highly Entangled Many-Body States that can be Efficiently Simulated", Phys. Rev. Lett. **112**, 240502 (2014).

[23]F. Verstraete and J. I. Cirac, "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions", arXiv preprint (2004).

[24]A. Honecker and I. Peschel, "Matrix-product states for a one-dimensional lattice gas with parallel dynamics", J. Stat. Phys **88**, 319–345 (1997).

[25]Y. Hieida, "Application of the Density Matrix Renormalization Group Method to a Non-Equilibrium Problem", J. Phys. Soc. Jpn. **67**, 369–372 (1998).

[26]A. Kemper, A. Gendiar, T. Nishino, A. Schadschneider, and J. Zittartz, "Stochastic light-cone CTMRG: a new DMRG approach to stochastic models", J. Phys. A: Math. Gen. **36**, 29 (2002).

[27]K. Ueda, R. Otani, Y. Nishio, A. Gendiar, and T. Nishino, "Snapshot Observation for 2D Classical Lattice Models by Corner Transfer Matrix Renormalization Group", J. Phys. Soc. Jpn. **74**, 111–114 (2005).

[28]A. Novikov, M. Trofimov, and I. Oseledets, "Exponential Machines", arXiv preprint (2016).

[29]S. T. Wauthier, B. Vanhecke, T. Verbelen, and B. Dhoedt, "Learning Generative Models for Active Inference using Tensor Networks", arXiv preprint (2022).

[30]E. Gillman, D. C. Rose, and J. P. Garrahan, "Reinforcement Learning with Tensor Networks: Application to Dynamical Large Deviations", arXiv preprint (2022).

[31]E. M. Stoudenmire, "Learning relevant features of data with multi-scale tensor networks", Quantum Sci. Technol. **3**, 034003 (2018).

[32]I. Glasser, N. Pancotti, and J. I. Cirac, "From probabilistic graphical models to generalized tensor networks for supervised learning", arXiv preprint (2018).

[33]Z.-Y. Han, J. Wang, H. Fan, L. Wang, and P. Zhang, "Unsupervised Generative Modeling Using Matrix Product States", Phys. Rev. X **8**, 031012 (2018).

[34]S. Cheng, L. Wang, T. Xiang, and P. Zhang, "Tree tensor networks for generative modeling", Phys. Rev. B **99**, 155131 (2019).

[35]E. Gillman, D. C. Rose, and J. P. Garrahan, "A Tensor Network Approach to Finite Markov Decision Processes", arXiv preprint (2020).

[36]R. Sengupta, S. Adhikary, I. Oseledets, and J. Biamonte, "Tensor networks in machine learning", arXiv preprint (2022).

[37]I. Convy and K. B. Whaley, "Interaction Decompositions for Tensor Network Regression", arXiv preprint (2022).

[38]T. Vieijra, L. Vanderstraeten, and F. Verstraete, "Generative modeling with projected entangled-pair states", arXiv preprint (2022).

[39]H. Touchette, "The large deviation approach to statistical mechanics", Phys. Rep **478**, 1–69 (2009).

[40]J. P. Garrahan, "Aspects of non-equilibrium in classical and quantum systems: Slow relaxation and glasses, dynamical large deviations, quantum non-ergodicity, and open quantum dynamics", Physica A **504**, 130–154 (2018).

[41]R. L. Jack, "Ergodicity and large deviations in physical systems with stochastic dynamics", **93**, 74 (2020).

[42]D. Limmer, C. Gao, and A. Poggioli, "A large deviation theory perspective on nanoscale transport phenomena", Eur. Phys. J. B **94**, 145 (2021).

[43]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).

[44]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).

[45]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).

[46]G. H. Fredrickson and H. C. Andersen, "Kinetic Ising Model of the Glass Transition", Phys. Rev. Lett. **53**, 1244–1247 (1984).

[47]R. G. Palmer, D. L. Stein, E. Abrahams, and P. W. Anderson, "Models of Hierarchically Constrained Dynamics for Glassy Relaxation", Phys. Rev. Lett. **53**, 958–961 (1984).

[48]J. Jäckle and S. Eisinger, "A hierarchically constrained kinetic Ising model", Z. Phys. B: Cond. Matt **84**, 115–124 (1991).

[49]G. Biroli, K. Miyazaki, and D. R. Reichman, "Dynamical Heterogeneity in Glass-Forming Liquids", arXiv preprint (2022).

[50]L. Berthier and G. Biroli, "Glasses and aging, a statistical mechanics perspective on", in *Encyclopedia of complexity and systems science*, edited by R. A. Meyers (Springer New York, New York, NY, 2009), pp. 4209–4240.

