# Tensor Network Methods for Nonequilibrium Statistical Mechanics

Thesis by

Phillip Laurence Helms

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2021
Defended June 4th, 2021

# ACKNOWLEDGEMENTS

To consider the family, friends, and mentors that have provided support and guidance throughout my time at Caltech is a humbling experience. None of my success here would have been possible without their influence in my life.

First, I would like to thank my advisor, Garnet Chan. To this day, it seems shocking that he happily welcomed me into his research group as a chemical engineering student with only one undergraduate-level course in quantum mechanics! His patience as I've slowly learned, often through error, has been incredibly encouraging. I am continually impressed by his flexibility, letting me pursue my independent interests, and providing thoughtful discussions and insight into problems he would not have normally been interested in. His kindness and understanding towards my family and me were of monumental importance and I cannot thank him enough for that. His scientific prowess, high expectations, and mentorship have set a high standard that I aspire to achieve as I begin my academic career.

I must also thank my thesis committee: John Brady, Zhen-Gang Wang, and Austin Minnich. Their feedback concerning my progress has helped guide my research and encouraged me to better understand how to frame it in the broader context of statistical physics.

Working alongside the other members of the Chan group has been an unforgettable experience. I am indebted to Ushnish Ray for introducing me to large deviation theory and nonequilibrium statistical mechanics. The members of the tensor network subgroup, Yang Gao, Johnnie Gray, Klaas Gunst, Reza Haghshenas, Henrik Larsson, Matt O'Rourke, and Erika Ye, have been a constant sounding board for scientific problems and I couldn't have completed this work without their insight and expertise. I have enjoyed mentorship from and friendship with all the other group members, past and present, and will always remember the group outings to dinner, Sequoia, and Universal Studios.

I am thankful for friends that have provided distractions from research problems and helped make my time at Caltech memorable. The residents in the student family housing have become a second family and the student parent group has provided both friendship and commiseration. Caltech's Graduate Student Council and members of its advocacy committee have given me a deeper sense of purpose and I am grateful for my experiences as an advocate for graduate students. To list all of the other

supportive friends would be impossible, but I would particularly like to thank the Pearson, Larsen, Paz, Wong-Lim, and Riveros families.

Last, I must express my immense gratitude for the support from my family. My parents have loved and encouraged me unconditionally; their support for my ambitions and encouragement to reach my full potential are invaluable. They are wonderful examples of kindness, support, and love. This sentiment carries over to my siblings and in-laws, who have provided much support and many adventures during their visits.

For the past year, my two daughters have served as my office mates at the kitchen table. Their bright personalities, positive attitudes, and perseverance have been examples to me in the most difficult times. Their love has brought me deep happiness, fun adventures, and inspiration.

To my wife Brittni, I owe all of my success at Caltech and in the future to you. You have been a constant source of support and kindness. Thank you for caring for me unconditionally, encouraging me to pursue my dreams, supporting me in my most difficult times, grounding me, and bringing joy to my daily life. Words cannot express my appreciation for you.

The work included in this thesis was supported through the ARCS Foundation through an ARCS Award and the National Science Foundation's Graduate Research Fellowship Program under grant DGE-1745301.

# ABSTRACT

Large deviation theory has emerged as a powerful mathematical scaffolding for studying nonequilibrium statistical mechanics, particularly for characterizing the macroscopic effects of microscopic fluctuations. While the large deviation approach is firmly established, it relates the effects of fluctuations to the likelihood of exponentially rare events, which naively requires exponentially large simulation costs. This, in turn, necessitates the development of appropriate numerical simulation techniques. While the standard Monte Carlo toolkit has expanded to incorporate methods towards making rare events typical, in this thesis I propose and evaluate a powerful unorthodox approach adopted from quantum simulation, namely tensor network algorithms, which can work in concert with standard methods to deepen our understanding of nonequilibrium phenomena.

As a testbed for this novel approach, I consider the dynamical phase behavior of several versions of the simple exclusion process, a paradigmatic model of classical driven diffusion. Using a matrix product state, a one-dimensional tensor network ansatz, and the density matrix renormalization group algorithm, a corresponding optimization routine, I characterize the dynamical phase transition between a jammed and maximal current phase in both the one-dimensional and multi-lane simple exclusion processes. The matrix product state is found to be an efficient representation of the nonequilibrium steady-state biased to arbitrarily rare currents via large deviation theory. Because the one-dimensional ansatz is limited to finite-width systems, I extend this success to study the fully two-dimensional simple exclusion process. There, the projected entangled pair state, a two-dimensional tensor network ansatz, is used with the time evolution via block decimation algorithm to demonstrate that the phase transition observed in one-dimension persists in the fully two-dimensional system.

Towards the goal of making tensor network methods adaptable for a broad range of physically important systems, both classical and quantum, I also present progress towards studying systems in the continuum limit with interacting particles in two dimensions. This builds upon previous work proposing tensor network representations of quantum operators with long-range interactions in two dimensions by evaluating three operator representations in practice and finding two competitive and viable approaches.

# PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Y. Gao, P. Helms, G. K.-L. Chan, E. Solomonik, *arXiv:2007.08056* **2020**, https://arxiv.org/abs/2007.08056
P.L.H. participated in the conception of the project, designing and implementing algorithms, running benchmarking calculations, analyzing data, and writing the manuscript.

[2] P. Helms, G. K.-L. Chan, *Physical Review Letters* **2020**, *125*, 140601, https://doi.org/10.1103/PhysRevLett.125.140601
P.L.H. participated in the conception of the project, selecting and implementing algorithms, performing calculations, analyzing data, and writing the manuscript.

[3] M. Motta, C. Genovese, F. Ma, Z.-H. Cui, R. Sawaya, G. K.-L. Chan, N. Chepiga, P. Helms, C. Jiménez-Hoyos, A. J. Millis, U. Ray, E. Ronca, H. Shi, S. Sorella, E. M. Stoudenmire, S. R. White, S. Zhang, *Physical Review X* **2020**, *10*, 031058,
https://doi.org/10.1103/PhysRevX.10.031058
P.L.H. participated in setting up and running coupled cluster calculations, analyzing results, creating figures, and writing supplementary material.

[4] P. Helms, U. Ray, G. K.-L. Chan, *Physical Review E* **2019**, *100*, 022101, https://doi.org/10.1103/PhysRevE.100.022101
P.L.H. participated in the conception of the project, selecting and implementing algorithms, performing calculations, analyzing data, and writing the manuscript.

# TABLE OF CONTENTS

# LIST OF ILLUSTRATIONS

# LIST OF TABLES

# Chapter 1

# INTRODUCTION

This chapter provides relevant background to properly contextualize the work reported in the following chapters. We provide a brief, broad overview of modern nonequilibrium statistical mechanics before discussing how it intersects with the mathematical framework provided by large deviation theory. An introduction to tensor networks in one and two dimensions is then provided, along with a description of standard tensor network ground state algorithms. This chapter is concluded with a discussion on the intersection between these three topics, motivating the use of tensor networks for studying large deviations in nonequilibrium statistical mechanics.

## 1.1 Introduction

Statistical mechanics provides a framework for computing the macroscopic properties of a system from an understanding of the microscopic behaviors of its constituent particles. With the scaffolding for studying statistical mechanics having been erected over a century ago and its use having evolved into a modern pillar of physical science, many of the key results have remained limited to systems in equilibrium. Because many physically important systems lie within regimes characterized as far from equilibrium and due to the dearth of robust, flexible, and appropriate formalisms, entities such as the National Academy of Sciences and the United States Department of Energy have labeled the characterization and control of matter far from equilibrium as a current grand challenge in basic science [1, 2]. In recent years, large deviation theory (LDT), a mathematical approach originally created to understand the occurrence and probability of rare events in stochastic systems, has been identified as a natural language and appropriate framework for nonequilibrium statistical mechanics (NESM) [3–6]. Recent work applying LDT to classical physical systems has led to significant discoveries such as insights into the structure and dynamics of glassy systems [7], design principles for both simple models [8] and molecular system [9], and deepened understanding of myriad nonequilibrium phenomena [10–16]

With the utility of LDT established, a present bottleneck towards its widespread use is the development of efficient and accurate computational methods for the numerical determination of relevant quantities [17–25]. While Monte Carlo algorithms of various flavors are commonly employed, this thesis centers on a fundamentally different approach, namely tensor networks (TNs), which are borrowed and adapted from the field of quantum simulation.

In this introductory chapter, we seek to provide all of the background information necessary to understand the details of the following chapters where TNs are used to characterize the behavior of paradigmatic model systems. To contextualize LDT within the larger field of NESM, the next section provides a high-level overview of this field. In Section 1.3 we give a primer on large deviation theory, focusing on its application to equilibrium and nonequilibrium statistical mechanics and providing a small number of example calculations to illustrate and highlight key facets. Last, Section 1.4 starts with an overview of tensor networks and their origin before moving into a technical description of the standard algorithms used in Chapters 2 and 3 of this thesis. We close there with a motivation for the use of TNs in the context of

LDT for NESM.

## 1.2 Overview of Nonequilibrium Statistical Mechanics

We begin now by reviewing some introductory NESM, starting in the next subsection with a review of the equilibrium approach, where we provide an outline of the standard approach for equilibrium systems and define the differences between equilibrium and nonequilibrium systems.

In the subsequent subsection, we discuss the current landscape of theoretical approaches for studying the statistical mechanics of systems out of equilibrium. We begin by describing the diversity of nonequilibrium systems and providing some broad and practical criteria for the establishment of a robust and generalizable framework for use in NESM. Last, we give a high-level overview of current theoretical approaches in specific physical regimes.

### Introductory Equilibrium Statistical Mechanics

As is typical in other fields of physical science, we must begin by specifying the system that we will study and the degrees of freedom it is allowed to explore. In statistical mechanics, this specification refers to the full microscopic state space of the system. A full specification of the state of the microscopic system is referred to as a microscopic configuration, or microstate, which we will denote here as $C$. For a system of $N$ interacting particles, a microstate consists of a specification in $d$ dimensions of the position $x_i$ and momentum $p_i$ for each particle, making $C$ a single point in a $2dN$-dimensional space. Alternatively, for a spin-1/2 lattice model, a microstate is the specification of a spin for each site in the lattice, meaning the number of possible configurations grows exponentially with the number of lattice sites, $N_{\text{microstates}} = 2^{N_{\text{sites}}}$.

To provide an introduction to the fundamentals of statistical mechanics, we can consider the system of interacting particles. For an isolated system in equilibrium, Ludwig Boltzmann hypothesized that, in the long time limit, every possible microstate has an equally likely chance of occurring, i.e.

$$P(C) = 1/\Omega(N, V, E), \tag{1.1}$$

where $\Omega(N, V, E)$ is called the microcanonical partition function and is the number of possible configurations with a fixed number of particles $N$, volume $V$, and energy $E$, which is given by the system's Hamiltonian, i.e. $E = \mathcal{H}(C)$. There are many possible configurations with a given $N$, $V$, and $E$ and we call the set of all such configurations the microcanonical ensemble.

By simply counting the number of configurations in this ensemble, we get the microcanonical partition function, and can immediately determine the entropy, a thermodynamic state variable of the macroscopic system, as

$$S = k_B \ln\left(\Omega(N, V, E)\right). \tag{1.2}$$

With the entropy and energy of the macroscopic system now specified, the principles of thermodynamics can be employed to obtain all other state variables and thus a full macroscopic description of the system. As a simple example, the temperature can be obtained by taking the derivative of the entropy with respect to the energy

$$T = \left(\frac{\partial S}{\partial E}\right)^{-1} = \frac{1}{k_B}\left(\frac{\partial \ln \Omega}{\partial E}\right)^{-1}, \tag{1.3}$$

thus tying its definition directly to the microcanonical partition function.

While working with an isolated system provides a simple conceptual introduction to statistical mechanics, the counting problem becomes dauntingly large for interesting chemical or physical systems and experience from thermodynamics indicates that working with other variables can often prove convenient. As an alternative, we can consider a system that is no longer isolated, but instead in contact with a thermal bath at temperature $T$, or inverse temperature $\beta = 1/k_B T$, and again attempt to determine the configurational probabilities. With the details beyond the scope of this brief review, it can be shown that the probability of having a given configuration occur in this system takes the form

$$P(C) \propto e^{-\beta H(C)}, \tag{1.4}$$

where $E_C = H(C)$ is the energy of the given configuration and the term on the right-hand side is frequently called a Boltzmann weight. To determine this probability exactly, we specify that $\sum_C P(C) = 1$ and thus

$$P(C) = \frac{e^{-\beta H(C)}}{Z(N, V, T)}, \tag{1.5}$$

where

$$Z(N, V, T) = \sum_C e^{-\beta H(C)} \tag{1.6}$$

is a normalization factor called the canonical partition function. Once again, by determining this partition function, we gain access to all thermodynamic state variables. As with the microcanonical partition function, the canonical partition function gives us directly a thermodynamic state variable, the Helmholtz free energy $F(N, V, T) = -\frac{1}{\beta}\ln(Z(N, V, T))$. Additionally, the free energy serves as a cumulant generating function for the cumulants of the energy, meaning, for example, that the average system energy can be determined by a derivative of the free energy

$$\langle E \rangle = \frac{\partial \beta F}{\partial \beta}. \tag{1.7}$$

While other ensembles are also commonly utilized when convenient, the microcanonical and canonical ensembles will be sufficient to explain the connection between large deviation theory and equilibrium statistical mechanics and the extension to NESM.