[51]D. Chandler and J. P. Garrahan, "Dynamics on the Way to Forming Glass: Bubbles in Space-Time", Annu. Rev. Phys. Chem. **61**, 191–217 (2010).

[52]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).

[53]T. Nemoto, É. Fodor, M. E. Cates, R. L. Jack, and J. Tailleur, "Optimizing active work: dynamical phase transitions, collective motion, and jamming", Phys. Rev. E **99**, 022605 (2019).

[54]T. Agranov, M. E. Cates, and R. L. Jack, "Entropy production and its large deviations in an active lattice gas", Journal of Statistical Mechanics: Theory and Experiment **2022**, 123201 (2022).

[55]J. P. Garrahan and I. Lesanovsky, "Thermodynamics of Quantum Jump Trajectories", Phys. Rev. Lett. **104**, 160601 (2010).

[56]T. Komorowski, C. Landim, S. Olla, T. Komorowski, C. Landim, and S. Olla, "The simple exclusion process", in *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), pp. 155–197.

[57]L. Bertini and G. Giacomin, "Stochastic Burgers and KPZ Equations from Particle Systems", Commun. Math. Phys **183**, 571–607 (1997).

[58]T. Bodineau and B. Derrida, "Cumulants and large deviations of the current through non-equilibrium steady states", C. R. Phys **8**, 540–555 (2007).

[59]C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van Wijland, "Universal cumulants of the current in diffusive systems on a ring", Phys. Rev. E **78**, 021122 (2008).

[60]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 (2009).

[61]M. Gorissen and C. Vanderzande, "Current fluctuations in the weakly asymmetric exclusion process with open boundaries", Phys. Rev. E **86**, 051114 (2012).

[62]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).

[63]M. Vanicat, E. Bertin, V. Lecomte, and E. Ragoucy, "Mapping current and activity fluctuations in exclusion processes: consequences and open questions", SciPost Phys. **10**, 028 (2021).

[64]V. Lecomte, J. Garrahan, and F. Wijland, "Inactive dynamical phase of a symmetric exclusion process on a ring", J. Phys. A: Math. Theor. **45**, 175001 (2012).

[65]L. Causer, I. Lesanovsky, M. C. Bañuls, and J. P. Garrahan, "Dynamics and large deviation transitions of the XOR-Fredrickson-Andersen kinetically constrained model", Phys. Rev. E **102**, 052132 (2020).

[66]C. Ates, T. Pohl, T. Pattard, and J. M. Rost, "Antiblockade in Rydberg Excitation of an Ultracold Lattice Gas", Phys. Rev. Lett. **98**, 023002 (2007).

[67]T. M. Wintermantel, Y. Wang, G. Lochead, S. Shevate, G. K. Brennen, and S. Whitlock, "Unitary and Nonunitary Quantum Cellular Automata with Rydberg Arrays", Phys. Rev. Lett. **124**, 070503 (2020).

[68]P. P. Mazza, R. Schmidt, and I. Lesanovsky, "Vibrational Dressing in Kinetically Constrained Rydberg Spin Systems", Phys. Rev. Lett. **125**, 033602 (2020).

[69]T. Amthor, C. Giese, C. S. Hofmann, and M. Weidemüller, "Evidence of Antiblockade in an Ultracold Rydberg Gas", Phys. Rev. Lett. **104**, 013001 (2010).

[70]I. Lesanovsky and J. P. Garrahan, "Kinetic Constraints, Hierarchical Relaxation, and Onset of Glassiness in Strongly Interacting and Dissipative Rydberg Gases", Phys. Rev. Lett. **111**, 215305 (2013).

[71]I. Lesanovsky and J. P. Garrahan, "Out-of-equilibrium structures in strongly interacting Rydberg gases with dissipation", Phys. Rev. A **90**, 011603 (2014).

[72]M. Hoening, W. Abdussalam, M. Fleischhauer, and T. Pohl, "Antiferromagnetic long-range order in dissipative Rydberg lattices", Phys. Rev. A **90**, 021603 (2014).

[73]A. Urvoy, F. Ripka, I. Lesanovsky, D. Booth, J. P. Shaffer, T. Pfau, and R. Löw, "Strongly Correlated Growth of Rydberg Aggregates in a Vapor Cell", Phys. Rev. Lett. **114**, 203002 (2015).

[74]M. M. Valado, C. Simonelli, M. D. Hoogerland, I. Lesanovsky, J. P. Garrahan, E. Arimondo, D. Ciampini, and O. Morsch, "Experimental observation of controllable kinetic constraints in a cold atomic gas", Phys. Rev. A **93**, 040701 (2016).

[75]M. Marcuzzi, J. ř. Miná ř, D. Barredo, S. de Léséleuc, H. Labuhn, T. Lahaye, A. Browaeys, E. Levi, and I. Lesanovsky, "Facilitation Dynamics and Localization Phenomena in Rydberg Lattice Gases with Position Disorder", Phys. Rev. Lett. **118**, 063606 (2017).

[76]M. Ostmann, M. Marcuzzi, J. P. Garrahan, and I. Lesanovsky, "Localization in spin chains with facilitation constraints and disordered interactions", Phys. Rev. A **99**, 060101 (2019).

[77]J. W. P. Wilkinson, T. ž. Prosen, and J. P. Garrahan, "Exact solution of the "rule 150" reversible cellular automaton", Phys. Rev. E **105**, 034124 (2022).

[78]S. Gopalakrishnan and B. Zakirov, "Facilitated quantum cellular automata as simple models with non-thermal eigenstates and dynamics", Quantum Sci. Technol. **3**, 044004 (2018).

[79]T. Iadecola and M. Schecter, "Quantum many-body scar states with emergent kinetic constraints and finite-entanglement revivals", Phys. Rev. B **101**, 024306 (2020).

[80]Z.-C. Yang, F. Liu, A. V. Gorshkov, and T. Iadecola, "Hilbert-Space Fragmentation from Strict Confinement", Phys. Rev. Lett. **124**, 207602 (2020).

[81]D. K. Mark, C.-J. Lin, and O. I. Motrunich, "Unified structure for exact towers of scar states in the Affleck-Kennedy-Lieb-Tasaki and other models", Phys. Rev. B **101**, 195131 (2020).

[82]F. Carollo, M. Gnann, G. Perfetto, and I. Lesanovsky, "Signatures of a quantum stabilized fluctuating phase and critical dynamics in a kinetically constrained open many-body system with two absorbing states", Phys. Rev. B **106**, 094315 (2022).

[83]L. Causer, J. P. Garrahan, and A. Lamacraft, "Slow dynamics and large deviations in classical stochastic Fredkin chains", Phys. Rev. E **106**, 014128 (2022).

[84]O. Salberger and V. Korepin, "Fredkin Spin Chain", arXiv preprint (2016).

[85]K. Adhikari and K. S. D. Beach, "Deforming the Fredkin spin chain away from its frustration-free point", Phys. Rev. B **99**, 54436 (2019).

[86]K. Adhikari and K. S. D. Beach, "Slow dynamics of the Fredkin spin chain", Phys. Rev. B **104**, 115149 (2021).

[87]C. M. Langlett and S. Xu, "Hilbert space fragmentation and exact scars of generalized Fredkin spin chains", Phys. Rev. B **103**, L220304 (2021).

[88]H. Singh, B. A. Ware, R. Vasseur, and A. J. Friedman, "Subdiffusion and Many-Body Quantum Chaos with Kinetic Constraints", Phys. Rev. Lett. **127**, 230602 (2021).

[89]J. Richter and A. Pal, "Anomalous hydrodynamics in a class of scarred frustration-free Hamiltonians", Phys. Rev. Research **4**, L012003 (2022).

[90]H. Singh, R. Vasseur, and S. Gopalakrishnan, "The Fredkin staircase: An integrable system with a finite-frequency Drude peak", arXiv preprint (2022).

[91]L. Dell'Anna, O. Salberger, L. Barbiero, A. Trombettoni, and V. E. Korepin, "Violation of cluster decomposition and absence of light cones in local integer and half-integer spin chains", Phys. Rev. B **94**, 10.1103/physrevb.94.155140 (2016).

[92]X. Chen, E. Fradkin, and W. Witczak-Krempa, "Gapless quantum spin chains: multiple dynamics and conformal wavefunctions", J. Phys. A: Math. Theor. **50**, 464002 (2017).

[93]O. Salberger, T. Udagawa, Z. Zhang, H. Katsura, I. Klich, and V. Korepin, "Deformed Fredkin spin chain with extensive entanglement", J. Stat. Mech.: Theory Exp **2017**, 63103 (2017).

[94]T. Udagawa and H. Katsura, "Finite-size gap, magnetization, and entanglement of deformed Fredkin spin chain", J. Phys. A: Math. Theor. **50**, 405002 (2017).

[95]R. Movassagh, "The gap of Fredkin quantum spin chain is polynomially small", Adv. Math. Sci. Appl. **3**, 531–562 (2018).

[96]F. Sugino and V. Korepin, "Rényi entropy of highly entangled spin chains", Int. J. Mod. Phys. B **32**, 1850306 (2018).