We intuitively expect that some macroscopic properties are simply averages over the microscopic degrees of freedom, for example, the density $\rho$ is the average number of particles per unit volume in the system. Equilibrium statistical mechanics also emphasizes the macroscopic importance of microscopic fluctuations. An example of this occurs when we consider the fluctuations in the energy, defined as $\langle(\delta E)^2\rangle = \langle(E - \langle E\rangle)^2\rangle$, for a system in contact with a thermal reservoir. We can show that this value can be determined via the second derivative of the free energy

$$\langle(\delta E)^2\rangle = -\frac{\partial^2 \beta F}{\partial \beta^2}. \tag{1.8}$$

Given that the first derivative of the free energy was equal to the average energy, via Equation 1.7, this expression becomes $\langle(\delta E)^2\rangle = -\frac{\partial\langle E\rangle}{\partial\beta}$, which is remarkably similar to the definition of the heat capacity, $C_v = \frac{\partial E}{\partial T}$. Combining these, we see that the heat capacity is directly related to fluctuations in the energy as provided by derivatives of the free energy

$$C_v = \frac{\langle(\delta E)^2\rangle}{k_B T^2}. \tag{1.9}$$

This emphasizes that while average behavior is obviously of fundamental importance, fluctuations in microscopic behaviors strongly influence the behavior of the macroscopic system.

With the fundamentals of equilibrium statistical mechanics established, we can transition into a discussion of the differences between equilibrium and nonequilibrium systems. To do this, we shift slightly and consider the time evolution of the lattice-based spin model mentioned earlier, though this can be generalized. For simplicity, we consider the behavior of a Markovian spin-1/2 system evolving discretely in time.

Given that the system begins in an initial configuration $C_0$, there exists some probability $P_t(C)$ that the system will inhabit a given microstate at time $t$. This probability is specified entirely by the system's evolutionary rules, which dictate how it transitions from one configuration to the next, i.e.

$$P_{t+\Delta t}(C') = w_{C\to C'} P_t(C), \tag{1.10}$$

where $w_{C\to C'}$ is the probability of the system transitioning from configuration $C$ to configuration $C'$ and is assumed here to be stationary in time. The probability current gives the flow of probability between configurations $C$ and $C'$ at time $t$ and is defined as

$$J_t(C \to C') = w_{C\to C'} P_t(C) - w_{C'\to C} P_t(C'). \tag{1.11}$$

It can be understood as the difference between the rate at which the system makes the transition $C \to C'$ and the reversed transition, $C' \to C$.

While initial dynamics are highly dependent on the beginning configuration, most systems eventually relax to a steady-state, where the configurational probability becomes constant in time, i.e. $P_{t\to\infty}(C) = P^*(C)$. Accordingly, we can also specify the steady-state probability current as

$$J^*(C \to C') = w_{C\to C'} P^*(C) - w_{C'\to C} P^*(C'). \tag{1.12}$$

A system with a vanishing probability current, i.e. $J_t(C \to C') = J^*(C \to C') = 0$, for all combinations of configurations is an equilibrium system. A system that obeys this condition is colloquially referred to as one that obeys detailed balance and must be in equilibrium, though there are also equilibrium states that break detailed balance.

It is also informative at this point to relate these results to the canonical partition function and Boltzmann weights introduced earlier. For a system that obeys detailed balance, we require that $w_{C\to C'} P^*(C) = w_{C'\to C} P^*(C')$. For the canonical partition function, we now know the form of these configurational probabilities, which dictates that the ratio of the transition rates be

$$\frac{w_{C\to C'}}{w_{C'\to C}} = e^{-\beta(H(C)-H(C'))}, \tag{1.13}$$

providing a simple constraint on the transition rates of a system in contact with a thermal bath to satisfy the conditions of detailed balance.

Failure to obey detailed balance can often be attributed to one of two conditions, both of which classify the system as being out of equilibrium. First, as briefly alluded to earlier, starting from a given initial configuration, the system will experience transient dynamics, meaning that it has not yet reached a steady-state and cannot be in equilibrium regardless of the nature of the system at long times. Second, when a system has reached a steady-state, its evolutionary rules can cause that state to exhibit finite probability currents and the system is said to be in a nonequilibrium steady-state (NESS). From a more phenomenological lens, both of these deviations from equilibrium require the system to be experiencing finite gradients in quantities such as particle density, chemical or electric potential, or thermal energy. For a transient state, gradients are artifacts of the initial configuration, for example, if dye is put into a cup of standing water, the dye will remain more concentrated at the point of insertion until it has diffused. Alternatively, for a NESS, gradients are caused by external perturbations to the system, with a common example being a thermally conducting rod connected to large hot and cold reservoirs at each end, giving a constant flow of heat through the rod.

With these highlighted foundational concepts, much of statistical mechanics can be simplified significantly, requiring only the computation of averages using Boltzmann weights.

### Introductory Nonequilibrium Statistical Mechanics

Unfortunately, when a system is not in equilibrium, the approach developed by Boltzmann, Gibbs, and others is invalidated. Not only are there no known universal rules for defining configurational probabilities *a priori*, but additional complications are introduced by inequivalence of the time and ensemble averages, implying the failure of the ergodic theorem and meaning we must take into account the spatial and dynamical behavior of a system. Despite these challenges and because myriad phenomena in physics, chemistry, biology, finance, and social science operate far from equilibrium, there has been significant effort towards extending the concepts of statistical mechanics to nonequilibrium systems.

These efforts often focus on developing an approach tailored towards a specific class of physical system or behavior. A handful of the choices that one must consider when attempting to develop a tailored approach include:

- **Distance from Equilibrium:** While the rules for what constitutes an equilibrium system are clearly defined, there are many ways a state can be out of equilibrium. Systems experiencing small perturbations from equilibrium, called near-equilibrium systems, are normally defined as perturbed systems whose responses to the perturbation are linear with respect to the perturbation strength. These are well-characterized and frequently discussed in standard texts on statistical mechanics [26–29]. On the other end of this spectrum are systems that are far from equilibrium, which can be thought of as systems where the phenomena of interest occur over a period that is orders of magnitude shorter than the time required for the equilibration of the system and environment. Between these two limits, are systems exhibiting slow relaxation, sometimes called aging kinetics [30, 31].

- **Dynamic vs. Steady-State:** For systems far from equilibrium, we must determine whether the physics of interest involve the dynamical behavior of a system or the time-independent steady-state. Frequently, studying steady-state behavior is seen as a stepping stone towards studying the full dynamic behavior, though there are physical phenomena where the transient nature is central to the problem.

- **Quantum vs. Classical:** As in equilibrium statistical mechanics, it must be determined to what extent quantum effects are expected to affect the behavior of the system. Often, treating the system classically is sufficient, though there are many cases where quantum effects play an important role in the system's behavior.

- **Deterministic vs. Stochastic Dynamics:** For nonequilibrium systems, the interplay between the environment and system complicates standard deterministic modeling. To perform a deterministic computation, we must specify the initial configuration of both the system and environment then use Hamiltonian mechanics to perform the time evolution of all particles in both the system and environment. The probabilistic part of the system's dynamics is then reintroduced by running trajectories with many different initial environment configurations. To overcome the extreme cost of tracking both the system and environment, this approach can be simplified to what is often called thermostatted dynamics [32]. Here, the system is initially in contact with a thermal bath. At the start of the evolution, the system is isolated from the environment and a field is applied. Because of the applied field, the deterministic system begins to heat up. To avoid this complication, friction terms are introduced in such a way as to maintain the correct energy. Alternatively, the system can be treated stochastically, where noise is introduced to match phenomenological observations and correctly model the effect of the environment on the system [33, 34] Within the context of stochastic systems, the master equation is the most flexible equation for time evolution. With some simplifications, one can arrive at the Langevin and Fokker-Planck equations, which are also frequently employed for modeling the dynamics of stochastic systems [33].

- **Microscopic vs. Mesoscopic:** Especially when dealing with master equation dynamics, it is prudent to carefully determine what is considered the system and environment. Frequently, this can result in an intermediate level of dynamics where many microscopic degrees of freedom are grouped into a single mesoscopic degree of freedom. By doing this, we can lose the effects of some of the microscopic behaviors but are often able to simulate larger systems by focusing on only the most important degrees of freedom. Using a microscopic model not only provides more accurate results, but is convenient when it is unclear how to define the relevant mesoscopic degrees of freedom.
- **Typical vs. Rare Behavior:** For many phenomena, we only need to study and understand the typical, or most probable, dynamics of a system. This greatly simplifies the analysis because, in simulations, we are most likely to generate typical behavior. As mentioned before, though, fluctuations often have significant effects on the macroscopic behavior of the system and for those systems, we must statistically characterize fluctuations more precisely.

With the complexity of the landscape of possible nonequilibrium systems, a universal framework for approaching NESM has remained elusive. Ideally, a universal theory would be entirely applicable to systems far from equilibrium, near-equilibrium, and in equilibrium. Additionally, the approach would be agnostic to whether the system's dynamics are microscopic or mesoscopic, quantum or classical, or deterministic or stochastic. While a steady-state approach would serve as a natural intermediate step, a fully unifying scaffolding would need to be extensible to dynamical behavior. Finally, any such approach must properly accommodate the effects of fluctuations on the macroscopic system behavior.

Given the vastness of possible nonequilibrium problems, preliminary work has focused on developing an array of theoretical treatments for specific problem classifications. The most successful regime by far has been for near-equilibrium systems. A central theoretical result in this regime is Onsager's regression hypothesis [28], which qualitatively proposed that a system prepared in a nonequilibrium state regresses, or relaxes, to equilibrium in the same way that an equilibrium system regresses from a random fluctuation into the same state. This hypothesis was later proven as a result of the fluctuation-dissipation theorem [35], another important result for systems in the linear response region. Of practical importance is the ability to determine transport coefficients, which dictate the strength of perturbation responses. The Green-Kubo relations build upon Onsager's regression hypothesis to relate transport coefficients of near-equilibrium systems to time correlations in the equilibrium system [26, 27].

For systems far from equilibrium, much work centers on understanding typical behavior of either simplified models, such as the simple exclusion process discussed in later chapters of this thesis, which are paradigmatic systems analagous to the Ising model in equilibrium statistical mechanics, or phenomenologically interesting systems [34]. Away from typicality, there are a handful of theoretical results that extend some of the key concepts from near-equilibrium systems; these include fluctuation theorems [36–39], thermodynamic uncertainty relations [40–42], and generalized fluctuation-dissipation theorems [43–45].

### 1.3 Large Deviation Theory and Statistical Mechanics

While the previous section emphasized the difficulty of developing a generalizable framework for studying NESM, recently, large deviation theory (LDT) has emerged as a strong candidate. Initially developed as a mathematical theory for quantifying the rarity of events in stochastic systems [46–51], it later became clear that it was a natural and rigorous framework for studying equilibrium statistical mechanics [52–56], providing, for example, a statistical justification for the use of the Legendre transform. In recent years, because of the generality of the mathematics, it has been extended to provide a novel approach for the theoretical treatment of NESM [3–5, 57] and its use over the past decade has led to improved understanding of various nonequilibrium phenomena [6–16].

In this section, we will highlight some of the major results of LDT, particularly with regards to its application to NESM. We begin by providing a simple introductory example that helps establish some of the fundamental concepts used later. Afterwards, we demonstrate how the central results from LDT underpin key aspects of equilibrium statistical mechanics. This leads naturally into a discussion of how these principles can be extended to nonequilibrium systems. Before finishing this section and because this thesis is centered on the development and use of novel numerical

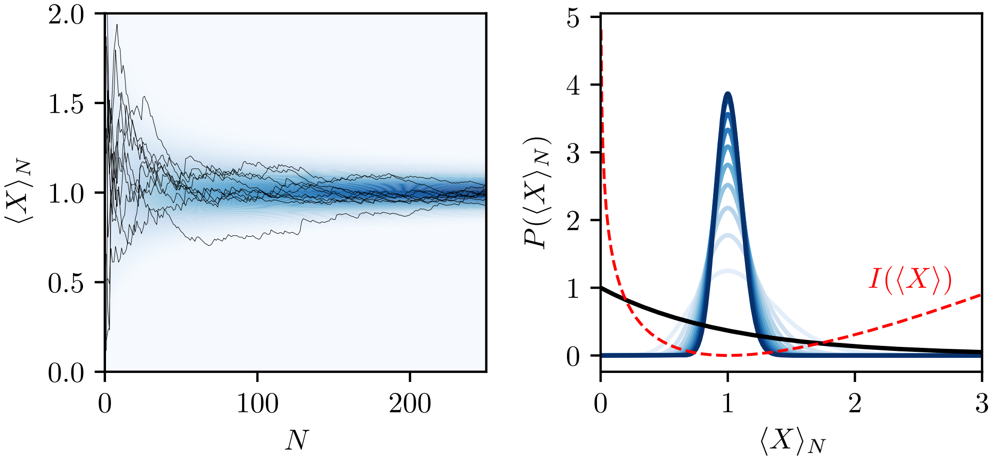

Figure 1.1: The large deviation behavior of the sample mean $\langle X \rangle_N$ of $N$ samples taken from the exponential distribution. (left) Lines show a handful of example trajectories, showing the resulting sample mean as a function of $N$. The shaded background shows the sample mean probability distribution as computed from Equation 1.17. (right) The black line shows the exponential distribution and the red dashed line shows the rate function. Blue lines show the resulting probability distribution as a function of $N$, with the lightest line corresponding to $N = 10$ up to the darkest line at $N = 100$.

approaches for calculations within LDT, we provide an overview of standard modern numerical techniques. Finally, to illustrate the importance of LDT, we highlight a few recent successful applications in NESM.

Before continuing, we should note that the contents of this section are intended only to provide an introduction to this field at the depth necessary for understanding subsequent chapters of this thesis. For those interested in further coverage, we point to a handful of reviews on introductory LDT [58], LDT and equilibrium statistical mechanics [3], and LDT and NESM [4, 5].

### Fundamentals of LDT

As a means of introducing the fundamental principles of LDT, we can follow [3] and consider the statistics of the sample mean

$$\langle X \rangle_N = \frac{1}{N} \sum_{i=1}^{N} X_i, \tag{1.14}$$

where we take $N$ samples of a random variable $X_i$ distributed according to an exponential distribution function

$$P(X_i = x) = \frac{1}{\mu} e^{-x/\mu}, \tag{1.15}$$

where $\mu$ is the sample mean, which we set to $\mu = 1$, and the resulting distribution is shown as the black line in the right plot of Figure 1.1. In the left plot of that figure, the black lines show example trajectories of the sample mean as a function of the number of samples that have been taken. The law of large numbers dictates that as $N \to \infty$, the probability that the sample mean is within a small interval around the random variable's distribution's mean approaches unity. Thus, as would be expected, as the number of samples taken increases, the sample mean for these trajectories converges towards $\mu = 1$. As a slight extension of this, it is also clear that the probability of getting a sample mean far from $\mu$ becomes increasingly unlikely as more samples are taken. Accordingly, it can be shown that in the limit of a large number of samples, the probability distribution of sample means has the form

$$P(\langle X \rangle) \propto e^{-N\phi(\langle X \rangle)}, \tag{1.16}$$

where

$$\phi(\langle X \rangle) = \frac{\langle X \rangle}{\mu} - \ln\left(\frac{\langle X \rangle}{\mu}\right) - 1, \tag{1.17}$$

which is shown by the red line in the right plot of Figure 1.1 and is called the rate function as it specifies the exponential rate at which we expect to see arbitrarily large fluctuations in the sample mean.

The properties of this rate function are of central importance. Because $\phi(\mu) = 0$ is the only zero, the probability of observing a sample mean away from $\mu$ becomes exponentially unlikely as more samples are taken. This is illustrated in both the left and right plots of Figure 1.1. In the left figure, the blue shading in the background shows the probability distribution of the sample means as a function of $N$, becoming increasingly concentrated near $\langle X \rangle_N = \mu$. Slices of this function at increasing values of $N \in [10, 100]$ are shown in the blue lines in the right plot where we see the probability density function sharpening around $\mu$.

A second important property of the rate function is that its first derivative with respect to $\langle X \rangle$ is 0 at $\mu$. Because of this and the previously mentioned property of $\phi$, when we do a second-order Taylor series approximation of $\phi$ around $\mu$

$$\phi(\langle X \rangle) \approx \phi(\mu) + \left.\frac{\partial \phi}{\partial \langle X \rangle}\right|_{\langle X \rangle = \mu} (\langle X \rangle - \mu) + \frac{1}{2} \left.\frac{\partial^2 \phi}{\partial \langle X \rangle^2}\right|_{\langle X \rangle = \mu} (\langle X \rangle - \mu)^2, \tag{1.18}$$

the first two terms vanish, giving

$$\phi(\langle X\rangle) \approx \frac{1}{2}\frac{\partial^2\phi}{\partial\langle X\rangle^2}\bigg|_{\langle X\rangle=\mu}(\langle X\rangle-\mu)^2. \tag{1.19}$$

With this result, we have now recovered the central limit theorem, meaning that approximating the rate function via a second-order Taylor series expansion results in Gaussian fluctuations of the sample mean around $\mu$. This approximation is then only valid for estimating sample mean probabilities near $\mu$, meaning it represents only small deviations from expected behavior accurately. With this, we can now realize that the moniker of *large deviation* refers to the incorporation of non-Gaussian fluctuations arbitrarily far from typical behavior.

We say that any system whose probability distribution has the form of Equation 1.16 obeys a large deviation principle, which can be expressed more precisely as

$$\lim_{N\to\infty}\left(-\frac{1}{N}\ln\left(P(\mathcal{O})\right)\right) = \phi(\mathcal{O}), \tag{1.20}$$

where we are using $\mathcal{O}$ to represent an arbitrary function or observable. By computing $\phi(\mathcal{O})$, we can access, in the large $N$ limit, the probability of arbitrarily large fluctuations, though, as can be imagined, this computation becomes difficult for even moderately complex random processes.

A complementary approach is frequently employed, to provide a second route to accessing large fluctuations, wherein a Laplace transform of Equation 1.20 is taken to give

$$\psi(\lambda) = \lim_{N\to\infty}\left(\frac{1}{N}\ln\left\langle e^{N\lambda\mathcal{O}}\right\rangle\right), \tag{1.21}$$

with $\lambda$ being the variable introduced in the Laplace transform. The resulting function, $\psi(\lambda)$ is the scaled cumulant generating function of the observable $\mathcal{O}$, meaning that fluctuations of $\mathcal{O}$ are encoded in $\psi(\lambda)$, i.e.

$$\begin{aligned} \frac{\partial\psi(\lambda)}{\partial\lambda}\bigg|_{\lambda=0} &= \langle\mathcal{O}\rangle \\ \frac{\partial^2\psi(\lambda)}{\partial\lambda^2}\bigg|_{\lambda=0} &= \langle\mathcal{O}\rangle - \langle\mathcal{O}\rangle^2 \\ &\vdots \end{aligned} \tag{1.22}$$

A remarkable result known as the Gärtner-Ellis Theorem relates the scaled cumulant generating function $\psi$ and the rate function $\phi$ via a Legendre-Fenchel transform

$$\phi(\mathcal{O}) = \max_{\lambda}\left(\lambda\mathcal{O} - \psi(\lambda)\right), \tag{1.23}$$

a result of a saddle point approximation and valid so long as $\psi(\lambda)$ and $\phi(\mathcal{O})$ are differentiable. If $\psi(\lambda)$ is convex, then this relation further simplifies to the standard Legendre transform

$$\phi(\mathcal{O}) = \lambda^* \mathcal{O} - \psi(\lambda^*), \tag{1.24}$$

where $\lambda^*$ is the point where $\partial_\lambda \psi(\lambda^*) = \mathcal{O}$.

With the large deviation functions, meaning the scaled cumulant generating function and the rate function, and their relation defined, we now are equipped with two important approaches to quantify fluctuations of random variables near and far from typical and can consider how they arise in statistical mechanics.

### LDT and Equilibrium Statistical Mechanics

To illustrate the connection between LDT and equilibrium statistical mechanics, we can recall the microcanonical ensemble introduced in Section 1.1 and consider the probability of observing the isolated system exhibiting a given energy per site $E_N = E/N$. This can be expressed in a simple sum

$$P(E_N) = \sum_{\{C\}} \delta(\mathcal{H}(C)/N - E_N) P(C), \tag{1.25}$$

where $\{C\}$ is the set of all possible configurations, $\delta$ is the Kronecker delta function, and, from Equation 1.1, $P(C) = \Omega^{-1}(N, V, E)$. Rearranging Equation 1.2, we see that

$$P(E_N) = e^{-S}, \tag{1.26}$$

and working with the intensive entropy, $s = S/N$, we recognize a large deviation principle

$$s(E) = \lim_{N \to \infty} -\frac{1}{N} \ln\left((P(E_N)\right). \tag{1.27}$$

similar to Equation 1.20 and we recognize the entropy to be a rate function of the energy. Because of this relation, the rate function is also frequently referred to as a generalized entropy.

Now, we can explore the meaning of the scaled cumulant generating function in this context. First, by analyzing the definition of the canonical partition function in Equation 1.6, $\left\langle e^{-\beta\mathcal{H}(C)}\right\rangle$ we can recognize this term has the same form as the generating function of $\mathcal{O}$ in Equation 1.32, $\left\langle e^{N\lambda\mathcal{O}}\right\rangle$, with $\lambda = \beta$ and $\mathcal{O} = \frac{1}{N}\mathcal{H}(C)$. By putting this into the definition of the scaled cumulant generating function

$$\psi(\lambda) = \lim_{N \to \infty} \frac{1}{N} \ln Z\left(N, V, T\right), \tag{1.28}$$

we immediately see a similarity to the definition of the free energy in Equation 1.7, and recognize that the intensive free energy is the scaled cumulant generating function of the intensive energy. Similar to the colloquial nomenclature of the rate function, the scaled cumulant generating function is frequently called the generalized free energy.

The recognition of the entropy and free energy of equilibrium statistical mechanics as large deviation functions allows tools from LDT to be employed in that setting. In particular, we see the origin of the Legendre transform as being purely statistical and understand better the limits of its applicability. Furthermore, this understanding can be used to rigorously support the use of variational principles, such as the maximum entropy principle, which help identify equilibrium states [3].

### LDT and Nonequilibrium Statistical Mechanics

Since the central findings of LDT were tied to equilibrium statistical mechanics, significant work has been put into extending these principles to nonequilibrium systems [4, 5]. In this section, we will show that LDT provides a flexible and rigorous framework for applications in nonequilibrium statistical mechanics, though it does not directly overcome many of the inherent complications. For example, because nonequilibrium thermodynamics is unsettled, we do not have sufficient prescience to select a handful of macroscopic functions of the microscopic degrees of freedom that fully specify the state of the macroscopic system, though we can use intuition to select variables that provide important insight. Furthermore, there remains the diverse landscape of nonequilibrium systems outlined in Section 1.1, so a current effort is to extend the LDT approach to many of these classes. For this reason, in the remainder of this thesis, we will limit our discussion to the steady-state behavior of classical stochastic systems arbitrarily far from equilibrium, with the theory agnostic to whether the dynamics generated are for a microscopic or mesoscopic system, which is the setting where the principles of LDT have been most established.

In equilibrium systems, defining the system's Hamiltonian also defines the Boltzmann weights that directly provide microstate probabilities. Unfortunately, two difficulties arise when extending this approach to NESM. First, because of the dynamical nature of nonequilibrium systems, single configurations representing a snapshot of the system at a single time do not provide enough information to understand dynamical properties such as current, activity, or entropy generation. Thus it is requisite that we consider the statistics of entire trajectories instead of individual configurations. This greatly increases the complexity of the problem, however, because while the number of potential configurations grows exponentially with system size and the number of particles, the number of trajectories grows exponentially with both the system size and the simulation time. This makes performing computations at long times and large system sizes quickly become prohibitively expensive.

Second, once a trajectory has been defined, there are no *a priori* rules for the probability of that trajectory. Instead, the equation that generates the stochastic dynamics of the system of interest can be used to compute the probability of a trajectory. While the stochastic dynamics of the systems we are interested in can frequently be modeled by Langevin or Fokker-Planck equations, the master equation approach is the most general of these and we will use it here. The discrete time form of the master equation was introduced in Equation 1.10. We will work with the continuous time version now, writing it more compactly as

$$\frac{\partial}{\partial t}|P_t\rangle = \mathcal{W}|P_t\rangle, \tag{1.29}$$

where $|P_t\rangle$ is a vector representing all configurational probabilities at time $t$ and $\mathcal{W}$ is the transition matrix, whose off-diagonal elements are the transition rates between configurations $\mathcal{W}_{C,C'} = w_{C'\to C}$ and diagonal entries are the so-called escape rates $\mathcal{W}_{C,C} = r_C = \sum_{C'\neq C} w_{C\to C'}$. We define a trajectory in continuous time as $\mathscr{C}(t_N) = \{C_0(0), C_1(t_1), C_2(t_2), \ldots, C_N(t_N)\}$, meaning the trajectory starts in configuration $C_0$ at time $t = 0$, then transitions to $C_1$ at time $t = t_1$, and so forth, with time step size $t_{i+1} - t_i$ obeying a Poisson distribution. The probability of an arbitrary trajectory can now be determined using the transition matrix $\mathcal{W}$ and the initial configurational probability distribution $|P_0\rangle$.

At this point, we can introduce the idea of a dynamical microcanonical ensemble of trajectories. If we are interested in measuring the probability of a trajectory exhibiting a dynamical observable $\mathscr{O}_{\text{Dyn.}}$, which can generally be an arbitrary function of the full trajectory, we define the dynamical microcanonical ensemble as the set of all trajectories of length $t_N$ with $\mathscr{O}_{\text{Dyn.}}(\mathscr{C}_{t_N}) = \mathscr{O}_{\text{Dyn.}}$, where $\mathscr{O}_{\text{Dyn.}}(\mathscr{C}_{t_N})$ is the measurement of the observable's value for a given trajectory. Having now defined both the probability of a trajectory and the dynamical microcanonical ensemble, we can determine the microcanonical partition function as the probability of observing a given value of the observable $\mathscr{O}_{\text{Dyn.}}$, namely

$$\Omega_{\text{Dyn.}} = \sum_{\{\mathscr{C}\}} P(\mathscr{C})\, \delta\left(\mathscr{O}_{\text{Dyn.}}(\mathscr{C}_{t_N}) - \mathscr{O}_{\text{Dyn.}}\right), \tag{1.30}$$

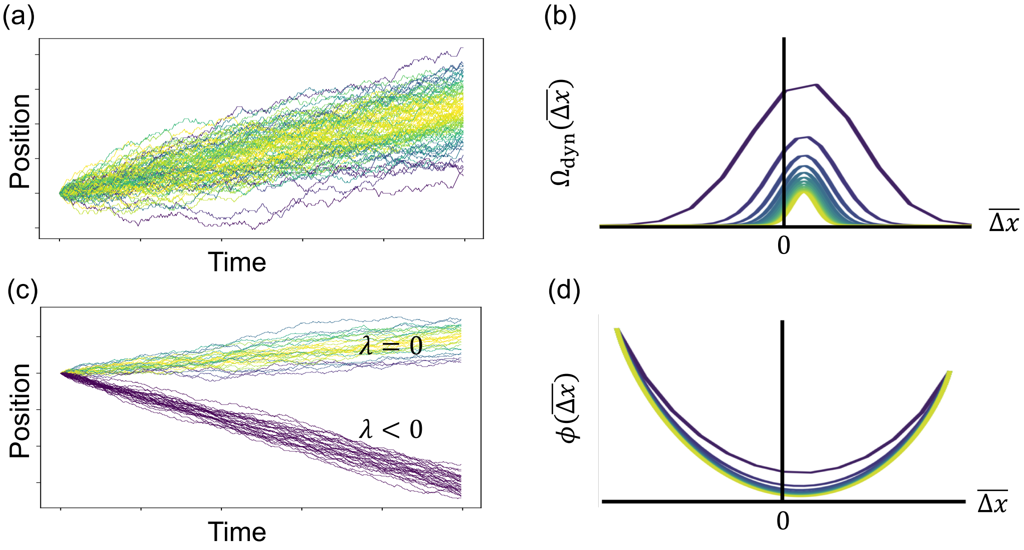

Figure 1.2: Results showing the large deviation behavior of a biased Brownian walker on a lattice. (a) Many example trajectories colored according to the probability of the resulting final displacement of the particle with yellow (blue) lines being the most (least) probable. (b) The microcanonical partition function, or probability distribution of time averaged displacements $\bar{\Delta x}$, with line color indicating the number of time steps from $N = 5$ (blue) to $N = 250$ (yellow). (c) A comparison between trajectories carried out with no bias ($\lambda = 0$) and a bias towards negative displacements ($\lambda < 0$), with the trajectory color again corresponding to the final displacement probability. (d) The rate function versus the time-averaged displacements, with line color again indicating the number of time steps.

where the sum represents a summation over all possible trajectories and $\delta$ is the Kronecker delta function. By selecting an observable that scales appropriately with time, for many systems, $\Omega_{\text{Dyn.}}$ obeys a large deviation principle

$$\phi(\mathscr{O}_{\text{Dyn.}}) = \lim_{t\to\infty} \frac{1}{t} \ln \Omega_{\text{Dyn.}} \left(\mathscr{O}_{\text{Dyn.}}\right), \tag{1.31}$$

where $\phi$ is the dynamical rate function and gives the logarithmic probability of fluctuations in the observable.