[97]L. Dell'Anna, "Long-distance entanglement in Motzkin and Fredkin spin chains", SciPost Phys **7**, 10.21468/scipostphys.7.4.053 (2019).

[98]L. Dell'Anna, L. Barbiero, and A. Trombettoni, "Dynamics and correlations in Motzkin and Fredkin spin chains", J. Stat. Mech.: Theory Exp. **2019**, 124025 (2019).

[99]L. Causer, M. C. Bañuls, and J. P. Garrahan, "Optimal sampling of dynamical large deviations via matrix product states", Phys. Rev. E **103**, 062144 (2021).

[100]L. Causer, M. C. Bañuls, and J. P. Garrahan, "Finite Time Large Deviations via Matrix Product States", Phys. Rev. Lett. **128**, 090605 (2022).

[101]L. Causer, M. C. Bañuls, and J. P. Garrahan, "Optimal sampling of dynamical large deviations in two dimensions via tensor networks", arXiv preprint (2022).

[102]C. Gardiner, *Stochastic Methods: a handbook for the natural and social sciences*, 4th ed. (Springer, Berlin, 2009).

[103]L. Peliti and S. Pigolotti, *Stochastic Thermodynamics* (Princeton University Press, 2021).

[104]J. R. Norris, *Markov Chains* (Cambridge University Press, 1997).

[105]F. J. Dyson, "The Radiation Theories of Tomonaga, Schwinger, and Feynman", Phys. Rev. **75**, 486–502 (1949).

[106]D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", J. Comput. Phys **22**, 403–434 (1976).

[107]D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", J. Phys. Chem. **81**, 2340–2361 (1977).

[108]A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, "A new algorithm for Monte Carlo simulation of Ising spin systems", J. Comput. Phys **17**, 10–18 (1975).

[109]V. Lecomte, C. Appert-Rolland, and F. van Wijland, "Thermodynamic Formalism for Systems with Markov Dynamics", J. Stat. Phys. **127**, 51–106 (2007).

[110]C. Maes, "Frenesy: Time-symmetric dynamical activity in nonequilibria", Phys. Rep. **850**, Frenesy: time-symmetric dynamical activity in nonequilibria, 1–33 (2020).

[111]D. Chandler, *Introduction to Modern Statistical Mechanics*, 1st (OUP USA, 1987).

[112]N. E. Strand, H. Vroylandt, and T. R. Gingrich, "Using tensor network states for multi-particle Brownian ratchets", J. Chem. Phys. **156**, 221103 (2022).

[113]N. E. Strand, H. Vroylandt, and T. R. Gingrich, "Computing time-periodic steady-state currents via the time evolution of tensor network states", J. Chem. Phys. **157**, 054104 (2022).

[114]J. Gu and F. Zhang, "Tensor-Network Approaches to Counting Statistics for the Current in a Boundary-Driven Diffusive System", (2022).

[115]T. H. E. Oakes, A. Moss, and J. P. Garrahan, "A deep learning functional estimator of optimal dynamics for sampling large deviations", Mach. Learn.: Sci. Technol. **1**, 035004 (2020).

[116]S. Whitelam, D. Jacobson, and I. Tamblyn, "Evolutionary reinforcement learning of dynamical large deviations", J. Chem. Phys. **153**, 044113 (2020).

[117]C. Casert, T. Vieijra, S. Whitelam, and I. Tamblyn, "Dynamical Large Deviations of Two-Dimensional Kinetically Constrained Models Using a Neural-Network State Ansatz", Phys. Rev. Lett. **127**, 120602 (2021).

[118]D. C. Rose, J. F. Mair, and J. P. Garrahan, "A reinforcement learning approach to rare trajectory sampling", New J. Phys **23**, 013013 (2021).

[119]A. Das, D. C. Rose, J. P. Garrahan, and D. T. Limmer, "Reinforcement learning of rare diffusive dynamics", J. Chem. Phys. **155**, 134105 (2021).

[120]J. Yan, H. Touchette, and G. M. Rotskoff, "Learning nonequilibrium control forces to characterize dynamical phase transitions", Phys. Rev. E **105**, 024115 (2022).

[121]S.-H. Lin and F. Pollmann, "Scaling of Neural-Network Quantum States for Time Evolution", Phys. Status Solidi B **259**, 2100172 (2022).

[122]S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, "Time-evolution methods for matrix-product states", Ann. Phys **411**, 167998 (2019).

[123]T. Oakes, S. Powell, C. Castelnovo, A. Lamacraft, and J. P. Garrahan, "Phases of quantum dimers from ensembles of classical stochastic trajectories", Phys. Rev. B **98**, 064302 (2018).