To illustrate this, we can consider the behavior of a biased Brownian particle on a lattice evolving discretely in time. At each time step, the particle hops upwards with probability 0.3, downwards with probability 0.2, and remains at the same site with probability 0.5. In Figure 1.2 (a), we show a number of these trajectories, traced as the particle position as a function of time step. We select the time-averaged displacement $\bar{\Delta x}$ as the variable of interest and color the trajectories according to the probability of their final displacement, which corresponds to the microcanonical partition function in the long time limit. The adherence to a large deviation principle is emphasized if we analytically compute $\Omega_{\text{Dyn.}}(\bar{\Delta x})$ as a function of the number of time steps, which is shown in Figure 1.2 (b) with each line showing the distribution for a different time, increasing from blue to yellow. The corresponding rate function $\phi(\bar{\Delta x})$ can be computed from this as well and is shown in Figure 1.2 (d). Here, we see that large fluctuations in the intensive displacement become increasingly unlikely with time. For example, at short times, it is somewhat likely for particles to be below the starting point, but as time proceeds, this becomes exponentially more unlikely.

Mimicking what is done for the equilibrium system, we can introduce the scaled cumulant generating function for our dynamical observable $\mathcal{O}_{\text{Dyn.}}$

$$\psi(\lambda) = \lim_{t\to\infty} \left( \frac{1}{t} \ln \left\langle e^{\lambda \mathcal{O}_{\text{Dyn.}}} \right\rangle \right), \tag{1.32}$$

whose derivatives encode the observable's fluctuations and, under the conditions listed previously, is related to the dynamical rate function via a Legendre-Fenchel transform. While in the equilibrium case the variable introduced during the Laplace transform is physically understood to be the inverse temperature, in the nonequilibrium case there is no such known physical interpretation of this parameter $\lambda$. We can also apply the Laplace transform used to define the scaled cumulant generating function to the master equation, which generates typical trajectories associated with arbitrarily large fluctuations

$$\frac{\partial}{\partial t} |P_t^\lambda\rangle = \boldsymbol{\mathcal{W}}^\lambda |P_t^\lambda\rangle, \tag{1.33}$$

with the new generator $\boldsymbol{\mathcal{W}}^\lambda$ being *tilted*, meaning it is statistically biased to create trajectories with atypical average values of $\mathcal{O}$, causing it to no longer preserve probability. To illustrate this, in Figure 1.2 (c), we show for the biased Brownian walker the type of trajectories generated for the typical dynamics of the master equation, labeled $\lambda = 0$, and the rare fluctuations associated with a specific value of $\lambda$, at $\lambda < 0$. The tilted generator enables affordable generation of arbitrarily rare fluctuations, whereas the cost of generating these trajectories using the bare generator would grow exponentially with the distance from typicality.

With these fundamental large deviation functions and their relation defined, we have the most essential aspects of LDT. In summary, this approach allows us to select, with a few conditions, a macroscopic dynamical observable and characterize its behavior entirely, including arbitrarily large fluctuations from typicality. To briefly illustrate the importance of this approach, in the next subsection we will highlight a few successful applications to interesting physical systems and in the next, we will discuss how these large deviation functions can be computed in practice. First, it is useful to point out a couple of interesting complexities when using LDT for NESM.

First, in the previous section on equilibrium statistical mechanics, the large deviation principles were all taken in the limit of large $N$. Here, they have all been in the limit of large $t$, though it is common to take the limit of large $N$ afterward. It is worth noting that here these limits are not commutable, so different results will arise from taking the large $N$ limit first.

Second, in equilibrium systems, we frequently see singularities arising in the free energy as a function of $\beta$, indicative of a phase transition. This behavior is also seen in the generalized entropy and is commonly referred to as a dynamical phase transition. Special consideration must be taken here because $\lambda$ is not a controllable physical parameter like $\beta$ in equilibrium systems. Instead, a phase transition as a function of $\lambda$ indicates that there are multiple classes of trajectories that are possible. If the dynamical phase transition is near typicality, i.e. near $\lambda = 0$, then it is likely that the system will seemingly stochastically transition between the behavior typical of one phase to the other. This phenomenon is known as dynamical heterogeneity and is commonly reported in many physical systems [7, 59, 60].

### Recent successes in LDT and NESM

To demonstrate the ubiquitous applicability of LDT in NESM, we will briefly review three domains where its use has led to important physical insights or numerical improvements. These highlighted application areas are

- Improved numerical techniques for near-equilibrium systems.
- Providing novel insight into physical phenomena.
- Nonequilibrium system design.

As mentioned at the end of Section 1.1, for systems near-equilibrium, responses are linear with respect to small perturbations, i.e. $J = LX$, where $J$ is the current, $X$ is the applied force that can be in the form of a temperature or particle gradient, and $L$ is called the transport coefficient, which specifies how the system responds to the perturbation. To quantify these responses, Green-Kubo relationships [26, 27] tie the transport coefficients to time correlation functions

$$L \propto \int_0^\infty \langle \mathcal{O}(0)\mathcal{O}(t)\rangle dt, \tag{1.34}$$

where $\mathcal{O}$ is a flux-like observable. While this relation provides a simple method for computing transport coefficients, the calculation costs are related to the correlation time, thus this approach becomes difficult for systems where time correlations fade slowly. While in the past, there have been attempts to overcome this by inducing small currents via boundary conditions or changing the equations of motion, these are usually highly tailored towards a specific system [61–63].

LDT has recently been shown to provide an efficient alternative route [64, 65]. In this approach, we bias the statistical ensemble of possible trajectories towards those that generate fluctuations in the time-average of the flux-like observable, i.e.

$$\bar{\mathcal{O}} = \frac{1}{t}\int_0^t \mathcal{O} dt, \tag{1.35}$$

where the parameter $\lambda$ is the statistical biasing parameter introduced earlier in this section. For small values of $\lambda$, the scaled cumulant generating function $\psi$ takes on a parabolic form

$$\psi(\lambda) = L\lambda^2 + O(\lambda^4), \tag{1.36}$$

that provides a simple relationship between the transport coefficient and large deviation functions.

When this method was introduced [64], it was used to compute the shear viscosity, interfacial friction coefficient, and thermal conductivity for molecular systems and, importantly, was shown to be more statistically efficient than direct use of the Green-Kubo relation, in the sense that fewer Monte Carlo walkers were required for the large deviation approach to provide results of the same accuracy. A second development [65] extended this concept to systems whose response to a perturbation is nonlinear and takes the form

$$J = L_1 X + L_2 X^2 + L_3 X^3 + \cdots, \tag{1.37}$$

where $L_1$ is the linear response coefficient discussed previously and higher-order coefficients are nonlinear response coefficients. The central finding there shows that the nonlinear response coefficients are given by derivatives of the scaled cumulant generating function at $\lambda = 0$. These nonlinear factors become important for either systems that inherently exhibit nonlinear response, such as nanoscale systems, or

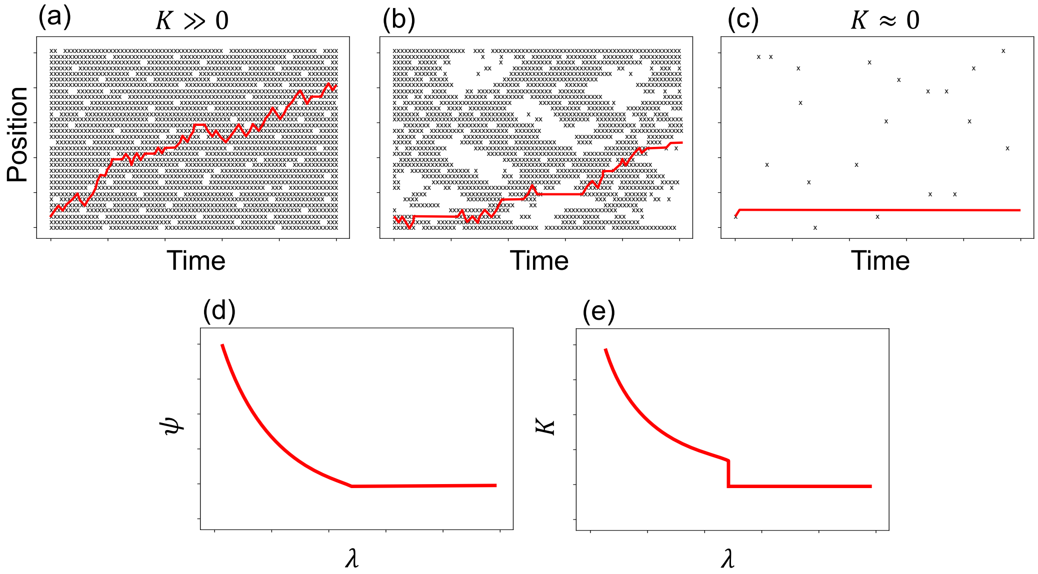

Figure 1.3: Illustration of dynamical heterogeneity exhibited by glassy systems with (a-c) showing example trajectories where each tick within the lattice corresponds to a hopping particle and the red lines trace the trajectory of a single particle. In (a) and (c) the activities $K$ are respectively large and near zero. In (b) the system exhibits dynamical heterogeneity with pockets of inactivity. (d) and (e) show, respectively, sketches of the scaled cumulant generating function and total activity for glassy system as a function of the statistical biasing parameter $\lambda$.

those driven out of the range of linear response and this approach has been used for calculation of these coefficients for models such as a driven Brownian ratchet and a thermal rectifier.

Having shown an example of LDT providing numerical advantage in the study of near-equilibrium systems, we now shift to highlighting how it can provide insight into fundamental phenomena of interest in various physical systems. In recent history, this is where the bulk of work in this field lies, thus providing a large array of possible examples. Some of these applications include crystal growth [13], active matter [10–12], carbon nanotubes [14], weather [66], and basic transport models [4]. Perhaps the most fundamentally important and interesting work, however, has been in the use of LDT in elucidating the behavior of glassy systems [7], which has been an extremely active area of interest for the past decade.

The key controversy within this field concerns whether the glass transition is purely configurational in nature or if there is also a required dynamical component. A central piece of evidence towards believing that the transition is dynamical in nature arises from a physical phenomenon known as dynamical heterogeneity, which we illustrate in Figure 1.3, which has become one of the hallmark characteristics of glassy behavior. This phenomenon occurs when dynamical activity, $K$, a measure that tracks the number of changes in the local state space of a system, becomes localized, meaning that there are some spatial regions where particles are moving, or active, and others where they are unable to move, often due to confinement, and are inactive. As exemplary trajectories, Figures 1.3 (a), (b), and (c) show constructed trajectories for a one-dimensional lattice of particles that hop between nearest-neighbor sites. Each mark in the plot represents a local configuration change, i.e. a particle hopping to a neighboring site, and the red lines track the trajectories of a single particle. While plots (a) and (c) exhibit globally active and inactive trajectories, plot (b) shows the regionally localized activity and inactivity that appear as *bubbles in space-time* [67], indicative of dynamical heterogeneity. From the perspective of a single particle, this behavior translates into fluctuations into and out of dynamical behavior, meaning that the particle experiences a random walk between confinement by other particles, leading to time-local inactive trajectories, and freedom to diffuse in time-local active trajectories.

LDT provides an important role in detecting and understanding dynamical heterogeneity in a wide array of glassy systems from simplified lattice models [68] to molecular models [69]. If the statistical biasing introduced via $\lambda$ is done against the total activity of the system, as sketched in Figures 1.3 (d) and (e), the scaled cumulant generating function exhibits a kink at $\lambda = 0$, which is indicative of a first-order dynamical phase transition, here from an active ($K \gg 0$) to an inactive ($K \approx 0$) phase. Because the typical dynamics of the system occur at $\lambda = 0$, this type of dynamical phase transition indicates dynamical heterogeneity will occur as the system's typical behavior does a random walk between the two types of behaviors.

As a final and particularly exciting example of success using LDT in NESM, we can consider recent work in the design of far from equilibrium molecular systems. Here, system design refers to our desire to determine microscopic system parameters, with broad examples including applied fields and intermolecular interactions, that will cause a macroscopic nonequilibrium system to behave as desired. While this has become a standard use of equilibrium statistical mechanics in various contexts, the extension to the nonequilibrium regime has been hampered by the lack of appropriate tools.

Because LDT deals with the occurrence of rare events, there are examples of rare

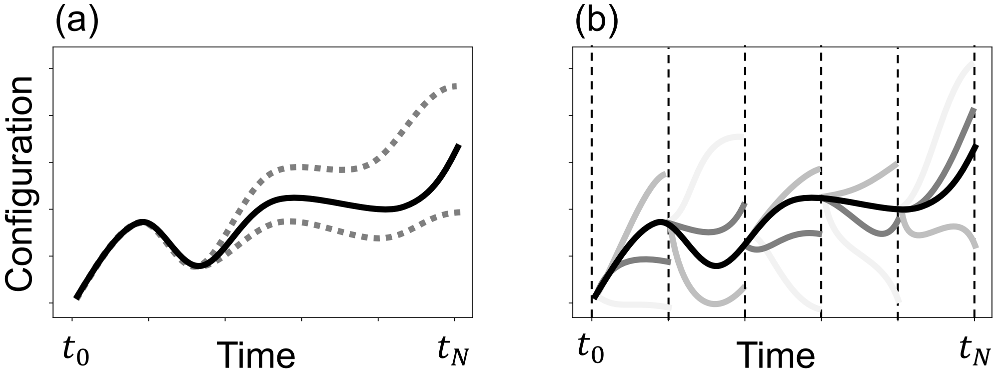

Figure 1.4: Sketches of common Monte Carlo simulation methods used for computing large deviation functions. (a) illustrates transition path sampling with the dark line showing the proposed trajectory and the gray dashed line representing proposed alterations. (b) illustrates diffusion Monte Carlo with the shade of the path determined by the weight factors accumulated for each trajectory.

events that we would like to cause and others that should be avoided. A recent paper took this approach to determine system parameters and design principles to cause the typical state of a set of interacting particles to self assemble into various configurations [9], which are unlikely in equilibrium. The statistical biasing was done using an indicator function, unique for each target cluster, and though we will not present the details here, a variational principle was used to determine the optimal system parameters for 21 distinct clusters. This is one of a couple of recent results where LDT has begun to be used to design far from equilibrium systems [8], which is likely to be a fruitful area of research over the next decade.

### Computing Large Deviation Functions

Having established the fundamentals of LDT and its importance in understanding the complex physics of nonequilibrium systems, we move into a brief discussion of computational approaches for computing large deviation functions. Current computational methods fall into two general classes, Monte Carlo algorithms or eigenproblems, both of which we discuss now. It should also be noted that some macroscopic approaches exist that attempt to incorporate the effects of fluctuations into macroscopic system behaviors using coarse-grained equations [70]. We will not include these in our discussion.

From a practitioner's perspective, Monte Carlo methods for studying rare events provide a natural framework. Here, trajectories are generated alongside a tracked reweighting factor of $e^{-\lambda\mathscr{O}}$, where $\mathscr{O}$ is the dynamical observable of interest [17]. While this allows for the use of standard molecular dynamics techniques, evaluating the probability of exponentially rare events, unfortunately, requires an exponentially large number of sampled trajectories. Applicable Monte Carlo approaches generally fall into two algorithm classes, namely transition path sampling (TPS) [71] and diffusion Monte Carlo (DMC) [72]; the general idea of each is illustrated in Figure 1.4. In TPS, the algorithm is initialized by proposing a full initial trajectory (black line in Figure 1.4(a)), which then stochastically explores space-time, with proposed trajectory updates (gray dashed lines in Figure 1.4(a)), being accepted or denied using a Metropolis rule that incorporates the exponential reweighting factor mentioned above.

DMC is a comparable method, in the sense that it is expected to require approximately the same computational effort as TPS, but with a different approach. In DMC, walkers are initialized at the initial time $t_0$ in a given configuration. The walkers then stochastically evolve in time while we track the reweighting factor. After a short time $\Delta t$, the walkers pause; they are then redistributed onto the trajectories that have acquired the largest weighting factor, meaning, alternatively, that trajectories with low weights are deleted and those with high weights are cloned. This process of evolution and redistribution at intervals of $\Delta t$ continues until the final simulation time is reached.

For completeness, we now describe in detail the DMC algorithm for a system evolving according to Equation 1.29, where we are interested in computing the statistics of an additive, time-local observable $\mathscr{O}\left(\mathscr{C}\left(t_N\right)\right) = \sum_{k=1}^{N} o\left(C_{t_{k-1}}, C_{t_k}\right)$. These types of observables are quite flexible including system activities, where $o\left(C_{t_{k-1}}, C_{t_k}\right) = 1$, which measure the frequency of configuration changes, and currents, where $o\left(C_{t_{k-1}}, C_{t_k}\right) = \pm 1$ depending on whether a particle moves forward or backward. We first consider the joint probability distribution $P\left(C, \mathscr{O}, t\right)$, defined as the probability of a trajectory at time $t$ being in configuration $C$ with its trajectory observable being $\mathscr{O}$. This joint probability evolves according to a constrained version of the master equation, which we write explicitly as

$$\partial_t P\left(C, \mathscr{O}, t\right) = \sum_{C'} w_{C' \to C} P\left(C', \mathscr{O} - o\left(C', C\right), t\right) - r_C P\left(C, \mathscr{O}, t\right), \tag{1.38}$$

where on the right-hand side the first term represents a transition into the given configuration and the second is the probability of remaining in the current configuration. The statistics of $\mathscr{O}$ can be directly accessed by summing this joint probability distribution over all configurations to get the dynamical microcanonical partition function.

Alternatively, the canonical partition function can be accessed by first taking the Laplace transform of the joint probability distribution

$$P\left(C, \lambda, t\right) = \sum_{\mathcal{O}} e^{-\lambda \mathcal{O}} P\left(C, \mathcal{O}, t\right), \tag{1.39}$$

giving the tilted master equation

$$\partial_t P\left(C, \lambda, t\right) = \sum_{C'} e^{-\lambda o(C',C)} w_{C' \to C} P\left(C', \lambda, t\right) - r_C P\left(C, \lambda, t\right), \tag{1.40}$$

shown compactly in Equation 1.33. Because of the exponential term, these tilted dynamics no longer conserve probability. We can instead define an analogous continuous-time Markovian process with generator

$$\tilde{\mathcal{W}}_{C,C'} = e^{-\lambda o(C',C)} w_{C' \to C} - r_{C,\lambda} \delta_{C',C} \tag{1.41}$$

with $r_{C,\lambda} = \sum_{C'} w_{C \to C'}$. Using these tilted, probability conserving stochastic dynamics, we can rewrite Equation 1.40 as

$$\partial_t P\left(C, \lambda, t\right) = \sum_{C'} \tilde{\mathcal{W}}_{C,C'} P\left(C', \lambda, t\right) + \left(r_{C,\lambda} - r_C\right) P\left(C, \lambda, t\right). \tag{1.42}$$

The computation of $P\left(C, \lambda, t\right)$ can then be thought of as an evolution according to the tilted, probability conserving generator $\tilde{\mathcal{W}}_{C,C'}$, where after each evolutionary step trajectories are cloned or destroyed according rules defined by the second term.

Unfortunately, the cost for calculating the scaled cumulant generating function grows exponentially as a function of $\lambda$, meaning as the sampled trajectories become exponentially rarer, the cost grows exponentially. To overcome this cost, two main techniques have been employed. First, when DMC is used, importance sampling schemes apply additional weights to walkers that are behaving in the desired way, causing them to be more likely to be cloned and exhibit the behavior of interest [17, 72]. Second, control forces can be introduced, which can be thought of as additional auxiliary forces that push the trajectories into desired behaviors [18, 23]. The influence of these control forces can be tracked and removed later, which allows users to figuratively meddle in the dynamics of the system while tracking and erasing the associated effects of the meddling. Together, these methods have supported much of the recent progress in this field.

An alternative approach to determining the large deviation functions for a given observable in a system of interest relies on the properties of the tilted generator to convert this into an eigenproblem. To illustrate these properties, we integrate Equation 1.33 to get

$$|P_t^\lambda\rangle = e^{\mathcal{W}^\lambda t}|P_0^\lambda\rangle. \tag{1.43}$$

By definition, a sum over these probabilities gives the ensemble average of $e^{-\lambda\mathcal{O}}$, meaning

$$\begin{aligned}\langle e^{-\lambda\mathcal{O}}\rangle &= \langle C|P_t^\lambda\rangle \\ &= \langle C|e^{\mathcal{W}^\lambda t}|P_0^\lambda\rangle,\end{aligned} \tag{1.44}$$

where $\langle C|$ is a vector of ones here that allows us to sum over the contributions from all configurations. By doing a spectral decomposition of the exponential term and taking the long time limit, this simplifies to

$$\langle e^{-\lambda\mathcal{O}}\rangle = Ce^{\psi_0(\lambda)t}, \tag{1.45}$$

where $\psi_0$ is the largest eigenvalue of $\mathcal{W}^\lambda$ and $C$ is a constant that arises from the summation over configurations. By comparing this equation with the large deviation principle obeyed by the scaled cumulant generating function, we see immediately that the largest eigenvalue of the tilted generator is equal to the scaled cumulant generating function, i.e. $\psi_0(\lambda) = \psi(\lambda)$.

With the eigenproblem defined, common algorithms for solving such problems can be adapted here, with the caveat that the tilted generator is non-Hermitian. The most obvious approach is exact diagonalization, which provides exact results for small systems but becomes overly expensive for even moderate sizes. In a similar vein, there are analytical solutions, often employing basis set techniques, that have been used frequently [73–75]. Unfortunately, these exact methods cannot be easily extended to study arbitrary states or complex molecular systems. The tensor network algorithms discussed in this thesis also belong to this eigenproblem category and provide systematically improvable methods that frequently scale linearly with system size. Finally, there are a handful of variational methods that compute the scaled cumulant generating function directly using properties of the tilted generator, but without an eigenproblem. Some algorithms used there include variational Monte Carlo [18] and machine learning techniques such as reinforcement learning [19, 22].

## 1.4 Introduction to Tensor Networks

With the basics of LDT introduced in the previous section, our focus now shifts to the introduction of the numerical ansätz and methods that are adapted in this

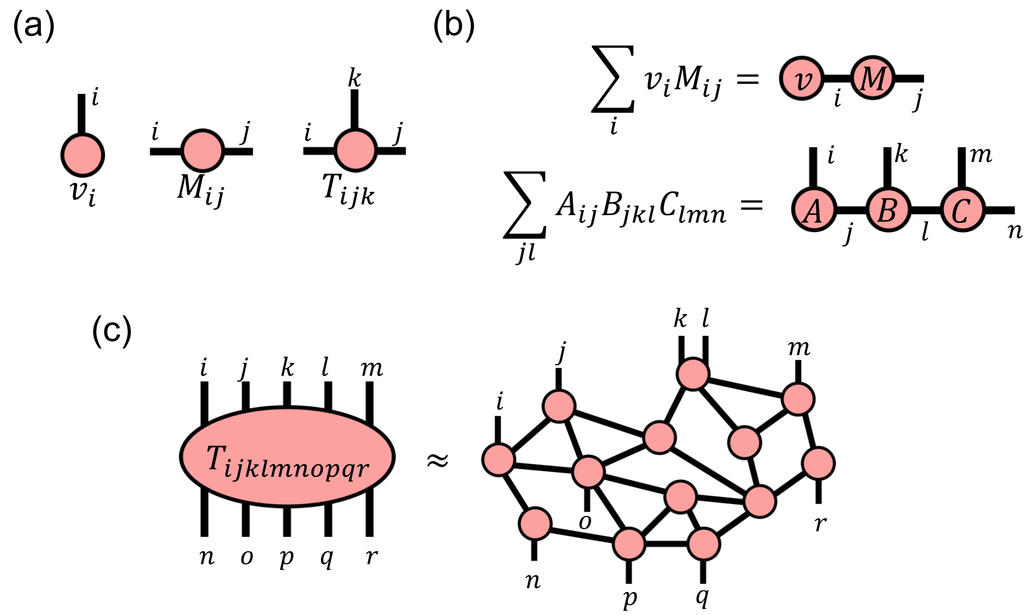

Figure 1.5: An introduction to tensor networks and tensor network diagrams with (a) showing how tensors, here a vector, matrix, and rank-3 tensor, are represented in tensor network diagrams. (b) shows how tensor contractions are represented in the diagrams. (c) illustrates how a large rank tensor can be transformed into several connected lower-rank tensors.

thesis to compute large deviation functions. We will begin in the next subsection with a basic introduction to tensor networks (TNs) by defining them, providing an overview of their historical development and use, and providing justification for their observed success. Because the algorithms are not introduced in detail in the following chapters, the following subsection here provides details of standard tensor network algorithms in both one and two dimensions. Finally, because Monte Carlo methods are standard in statistical mechanics, we close this introduction by discussing reasons why we suggest that TNs become a common tool in numerical studies of nonequilibrium statistical mechanics.

### Basics of Tensor Networks

To provide an intuitive representation of tensor networks, we begin by introducing tensor network diagrams in Figures 1.5(a) and (b). In (a), we show how a tensor is represented as an object, here circles, with legs that each represent an index of that tensor. Using these diagrams, Figure 1.5(a) shows a rank-1 tensor $v_i$, i.e. a vector, as an object with a single leg, a rank-2 tensor $M_{ij}$, i.e. a matrix, as an object with two legs, and a rank-3 tensor $T_{ijk}$ as an object with three legs. In Figure 1.5(b),

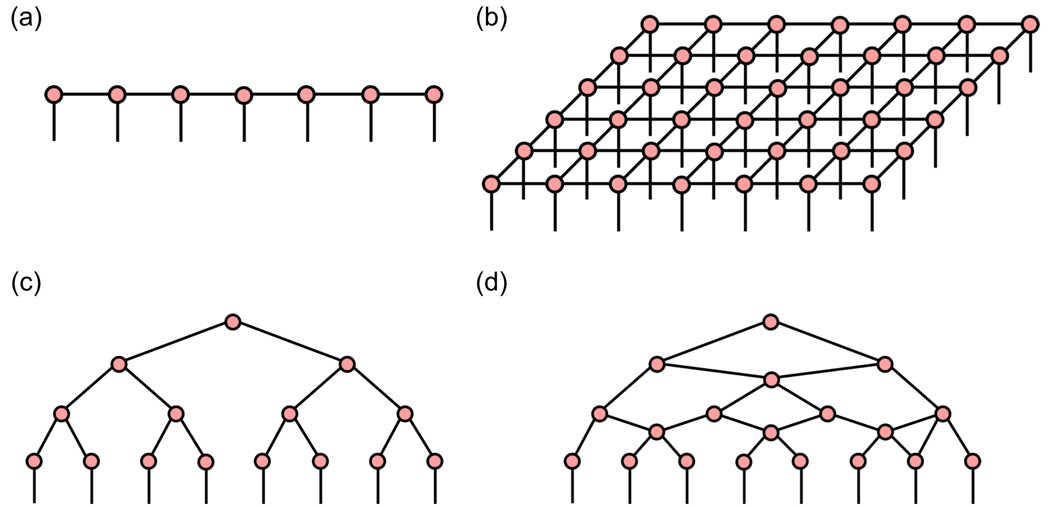

Figure 1.6: Diagrams of common tensor network classes, showing the (a) matrix product state, (b) projected entangled pair state, (c) tree tensor network, and (d) multiscale entanglement renormalization ansatz.

we show how tensor contractions are represented by objects with shared legs, often called auxiliary bonds. As examples, the top diagram shows a vector-matrix product whose result gives another vector, i.e $(vM)_j = \sum_i v_i M_{ij}$, while the bottom figure shows a more complex contraction between three tensors resulting in a rank-4 tensor.