[124]L. M. Vasiloiu, T. H. E. Oakes, F. Carollo, and J. P. Garrahan, "Trajectory phase transitions in noninteracting spin systems", Phys. Rev. E **101**, 042115 (2020).

[125]R. Hurtado-Gutiérrez, F. Carollo, C. Pérez-Espigares, and P. I. Hurtado, "Building Continuous Time Crystals from Rare Events", Phys. Rev. Lett. **125**, 160601 (2020).

[126]J. Dolezal and R. L. Jack, "Mechanical analysis of a dynamical phase transition for particles in a channel", Phys. Rev. Research **4**, 033134 (2022).

[127]T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, "Population-dynamics method with a multicanonical feedback control", Phys. Rev. E **93**, 062123 (2016).

[128]G. Ferré and H. Touchette, "Adaptive Sampling of Large Deviations", J. Stat. Phys **172**, 1525–1544 (2018).

[129]D. A. Levin, E. Wilmer, and Y. Peres, *Markov Chains and Mixing Times* (American Mathematical Society, 2009).

[130]R. Chetrite and H. Touchette, "Variational and optimal control representations of conditioned and driven processes", J. Stat. Mech.: Theory Exp **2015**, P12001 (2015).

[131]J. P. Garrahan, "Classical stochastic dynamics and continuous matrix product states: gauge transformations, conditioned and driven processes, and equivalence of trajectory ensembles", J. Stat. Mech.: Theory Exp **2016**, 073208 (2016).

[132]U. Ray and G. Kin-Lic Chan, "Constructing auxiliary dynamics for nonequilibrium stationary states by variance minimization", J. Chem. Phys. **152**, 104107 (2020).

[133]S. Marcantoni, C. Pérez-Espigares, and J. P. Garrahan, "Symmetry-induced fluctuation relations in open quantum systems", Phys. Rev. E **104**, 014108 (2021).

[134]F. Carollo, J. P. Garrahan, and R. L. Jack, "Large Deviations at Level 2.5 for Markovian Open Quantum Systems: Quantum Jumps and Quantum State Diffusion", J. Stat. Phys **184**, 13 (2021).

[135]B. J. Zhang, T. Sahai, and Y. M. Marzouk, "A Koopman framework for rare event simulation in stochastic differential equations", J. Comput. Phys **456**, 111025 (2022).

[136]C. Dellago, P. G. Bolhuis, and D. Chandler, "Efficient transition path sampling: Application to Lennard-Jones cluster rearrangements", J. Chem. Phys. **108**, 9236–9245 (1998).

[137]P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, "Transition Path Sampling: Throwing Ropes Over Rough Mountain Passes, in the Dark", Annu. Rev. Phys. Chem. **53**, 291–318 (2002).

[138]C. Dellago, P. G. Bolhuis, and P. L. Geissler, "Transition Path Sampling Methods", in *Computer simulations in condensed matter systems: from materials to chemical biology volume 1*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006), pp. 349–391.

[139]J. T. Berryman and T. Schilling, "Sampling rare events in nonequilibrium and nonstationary systems", J. Chem. Phys. **133**, 244101 (2010).

[140]P. G. Bolhuis and D. W. H. Swenson, "Transition Path Sampling as Markov Chain Monte Carlo of Trajectories: Recent Algorithms, Software, Applications, and Future Outlook", Adv. Theory Simul. **4**, 2000237 (2021).

[141]J. F. Mair, D. C. Rose, and J. P. Garrahan, "Training neural network ensembles via trajectory sampling", arXiv preprint (2022).

[142]F. Krzakala, A. Rosso, G. Semerjian, and F. Zamponi, "Path-integral representation for quantum spin models: Application to the quantum cavity method and Monte Carlo simulations", Phys. Rev. B **78**, 134428 (2008).

[143]T. Mora, A. M. Walczak, and F. Zamponi, "Transition path sampling algorithm for discrete many-body systems", Phys. Rev. E **85**, 036710 (2012).

[144]J. P. Garrahan and D. Chandler, "Coarse-grained microscopic model of glass formers.", eng, Proceedings of the National Academy of Sciences of the United States of America **100**, 9710–4 (2003).

[145]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: Math. Theor **42**, 075007 (2009).

[146]N. Pancotti, G. Giudice, J. I. Cirac, J. P. Garrahan, and M. C. Bañuls, "Quantum East Model: Localization, Nonthermal Eigenstates, and Slow Dynamics", Phys. Rev. X **10**, 021051 (2020).