The central idea of a tensor network is illustrated in Figure 1.5(c). Here, a large, rank-10 tensor is represented, either exactly or approximately, as a network of smaller tensors. This decomposition of the tensor into a TN is only one possibility and a single tensor can be represented by infinitely many different TNs.

Over the past two decades as TNs have become a commonly employed numerical tool, several standard tensor networks have emerged as particularly useful in representing physical systems. A few of the most popular TNs are shown in Figure 1.6.

The matrix product state (MPS), shown in Figure 1.6(a) and also called the tensor train ansatz, is a one-dimensional tensor network with origins in the calculation of the ground state of quantum lattice models. Its popularity emerged subtly as the underpinning ansatz for the extremely successful density matrix renormalization group (DMRG) algorithm [76], where the eigenvalues of the density matrix of a quantum state are used as the metric for renormalization. Since the connection between DMRG and the MPS was established [77, 78], the MPS has become widely studied and used. The MPS has become a common ansatz for diverse applications including quantum lattice models [79], quantum chemistry [80], classical statistical mechanics [81, 82], and machine learning [83–85]. Additionally, algorithm development has furthered its use, allowing extensions to time evolution [86, 87], thermal states [88, 89], infinite systems [90], and continuum systems [91, 92].

In quantum physical applications, the remarkable success of the MPS has led to significant interest in exploring its range of applicability [93], which we will discuss briefly here since it motivates the continued use of TNs in various settings. We begin by considering the ground state of a quantum lattice system with $N$ sites, which is expressed as

$$|\psi\rangle = \sum_{\sigma_1,\ldots,\sigma_N} t_{\sigma_1,\ldots,\sigma_N} |\sigma_1, \ldots, \sigma_N\rangle, \tag{1.46}$$

where $\sigma_i$ specifies the local state space for site $i$ (for example, in a spin-1/2 system, $\sigma_i \in \{\uparrow, \downarrow\}$) of dimension $d$, and $t_{\sigma_1,\ldots,\sigma_N}$ is a rank-$N$ tensor, shown in Figure 1.7(a), that grows exponentially with $N$. To begin to convert this state into the form of an MPS, a singular value decomposition (SVD) is employed to bipartition the state tensor into a left and a right tensor ($U$ and $V$) with a diagonal matrix $S$ of singular values connecting the two

$$t_{\sigma_1,\ldots,\sigma_N} = U_{\sigma_1,\ldots,\sigma_i,a_i} S_{a_i,a_i} V_{a_i,\sigma_i+1,\ldots,\sigma_N}, \tag{1.47}$$

shown in Figure 1.7(b). Mathematically, discarding the smallest singular values provides the most optimal approximation to $t$ in the Frobenius norm. Because the singular values are also equal to the square of the eigenvalues of the density matrix, this matrix provides easy access to the bipartite von Neumann entanglement entropy

$$S_{L|R}(\psi) = \sum_{a_i}^{D} S^2_{a_i,a_i} \log_2 S^2_{a_i,a_i} \tag{1.48}$$

which is a measure of the amount of entanglement between the left and right blocks, where $D$ is the size of the auxiliary bond introduced via the SVD. Consequently, for an approximation to the state of fixed accuracy, the required $D$ is directly related to the entanglement entropy. By applying successive SVDs, we can continue to decompose, either exactly or approximately, the state tensor into an MPS (Figure 1.7(c)). The usefulness of the MPS algorithm is now directly related to the amount of entanglement between sites within the lattice.

The central goal in introducing the MPS ansatz is to avoid the exponential scaling of the representation of the state with lattice size. If a quantum state exhibits arbitrarily large entanglement between bipartitions of the lattice, then $D$ must scale

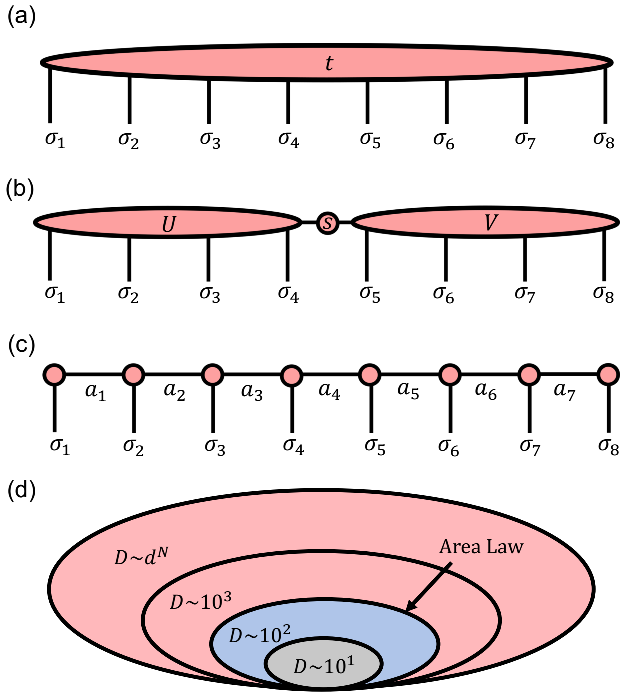

Figure 1.7: An introduction to the MPS ansatz with (a) showing the tensor that represents an arbitrary quantum state. (b) illustrates that an SVD of the original state tensor can bipartition the system into left and right halves connected by singular values. (c) shows the MPS ansatz itself. (d) provides an intuitive illustration of area law behavior where the full possible Hilbert space of the system is the largest circle and reducing $D$, the tensor's auxiliary bond dimension, can limit the effective Hilbert space appropriately.

exponentially with $N$, failing to circumvent the exponential cost. Fortunately, it can be shown that for gapped Hamiltonians acting on a one-dimensional lattice the entanglement entropy obeys an *area law* [93]. Adherence to this law means that the entanglement between two bipartitions of a system scales with the area of the boundary between the two systems, as opposed to *volume law* behavior, where entanglement grows with the size, or volume of the two parts. For one-dimensional systems, this means that the required $D$ for approximating a state that adheres to area law entanglement is constant with $N$. As illustrated in Figure1.7(d), by changing the auxiliary bond dimension, we can modify the effective Hilbert space of solutions. While the full Hilbert space grows exponentially with $N$, adherence to an area law limits the effective space of the state and $D$ can be adjusted to correctly tailor the size of the Hilbert space. The effective Hilbert space can also be modified to be smaller than the Hilbert space needed for the considered state. While it is now known that an MPS provides an optimal approximation to systems that adhere to an area law, there is also significant evidence that it is a powerful ansatz for other systems [94]. In the second chapter of this thesis, we will use an MPS to compute scaled cumulant generating functions for NESM.

Because of the wide success of the MPS ansatz in one dimension, a similarly structured ansatz was proposed for lattice systems in two dimensions, which is called the projected entangled pair state (PEPS) [95] and shown for a square lattice in Figure1.6(b). Extensions to three and higher dimensions are also included within the umbrella of PEPS ansätz as are other lattice configurations, such as hexagonal lattices. While a number of challenges exist in extending the algorithms associated with the MPS ansatz to PEPS, such as a lack of an obvious canonical form [96, 97], inability to exactly contract norms or expectation values [98, 99], and difficulty in representing long-range interaction [100–102], significant algorithmic progress has standardized a handful of algorithms, making PEPS calculations possible for difficult problems [98, 99]. In the third chapter of this thesis, we demonstrate how these standard PEPS algorithms can be applied to compute large deviation functions. Additionally, in the fourth chapter, we introduce and use algorithms seeking to overcome difficulties associated with long-range interactions and continuous degrees of freedom in two-dimensional PEPS calculations.

The other TNs included in Figures 1.6 (c) and (d) are the tree tensor network (TTN) [103] and multiscale entanglement renormalization ansatz (MERA) [104]. Many of the extensions that exist for the MPS and PEPS have also been formulated for the TTN and MERA, including extensions to two dimensions [105]. The TTN is a slight extension of the MPS that, by design, avoids any loops within the network while seeking to allow a more flexible representation of entangled quantum systems. If one looks closely, it becomes clear that the MERA has a structure similar to the TTN, with the addition of rank-4 disentangling tensors, introduced to allow greater entanglement. Because it allows greater entanglement, the MERA is generally used to study critical systems. While we expect that these TNs could be extended to study NESM, we only use the MPS and PEPS here, and thus the rest of our discussion will focus there.

### Standard Algorithms for TNs

As mentioned, because the TN algorithms used in later chapters of this thesis are only introduced superficially, this section provides a brief introduction into standard ground state methods that we adapt for computing large deviation functions in

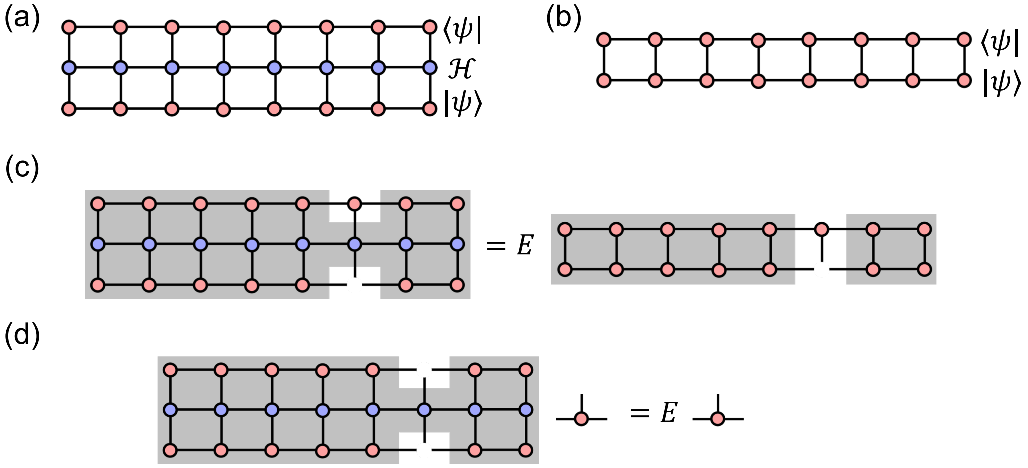

Figure 1.8: Various aspects of the DMRG algorithm from the MPS perspective. (a) and (b) show respectively the energy and norms as a TN, with red tensors being in the MPS and blue tensors being in the MPO. (c) The resulting equation from extremization with respect to a single MPS tensor, showing in gray the local Hamiltonian and norm tensors. (d) Conversion of the equation in (c) to an eigenproblem.

Chapters 2 and 3. We will first introduce the DMRG algorithm for determining the ground state of the MPS, which has been canonicalized and summarized in its entirety elsewhere [79]. We then discuss the time-evolving block decimation (TEBD) algorithm frequently used for PEPS calculations, drawing from various works where different algorithmic details are introduced [98, 99, 106].

To determine the ground state of a quantum system, the variational principle can be employed, which requires us to minimize the system's energy

$$E = \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{1.49}$$

with respect to the parameters in $|\psi\rangle$. To perform a DMRG calculation, the Hamiltonian can be converted into a matrix product operator (MPO), which is analogous to an MPO but with incoming and outgoing bonds. Thus Figures 1.8(a) and (b) represent the numerator and denominator in Equation 1.49, with the MPO and MPS represented respectively by blue and red tensors. To solve the variational problem, we iteratively traverse the lattice, solving variational problems at each lattice site for individual tensors in the MPS until convergence. The local variational problem for an MPS tensor can be converted into an eigenproblem by introducing a Lagrange multiplier to extremize the right-hand side of Equation 1.49 and take the derivative with respect to one of the MPS tensors

$$\langle\psi|\mathcal{H}|\delta\psi\rangle = E\langle\psi|\delta\psi\rangle. \tag{1.50}$$

This is easily understood via the TN diagram equation shown in Figure 1.8(c), where we see the site of interest in the ket has been removed by the derivative. To entirely localize the problem, we can exactly contract a local version of the Hamiltonian, the tensor network for this is highlighted in gray in Figure 1.8(c). In practice, since we are doing optimization sweeps through the lattice sites, we normally build left and right environments iteratively then contract with the local MPO tensor to construct the local Hamiltonian. On the right-hand side, the contraction of the environment tensors can be done exactly, but we exploit properties of the MPS, namely the mixed canonical gauge that will not be described in detail here, to force these environments to be identities. By pulling the MPS tensor from the local site, it becomes clear that the local variational problem is an eigenproblem

$$\mathcal{H}_{\text{loc}}\psi_{\text{loc}} = E\psi_{\text{loc}}, \tag{1.51}$$

where $\mathcal{H}_{\text{loc}}$ is the local Hamiltonian (defined by the TN diagram in Figure 1.8(c)) and $\psi_{\text{loc}}$ is the local MPS tensor.

As a rough sketch of the DMRG algorithm, it proceeds by iteratively sweeping through the lattice sites from the left to the right then back again. Before beginning the sweep to the right, the right environments are built iteratively from site $N$ to site 2. During the right sweep, we iterate over all sites, using an implicit diagonalization routine to find the optimized MPS tensor (the local eigenstate) and the system energy (the local eigenvalue) at each site. The left environment must then be updated by contracting the previous left environment with the newly determined MPS tensors and the MPO tensor. When the right boundary is reached, the procedure reverses, with all of the left boundaries having been computing in the sweep to the right, and the sweeps continue back and forth until the energy converges. While this covers the fundamental algorithm, we note that various techniques have been proposed to help calculations converge well. For the work in Chapter 2, for example, we use state-averaging to compute the ground state and first excited state when the gap becomes very small [107].

With the theoretically sound footing of the DMRG algorithm, the prospects of extending this success to optimizing a two-dimensional PEPS seem promising. We can first discuss a few central problems that prevent the simple extension of the

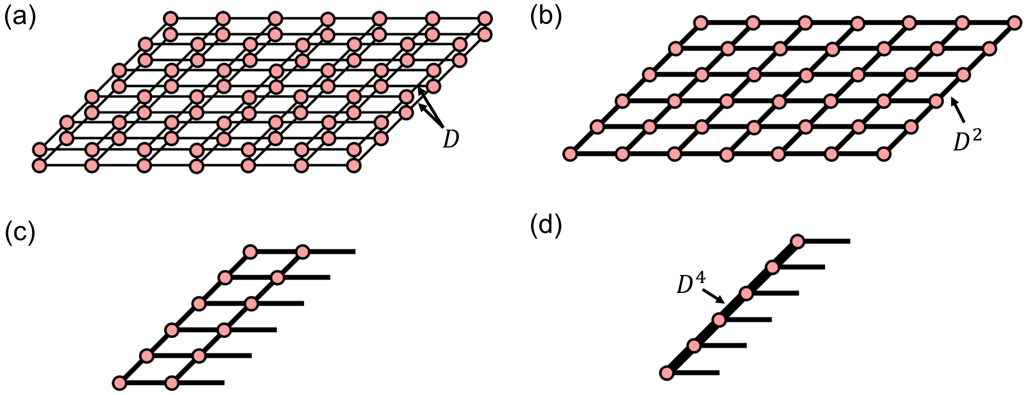

Figure 1.9: Illustration of steps towards contracting a PEPS norm with (a) showing the TN for a PEPS norm and (b) shows a version of this where the physical bond dimensions have been contracted out (requiring the bond dimension to grow). (c) is the first step of the boundary contraction approach where the left-most column is contracted with its neighbor. (d) an illustration of the increased bond dimension that eventually prohibits exact calculation.

DMRG algorithm, which motivates the standard TEBD algorithm which will then be introduced. While this has been the most used PEPS algorithm in recent years, there are many alternatives, including DMRG style optimizations [97, 108] and adaptation of automatic differentiation (AD) methods [109].

The first difficulty can be seen when we compute the norm of a PEPS

$$N = \langle\psi|\psi\rangle, \tag{1.52}$$

the TN for which is shown in Figure 3.1(a), where we assume that bulk tensors within the PEPS are rank-5 of size $dD^4$, where $D$ is the size of the auxiliary bonds connecting PEPS tensors and $d$ is the local state space dimension. To compute the norm, we can first contract over the physical bond that connects the bra and ket PEPS tensors. This results in a rank-4 tensor of size $D^8$, where each left, right, up, and down bond's dimension is $D^2$, as shown in Figure 3.1(b). With this flattened version of the PEPS, a boundary method prescribes starting with the left-most column and sequentially contracting PEPS tensors from left to right. The first step in this process is shown in Figure 3.1(c), where the operation is identical to contract an MPO into an MPS. Unfortunately, the exact contraction of this results in the MPS shown in Figure 3.1(d), where the MPS auxiliary bond dimension has doubled. With the contraction of each additional column, the auxiliary bond dimension continues to grow, making the exact contraction impossible for all but the smallest PEPS.

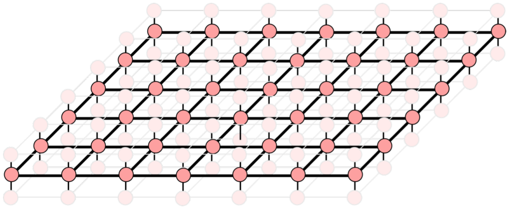

Figure 1.10: The TN for the energy $\langle\psi|H|\psi\rangle$ with the PEPS tensor in the bra and ket lightened to emphasize the structure of the PEPO.

While this example demonstrates that the norm cannot in practice be contracted exactly, this issue extends to energy expectation values, local Hamiltonians (the two-dimensional version of the left-hand side of Figure 1.7(d)), and local norms (the analogy to a local Hamiltonian, but without operator tensors). Fortunately, we can borrow the standard approach for approximately applying an MPO to an MPS to use here. In this approach, the tensors are first contracted together to form the boundary MPS with increased auxiliary bond dimensions. To optimally approximate this MPS with another MPS with maximum auxiliary bond dimension $\chi$ the boundary MPS is then *canonicalized* via a successive application of QR decompositions on the MPS tensors and finally *truncated* by a successive application of SVD decompositions to the canonicalized MPS tensors, with only the $\chi$ largest singular values retained between each set of tensors [98]. While this algorithm is nearly universally used for PEPS contractions, there are some slight variations on this which can make the contraction more accurate for similar numerical cost. These include an approach where the bra and ket are contracted into the boundary MPS separately, lowering the size of the MPS tensors that are being canonicalized and truncated [98], and a related single layer method, where the PEPS norm is flattened into a single layer lattice with rerouting identities introduced to keep the tensors appropriately connected [110], which is used in Chapter 3 of this thesis.

A second difficulty in dealing with PEPS calculations is related to the first and discussed further in Chapter 4. One of the strengths of the MPS ansatz in combination with DMRG that we did not mention previously is the ease of converting many Hamiltonians into MPOs. Many can be represented exactly, some constructed using finite automata [111] while others, such as exponentially decaying interactions, having naturally compact forms. Additionally, for long-range interactions that do not admit an obvious compact form, there has been significant progress in developing general MPO truncation schemes that are now widely available and used [112]. The extension of these approaches to create projected entangled pair operators (PEPOs), the TN of which is shown as the highlighted tensors in Figure 1.10, has been mildly successful, with the introduction of loops making a determination of a compact representation more cumbersome, but frequently possible.

Unfortunately, an additional difficulty arises when we consider how to compute the energy of a state using a PEPO, the TN representation of which is shown as the full TN in Figure 1.10 with the PEPO emphasized. When a boundary contraction approach is attempted with the three-layered TN both the PEPS and PEPO bonds are truncated. As was observed with the approximate norm contraction, this truncation also then affects the accuracy of the local Hamiltonian, which, if used for the local eigenproblem, may cause numerical issues. Indeed, this truncation has previously been shown to significantly affect the accuracy of the energy contractions [102, 113], making it seemingly impractical to use PEPOs for energy optimizations. While some previous work has proposed methods of including long-range operators with varying representations [100, 102, 114] and Chapter 4 of this thesis discusses the further improvement and use of one of these, this has limited standard PEPS algorithms to including only nearest-neighbor (NN) and occasionally next-nearest-neighbor (NNN) interactions.

The final complication that prevents a direct extension of the DMRG algorithm for a PEPS ground state optimization arises from the contraction of a local norm. As illustrated in Figures 1.7(c) and (d), the properties of the MPS, namely the canonicalization previously mentioned, allows for the left and right norm environments around the optimization site to contract away to an identity. This allows us to arrive at the local eigenproblem in Equation 1.51. Unfortunately, while some work has attempted to convert a PEPS into a canonical form [96, 97, 115] and proposals have been made for using *local canonicalization* [99], to make the environment close to an identity, this remains an open question. Until that is solved, the local site update becomes a generalized eigenproblem

$$\mathcal{H}_{\text{loc}}\psi_{\text{loc}} = E\mathcal{N}_{\text{loc}}\psi_{\text{loc}}, \tag{1.53}$$

where $\mathcal{N}_{\text{loc}}$ is the local environment (which is the norm TN in Figure 3.1(a) with the PEPS tensor for the site of interest removed from both the bra and ket).

With the limitations introduced by the mentioned complications associated with optimizing a PEPS, the TEBD algorithm has arisen as the most robust optimization approach. TEBD was initially introduced to perform imaginary time evolution (ITE) on an MPS to arrive at the ground state and, while not necessary for our discussion

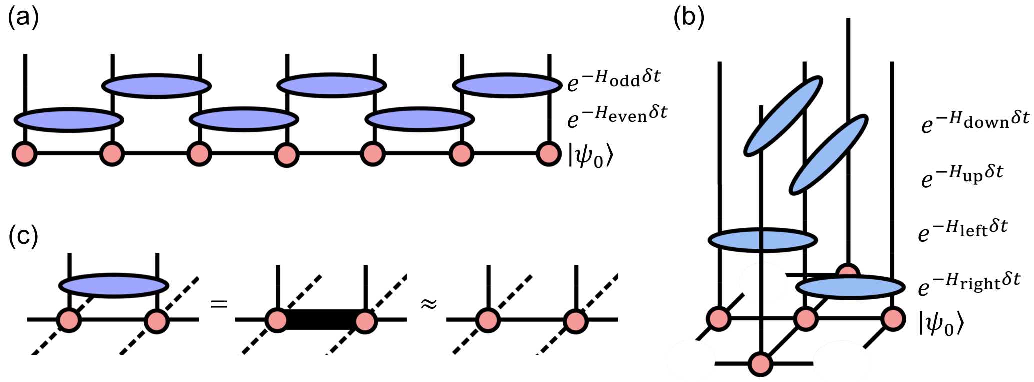

Figure 1.11: TN diagrams related to the TEBD algorithm. (a) shows the fundamental idea using suzuki-trotter decomposed local gates to do an approximate ITE. (b) shows how this idea can be extended to two-dimensional systems where we are required to do left, right, down, and up gate applications. (c) Illustrates that for a MPS or PEPS, the application of a gate increases the bond between those tensors. This bond can then be approximated using various methods.

here, it can also be extended to model time-dependent behavior and compute thermal states [116, 117]. The fundamental idea behind ITE is that the ground state of a system can be expressed as the long-time limit of imaginary time evolution via the Schrödinger equation

$$|\psi_{\text{GS}}\rangle = e^{-\mathcal{H}t}|\psi_0\rangle, \tag{1.54}$$

where $|\psi_{\text{GS}}\rangle$ is the ground state of the system and $\psi_0\rangle$ is an initial guess with some overlap with the ground state. Within the context of this paper, it is also helpful to note that this ITE for the quantum Schrödinger equation is mathematically equivalent to doing time evolution of the master equation to arrive at the steady-state. TEBD further simplifies this approach and can be easiest understood when considering the one-dimensional version, which we will do here. Locally acting Hamiltonians can be broken into separate terms, for example, a nearest-neighbor Hamiltonian can be broken into commuting terms $\mathcal{H}_{\text{even}}$ and $\mathcal{H}_{\text{odd}}$, which contain the terms that respectively act on the even and odd bonds in the MPS. We can do a Suzuki-Trotter decomposition of the evolution operator for such a Hamiltonian

$$e^{-(\mathcal{H}_{\text{even}}+\mathcal{H}_{\text{odd}})t} = \lim_{\delta t\to 0}\prod_{0}^{t/\delta t} e^{-\mathcal{H}_{\text{even}}\delta t}e^{-\mathcal{H}_{\text{odd}}\delta t} + O(\delta t^2), \tag{1.55}$$

where we have now divided the evolution into two separate steps, illustrated for an MPS in Figure 1.11(a), where each step involves applying the exponentiated local Hamiltonian term to the two relevant sites. This can be extended to 2D systems where, for nearest-neighbor interactions, we get four steps as shown in Figure 1.11(b). After each gate application, the bond dimension of the MPS grows, severely limiting the number of time steps that can be taken exactly. Accordingly, after each gate application, the bond is truncated back to a maximum bond dimension to make the problem tractable, shown in Figure 1.11(c), where the central problem now becomes the best way to make this approximation. In an MPS, because of our ability to specify an MPS representation where the norm environment tensors contract to identities, a truncated SVD provides an optimal truncation.

For the application of a gate within a PEPS, because there is no canonical form, we are no longer able to guarantee that a naive SVD provides the optimal approximation. Two approaches are thus commonly employed. The simple update algorithm acts under the assumption that the PEPS is at least near a canonical form and chooses to use only an SVD to approximate the state with a smaller bond dimension. While this is often a rough approximation, it circumvents the need to contract the environment around each bond before applying the gate resulting in significantly smaller computational costs. The full update algorithm, in contrast, takes the full (approximate) environment into account when doing the truncation by using an alternating least squares approach to arrive at the closest overlap between the full PEPS before and after truncation. Because the simple update is significantly cheaper, it is often used to obtain a good initial guess before refining is done through the full update.

In summary, the PEPS TEBD ground state search proceeds by first doing a Trotter decomposition of the Hamiltonian, breaking it into groups of commuting exponentiated local gates. We then iterate over all gates, applying it to the local PEPS tensors then using the simple update or full update approach to truncate the increased bond dimension. After having applied all Hamiltonian gates, a single time step has been completed and this step is repeated until the energy converges. While this makes up the fundamentals of the PEPS TEBD algorithm, there are a handful of numerical tricks that can improve convergence, such as the use of so-called *reduced* PEPS tensors [106] and local conditioning on the norm environment that are used in practice but too technical for coverage here.

### Combining TNs and LDT

Before finishing this section on TNs, it seems appropriate to discuss briefly some motivations for using TNs for computing large deviation functions in nonequilibrium systems. We would like to address two questions in this section:

1. While the use of TNs for quantum lattice systems is well-justified, why would we expect them to work in this setting?
2. Why are TN methods needed in this field when Monte Carlo methods are established and widely successful?

For the former, we provide separately intuitive, mathematical, and historical answers. From intuition, a TN is designed to mimic the physical layout of the system of interest, with the general idea being that sites are most entangled with nearby sites and any entanglement to sites at further distances can be passed through intermediate sites. While the mathematical definition of the entanglement entropy does not have a known physical interpretation for nonequilibrium steady-states, intuition roughly relates it to correlations within the system, which we expect to behave similarly. We know that the steady-state of a nonequilibrium probability distribution can be expressed as a large rank tensor. If we bipartition the system, a truncated SVD is guaranteed to provide the optimal approximation to the exact steady-state, in essence providing the same motivation that supported TN approaches before area laws were proven. Historically, it turns out that the MPS has been used as a standard analytical tool in studying driven nonequilibrium lattice systems in the form of the matrix product ansatz [118]. Given its successful analytical use, it is natural that numerical algorithms are adapted for use when an analytical solution is not apparent.