[147]L. Berthier and J. P. Garrahan, "Numerical Study of a Fragile Three-Dimensional Kinetically Constrained Model", J. Phys. Chem. B **109**, 3578–3585 (2005).

[148]S. Butler and P. Harrowell, "The origin of glassy dynamics in the 2D facilitated kinetic Ising model", J. Chem. Phys. **95**, 4454–4465 (1991).

[149]I. Hartarsky, F. Martinelli, and C. Toninelli, "Fredrickson–Andersen model in two dimensions", arXiv preprint (2022).

[150]G. Biroli and J. P. Garrahan, "Perspective: The glass transition", J. Chem. Phys. **138**, 12A301 (2013).

[151]D. J. Ashton, L. O. Hedges, and J. P. Garrahan, "Fast simulation of facilitated spin models", J. Stat. Mech.: Theory Exp **2005**, P12010 (2005).

[152]S. Kim, D. G. Thorpe, C. Noh, J. P. Garrahan, D. Chandler, and Y. Jung, "Study of the upper-critical dimension of the East model through the breakdown of the Stokes-Einstein relation.", eng, J. Chem. Phys **147**, 084504 (2017).

[153]M. D. Ediger, "Spatially Heterogeneous Dynamics in Supercooled Liquids", Annu. Rev. Phys. Chem. **51**, 99–128 (2000).

[154]S. C. Glotzer, "Spatially heterogeneous dynamics in liquids: insights from simulation", J. Non-Cryst. Solids **274**, 342–355 (2000).

[155]T. Bodineau, V. Lecomte, and C. Toninelli, "Finite Size Scaling of the Dynamical Free-Energy in a Kinetically Constrained Model", J. Stat. Phys **147**, 1–17 (2012).

[156]T. Bodineau and C. Toninelli, "Activity Phase Transition for Constrained Dynamics", Commun. Math. Phys **311**, 357–396 (2012).

[157]R. L. Jack and P. Sollich, "Large deviations of the dynamical activity in the east model: analysing structure in biased trajectories", J. Phys. A: Math. Theor. **47**, 015003 (2013).

[158]T. Nemoto, R. L. Jack, and V. Lecomte, "Finite-Size Scaling of a First-Order Dynamical Phase Transition: Adaptive Population Dynamics and an Effective Model", Phys. Rev. Lett. **118**, 115702 (2017).

[159]R. L. Jack, T. Nemoto, and V. Lecomte, "Dynamical phase coexistence in the Fredrickson–Andersen model", J. Stat. Mech.: Theory Exp **2020**, 053204 (2020).

[160]C. Flindt and J. P. Garrahan, "Trajectory Phase Transitions, Lee-Yang Zeros, and High-Order Cumulants in Full Counting Statistics", Phys. Rev. Lett. **110**, 050601 (2013).

[161]L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Large Deviations for the Boundary Driven Symmetric Simple Exclusion Process", Math. Phys. Anal. Geom. **6**, 231–267 (2003).

[162]V. Popkov, M. Salerno, and G. M. Schütz, "Asymmetric simple exclusion process with periodic boundary driving", Phys. Rev. E **78**, 011122 (2008).

[163]B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, "Exact solution of a 1D asymmetric exclusion model using a matrix formulation", J. Phys. A: Math. Gen. **26**, 1493 (1993).

[164]T. Bodineau and B. Derrida, "Current Large Deviations for Asymmetric Exclusion Processes with Open Boundaries", J. Stat. Phys **123**, 277–300 (2006).

[165]L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Large Deviations of the Empirical Current in Interacting Particle Systems", Theory Probab. Appl. **51**, 2–27 (2007).

[166]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).

[167]B. Derrida, O. Hirschberg, and T. Sadhu, "Large Deviations in the Symmetric Simple Exclusion Process with Slow Boundaries", Journal of Statistical Physics **182**, 15 (2021).

[168]T. Brewer, S. R. Clark, R. Bradford, and R. L. Jack, "Efficient characterisation of large deviations using population dynamics", J. Stat. Mech.: Theory Exp **2018**, 053204 (2018).

[169]R. A. Blythe and M. R. Evans, "Nonequilibrium steady states of matrix-product form: a solver's guide", J. Phys. A: Math. Theor. **40**, R333 (2007).

[170]H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction* (Oxford Graduate Texts, 2004).

[171]U. Kaldor, *Many-Body Methods in Quantum Chemistry* (Springer Berlin, Heidelberg, 1988).

[172]E. Rieffel and W. Polak, *Quantum Computing: A Gentle Introduction* (The MIT Press, 2014).

[173]S. Rommer and S. Östlund, "Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group", Phys. Rev. B **55**, 2164–2181 (1997).

[174]G. Evenbly and G. Vidal, "Algorithms for entanglement renormalization", Phys. Rev. B **79**, 144108 (2009).

[175]L. Cincio, J. Dziarmaga, and M. M. Rams, "Multiscale Entanglement Renormalization Ansatz in Two Dimensions: Quantum Ising Model", Phys. Rev. Lett. **100**, 240603 (2008).

[176]J. Haferkamp, D. Hangleiter, J. Eisert, and M. Gluza, "Contracting projected entangled pair states is average-case hard", Phys. Rev. Research **2**, 013010 (2020).

[177]R. Orús and G. Vidal, "Simulation of two-dimensional quantum systems on an infinite lattice revisited: corner transfer matrix for tensor contraction", Phys. Rev. B **80**, 094403 (2009).

[178]M. Lubasch, J. I. Cirac, and M.-C. Bañuls, "Unifying projected entangled pair state contractions", New J. Phys **16**, 033014 (2014).

[179]M. B. Hastings, "Solving gapped hamiltonians locally", Phys. Rev. B **73**, 085115 (2006).

[180]F. Verstraete and J. I. Cirac, "Matrix product states represent ground states faithfully", Phys. Rev. B **73**, 094423 (2006).

[181]C. Fernández-González, N. Schuch, M. M. Wolf, J. I. Cirac, and D. Pérez-García, "Frustration Free Gapless Hamiltonians for Matrix Product States", Communications in Mathematical Physics **333**, 299–333 (2015).

[182]J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pi žorn, H. Verschelde, and F. Verstraete, "Time-Dependent Variational Principle for Quantum Lattices", Phys. Rev. Lett. **107**, 070601 (2011).

[183]F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, "Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems", Phys. Rev. Lett. **93**, 207204 (2004).

[184]D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, "Matrix Product State Representations", 10.48550/ARXIV.QUANT-PH/060819 (2006).

[185]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**, 143–224 (2008).

[186]J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: area laws for the entanglement entropy", Rev. Mod. Phys. **82**, 277–306 (2010).

[187]B. Pirvu, V. Murg, J. I. Cirac, and F. Verstraete, "Matrix product operator representations", New J. Phys **12**, 025012 (2010).

[188]G. M. Crosswhite and D. Bacon, "Finite automata for caching in matrix product algorithms", Phys. Rev. A **78**, 012356 (2008).

[189]A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, "Variational matrix-product-state approach to quantum impurity models", Phys. Rev. B **80**, 165117 (2009).

[190]V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, "Variational optimization algorithms for uniform matrix product states", Phys. Rev. B **97**, 045145 (2018).

[191]B. Tuybens, J. De Nardis, J. Haegeman, and F. Verstraete, "Variational Optimization of Continuous Matrix Product States", Phys. Rev. Lett. **128**, 020501 (2022).

[192]C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", J. Res. Natl. Bur. Stand **45**, 255 (1950).

[193]P. Schmitteckert, "Nonequilibrium electron transport using the density matrix renormalization group method", Phys. Rev. B **70**, 121302 (2004).

[194]A. E. Feiguin and S. R. White, "Time-step targeting methods for real-time dynamics using the density matrix renormalization group", Phys. Rev. B **72**, 020404 (2005).

[195]R. Orús and G. Vidal, "Infinite time-evolving block decimation algorithm beyond unitary evolution", Phys. Rev. B **78**, 155117 (2008).

[196]M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, "Time-evolving a matrix product state with long-ranged interactions", Phys. Rev. B **91**, 165112 (2015).

[197]M. Suzuki, "Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems", Commun. Math. Phys. **51**, 183–190 (1976).

[198]H. C. Jiang, Z. Y. Weng, and T. Xiang, "Accurate Determination of Tensor Network State of Quantum Lattice Models in Two Dimensions", Phys. Rev. Lett. **101**, 090603 (2008).

[199]P. Corboz, R. Orús, B. Bauer, and G. Vidal, "Simulation of strongly correlated fermions in two spatial dimensions with fermionic projected entangled-pair states", Phys. Rev. B **81**, 165104 (2010).

[200]H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, and R. Orús, "Infinite projected entangled pair states algorithm improved: Fast full update and gauge fixing", Phys. Rev. B **92**, 035142 (2015).

[201]D. T. Limmer and D. Chandler, "Theory of amorphous ices", Proc. Natl. Acad. Sci. **111**, 9413–9418 (2014).

[202]A. S. Keys, D. Chandler, and J. P. Garrahan, "Using the s ensemble to probe glasses formed by cooling and aging", Phys. Rev. E **92**, 022304 (2015).

[203]P. Brighi, M. Ljubotina, and M. Serbyn, "Hilbert space fragmentation and slow dynamics in particle-conserving East model", arXiv preprint (2022).

[204]H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, "Probing many-body dynamics on a 51-atom quantum simulator", Nature **551**, 579–584 (2017).

[205]C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, "Quantum scarred eigenstates in a Rydberg atom chain: Entanglement, breakdown of thermalization, and stability to perturbations", Phys. Rev. B **98**, 155134 (2018).

[206]C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, "Weak ergodicity breaking from quantum many-body scars", Nature Physics **14**, 745–749 (2018).

[207]C.-J. Lin and O. I. Motrunich, "Exact Quantum Many-Body Scar States in the Rydberg-Blockaded Atom Chain", Phys. Rev. Lett. **122**, 173401 (2019).

[208]Y. Tang, J. Liu, J. Zhang, and P. Zhang, "Solving nonequilibrium statistical mechanics by evolving autoregressive neural networks", arXiv preprint (2022).

[209]J. P. Garrahan and D. Chandler, "Geometrical explanation and scaling of dynamical heterogeneities in glass forming systems", Phys. Rev. Lett. **89**, 035704 (2002).

[210]F. Carollo, J. P. Garrahan, I. Lesanovsky, and C. Pérez-Espigares, "Making rare events typical in Markovian open quantum systems", Phys. Rev. A **98**, 010103 (2018).

[211]F. Carollo, R. L. Jack, and J. P. Garrahan, "Unraveling the Large Deviation Statistics of Markovian Open Quantum Systems", Phys. Rev. Lett. **122**, 130605 (2019).

[212]S. Marcantoni, C. Pérez-Espigares, and J. P. Garrahan, "Symmetry-induced fluctuation relations for dynamical observables irrespective of their behavior under time reversal", Phys. Rev. E **101**, 062142 (2020).

[213]C. Ates, B. Olmos, J. P. Garrahan, and I. Lesanovsky, "Dynamical phases and intermittency of the dissipative quantum Ising model", Phys. Rev. A **85**, 043620 (2012).

[214]S. Genway, J. M. Hickey, J. P. Garrahan, and A. D. Armour, "Trajectory phases of a quantum dot model", J. Phys. A: Math. Theor. **47**, 505001 (2014).

[215]F. Carollo, J. P. Garrahan, I. Lesanovsky, and C. Pérez-Espigares, "Fluctuating hydrodynamics, current fluctuations, and hyperuniformity in boundary-driven open quantum chains", Phys. Rev. E **96**, 052118 (2017).

[216]C. Pérez-Espigares, I. Lesanovsky, J. P. Garrahan, and R. Gutiérrez, "Glassy dynamics due to a trajectory phase transition in dissipative Rydberg gases", Phys. Rev. A **98**, 021804 (2018).

[217]J. Cui, J. I. Cirac, and M. C. Bañuls, "Variational Matrix Product Operators for the Steady State of Dissipative Quantum Systems", Phys. Rev. Lett. **114**, 220601 (2015).

[218]E. Mascarenhas, H. Flayac, and V. Savona, "Matrix-product-operator approach to the nonequilibrium steady state of driven-dissipative quantum arrays", Phys. Rev. A **92**, 022116 (2015).

[219]T. Prosen and M. Žnidarič, "Matrix product simulations of non-equilibrium steady states of quantum spin chains", J. Stat. Mech.: Theory Exp. **2009**, P02035 (2009).

[220]Z. Cai and T. Barthel, "Algebraic versus Exponential Decoherence in Dissipative Many-Particle Systems", Phys. Rev. Lett. **111**, 150403 (2013).

[221]L. Bonnes, D. Charrier, and A. M. Läuchli, "Dynamical and steady-state properties of a bose-hubbard chain with bond dissipation: a study based on matrix product operators", Phys. Rev. A **90**, 033612 (2014).

[222]J. J. Mendoza-Arenas, T. Grujic, D. Jaksch, and S. R. Clark, "Dephasing enhanced transport in nonequilibrium strongly correlated quantum systems", Phys. Rev. B **87**, 235130 (2013).

[223]I. Pižorn, "One-dimensional Bose-Hubbard model far from equilibrium", Phys. Rev. A **88**, 043635 (2013).

[224]A. H. Werner, D. Jaschke, P. Silvi, M. Kliesch, T. Calarco, J. Eisert, and S. Montangero, "Positive Tensor Network Approach for Simulating Open Quantum Many-Body Systems", Phys. Rev. Lett. **116**, 237201 (2016).