To answer the second question, it should first be noted that throughout the physical sciences, it is common to develop multiple methods to solve the same problem. In doing this, it frequently becomes clear that there are regimes where each method performs best and thus a toolkit is created that allows for the study of myriad systems and phenomena. From a naive comparison, TN approaches have an advantage because they are entirely noiseless, circumventing some difficulties associated with observing and quantifying rare events using Monte Carlo methods. Additionally, while Monte Carlo algorithms require simulation times that scale with the system's characteristic time scales, TN approaches can find the steady-state behavior formally in the infinite-time limit, avoiding finite time errors.

In this discussion we have thus far assumed that Monte Carlo and TN will work entirely independently; it can be helpful to consider additionally how they can be used in concert. One possible combination would be to use TNs to compute the eigenstate of the tilted generator, then use this resulting state to determine near-ideal control forces for Monte Carlo simulations. Indeed, work complementary to ours has recently demonstrated success in this area for glassy models [119, 120].

Finally, we can imagine some areas of future work that will prove TNs worthwhile in this field. In simulating quantum systems, an active area of work has centered on developing TN algorithms that give results formally in the infinite system size limit [90, 106, 108]. Extending these algorithms to study, for example, driven lattice models is nontrivial because one must account for boundary effects not present in equilibrium quantum systems. If possible, though, the ability to simulate classical nonequilibrium systems formally in the infinite size limit would undoubtedly be an important development. A second possible area of future research could be to utilize the framework of TNs to gain fundamental theoretical insight into the physics of nonequilibrium systems. In quantum systems, the paradigm of TNs has enabled a deeper understanding of area law behavior and topological phase transitions [93, 121]; it is certainly conceivable that such insight could be provided in this field, where we possibly don't have the correct toolkit developed as yet.

Thus while we fully expect Monte Carlo to remain widely used, there is ample evidence for us to expect that TNs will play a complementary role in both numerical and theoretical progress.

## 1.5 Overview of Thesis

In the first chapter of this thesis, we have sought to provide sufficient theoretical and technical background for the reader to feel familiar and comfortable with NESM, large deviation theory, and tensor networks. The following three chapters are each entirely self-contained, but compose a continuous trajectory, evolving from one-dimensional to multi-lane lattice models in Chapter 2, to fully two-dimensional models in Chapter 3, before demonstrating progress towards calculations for continuum models with long-range interactions in Chapter 4. In Chapters 2 and 3, we use standard TN methods to study a canonical driven lattice model, namely the asymmetric simple exclusion process (ASEP). Both of these serve as a proof of principle that TN methods can be adapted and successfully used to compute large deviation functions. We then use these to study the physics of the ASEP, analyzing the behavior of a dynamical phase transition between a jammed and flowing phase, similar to what has been observed in glassy systems that exhibit dynamical heterogeneity. Because many of the most interesting nonequilibrium systems occur with continuous spatial degrees of freedom and involve particles that interact with one another over long ranges, Chapter 4 presents work towards developing numerical methods appropriate for that setting. A particularly difficult challenge for these calculations is the development of compact representations of long-range operators that can be used in practical calculations. In Chapter 4, we present three long-range operator representations, designed for use in a DMRG-style optimization algorithm, and characterize their utility and efficiency in practice using a coulombic Heisenberg model. While the test case here is quantum, our approach is agnostic to whether we are studying quantum or classical nonequilibrium systems.

# BIBLIOGRAPHY

[1] J. Hemminger, G. Fleming, M. Ratner, *Directing Matter and Energy: Five Challenges for Science and the Imagination*, US DOE - Office of Basic Energy Sciences, **2007**.

[2] *Condensed-Matter and Materials Physics: The Science of the World Around Us*, National Research Council, The National Academies Press, Washington, DC, **2007**.

[3] H. Touchette, *Physics Reports* **2009**, *478*, 1–69.

[4] B. Derrida, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, *2007*, P07023.

[5] H. Touchette, R. J. Harris, *arXiv:1110.5216* **2011**.

[6] D. T. Limmer, C. Y. Gao, A. R. Poggioli, *arXiv:2104.05194* **2021**.

[7] G. Biroli, J. P. Garrahan, *The Journal of Chemical Physics* **2013**, *138*, 12A301.

[8] F. Cagnetta, E. Mallmin, *Physical Review E* **2020**, *101*, 022130.

[9] A. Das, D. T. Limmer, *The Journal of Chemical Physics* **2021**, *154*, 014107.

[10] S. Whitelam, K. Klymko, D. Mandal, *The Journal of Chemical Physics* **2018**, *148*, 154902.

[11] Y.-E. Keta, É. Fodor, F. van Wijland, M. E. Cates, R. L. Jack, *Physical Review E* **2021**, *103*, 022603.

[12] T. GrandPre, K. Klymko, K. K. Mandadapu, D. T. Limmer, *Physical Review E* **2021**, *103*, 012613.

[13] K. Klymko, P. L. Geissler, J. P. Garrahan, S. Whitelam, *Physical Review E* **2018**, *97*, 032123.

[14] U. Ray, D. T. Limmer, *Physical Review B* **2019**, *100*, 241409.

[15] N. E. Strand, R.-S. Fu, T. R. Gingrich, *Physical Review E* **2020**, *102*, 012141.

[16] T. R. Gingrich, G. M. Rotskoff, S. Vaikuntanathan, P. L. Geissler, *New Journal of Physics* **2014**, *16*, 102003.

[17] U. Ray, G. K.-L. Chan, D. T. Limmer, *The Journal of Chemical Physics* **2018**, *148*, 124120.

[18] A. Das, D. T. Limmer, *The Journal of Chemical Physics* **2019**, *151*, 244123.

[19] D. C. Rose, J. F. Mair, J. P. Garrahan, *New Journal of Physics* **2021**, *23*, 013013.

[20] U. Ray, G. Kin-Lic Chan, *The Journal of Chemical Physics* **2020**, *152*, 104107.

[21] C. Casert, T. Vieijra, S. Whitelam, I. Tamblyn, *arXiv:2011.08657* **2020**.

[22] S. Whitelam, D. Jacobson, I. Tamblyn, *The Journal of Chemical Physics* **2020**, *153*, 044113.

[23] D. Jacobson, S. Whitelam, *Physical Review E* **2019**, *100*, 052139.

[24] A. Guyader, H. Touchette, *Journal of Statistical Physics* **2020**, *181*, 551–586.

[25] G. M. Rotskoff, E. Vanden-Eijnden, *arXiv:2008.06334* **2020**.

[26] R. Kubo, *Journal of the Physical Society of Japan* **1957**, *12*, 570–586.

[27] M. S. Green, *The Journal of Chemical Physics* **1954**, *22*, 398–413.

[28] L. Onsager, *Physical Review* **1931**, *37*, 405.

[29] D. Chandler, *Introduction to Modern Statistical Mechanics*, *Vol. 40*, Oxford University Press, Oxford, UK, **1987**.

[30] G. L. Daquila, U. C. Täuber, *Physical Review E* **2011**, *83*, 051107.

[31] P. Calabrese, A. Gambassi, *Physical Review B* **2002**, *66*, 212407.

[32] D. J Evans, G. P Morriss, *Statistical Mechanics of Nonequilbrium Liquids*, ANU Press, **2007**.

[33] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, *Vol. 1*, Elsevier, **1992**.

[34] B. Schmittmann, R. K. Zia, *Statistical Mechanics of Driven Diffusive Systems*, *Vol. 17*, Elsevier, **1995**, pages 3–214.

[35] R. Kubo, *Reports on Progress in Physics* **1966**, *29*, 255.

[36] E. M. Sevick, R. Prabhakar, S. R. Williams, D. J. Searles, *Annual Review of Physical Chemistry* **2008**, *59*, 603–633.

[37] G. Gallavotti, E. G. D. Cohen, *Physical Review Letters* **1995**, *74*, 2694.

[38] G. E. Crooks, *Physical Review E* **1999**, *60*, 2721.

[39] J. L. Lebowitz, H. Spohn, *Journal of Statistical Physics* **1999**, *95*, 333–365.

[40] A. C. Barato, U. Seifert, *Physical Review Letters* **2015**, *114*, 158101.

[41] T. R. Gingrich, J. M. Horowitz, N. Perunov, J. L. England, *Physical Review Letters* **2016**, *116*, 120601.

[42] G. Falasco, M. Esposito, J.-C. Delvenne, *New Journal of Physics* **2020**, *22*, 053046.

[43] T. Harada, S.-i. Sasa, *Physical Review Letters* **2005**, *95*, 130602.

[44] U. Seifert, T. Speck, *EPL (Europhysics Letters)* **2010**, *89*, 10007.

[45] J. Prost, J.-F. Joanny, J. M. Parrondo, *Physical Review Letters* **2009**, *103*, 090601.

[46] H. Cramér, H. Touchette, *arXiv:1802.05988* **2018**.

[47] M. D. Donsker, S. S. Varadhan, *Communications on Pure and Applied Mathematics* **1975**, *28*, 1–47.

[48] M. Donsker, S. Varadhan, *Communications on Pure and Applied Mathematics* **1975**, *28*, 279–301.

[49] M. D. Donsker, S. S. Varadhan, *Communications on Pure and Applied Mathematics* **1976**, *29*, 389–461.

[50] M. D. Donsker, S. S. Varadhan, *Communications on Pure and Applied Mathematics* **1983**, *36*, 183–212.

[51] Y. Kifer, *Nonlinear Problems in Future Particle Accelerators* **1988**, 189.

[52] O. E. Lanford in *Statistical Mechanics and Mathematical Problems*, Springer, **1973**, pages 1–113.

[53] Y. Oono, *Progress of Theoretical Physics Supplement* **1989**, *99*, 165–205.

[54] R. S. Ellis, *Scandinavian Actuarial Journal* **1995**, *1995*, 97–142.

[55] R. S. Ellis, *Physica D: Nonlinear Phenomena* **1999**, *133*, 106–136.

[56] R. S. Ellis, *Entropy, large deviations, and statistical mechanics*, *Vol. 1431*, Taylor & Francis, **2006**.

[57] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, *2007*, P07014.

[58] H. Touchette, *arXiv:1106.4146* **2011**.

[59] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, *Journal of Physics A: Mathematical and Theoretical* **2009**, *42*, 075007.

[60] T. Nemoto, É. Fodor, M. E. Cates, R. L. Jack, J. Tailleur, *Physical Review E* **2019**, *99*, 022605.

[61] A. Tenenbaum, G. Ciccotti, R. Gallico, *Physical Review A* **1982**, *25*, 2778–2787.

[62] W. G. Hoover, D. J. Evans, R. B. Hickman, A. J. C. Ladd, W. T. Ashurst, B. Moran, *Physical Review A* **1980**, *22*, 1690–1697.

[63] D. J. Evans, *Physics Letters A* **1982**, *91*, 457–460.

[64] C. Y. Gao, D. T. Limmer, *Entropy* **2017**, *19*, 571.

[65] C. Y. Gao, D. T. Limmer, *The Journal of Chemical Physics* **2019**, *151*, 014101.

[66] F. Ragone, J. Wouters, F. Bouchet, *Proceedings of the National Academy of Sciences* **2018**, *115*, 24–29.

[67] D. Chandler, J. P. Garrahan, *Annual Review of Physical Chemistry* **2010**, *61*, 191–217.

[68] J. P. Garrahan, P. Sollich, C. Toninelli, *Dynamical Heterogeneities in Glasses Colloids and Granular Media* **2011**, *150*, 111–137.

[69] L. O. Hedges, R. L. Jack, J. P. Garrahan, D. Chandler, *Science* **2009**, *323*, 1309–1313.

[70] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *Reviews of Modern Physics* **2015**, *87*, 593.

[71] P. G. Bolhuis, D. Chandler, C. Dellago, P. L. Geissler, *Annual Review of Physical Chemistry* **2002**, *53*, 291–318.

[72] C. Giardina, J. Kurchan, V. Lecomte, J. Tailleur, *Journal of Statistical Physics* **2011**, *145*, 787–811.

[73] A. Dhar, K. Saito, B. Derrida, *Physical Review E* **2013**, *87*, 010103.

[74] H. C. Fogedby, A. Imparato, *Journal of Statistical Mechanics: Theory and Experiment* **2011**, *2011*, P05015.

[75] T. GrandPre, D. T. Limmer, *Physical Review E* **2018**, *98*, 060601.

[76] S. R. White, *Physical Review Letters* **1992**, *69*, 2863.

[77] S. Östlund, S. Rommer, *Physical Review Letters* **1995**, *75*, 3537.

[78] J. Dukelsky, M. A. Martın-Delgado, T. Nishino, G. Sierra, *EPL (Europhysics Letters)* **1998**, *43*, 457.

[79] U. Schollwöck, *Annals of Physics* **2011**, *326*, 96–192.

[80] G. K.-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, T. Yanai in *Frontiers in Quantum Systems in Chemistry and Physics*, Springer, **2008**, pages 49–65.

[81] T. Nishino, *Journal of the Physical Society of Japan* **1995**, *64*, 3598–3601.

[82] T. Johnson, S. Clark, D. Jaksch, *Physical Review E* **2010**, *82*, 036702.

[83] E. M. Stoudenmire, D. J. Schwab, *arXiv:1605.05775* **2016**.

[84] Z.-Y. Han, J. Wang, H. Fan, L. Wang, P. Zhang, *Physical Review X* **2018**, *8*, 031012.

[85] S. Efthymiou, J. Hidary, S. Leichenauer, *arXiv:1906.06329* **2019**.

[86] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, C. Hubig, *Annals of Physics* **2019**, *411*, 167998.

[87] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, F. Verstraete, *Physical Review B* **2016**, *94*, 165116.

[88] F. Verstraete, J. J. Garcia-Ripoll, J. I. Cirac, *Physical Review Letters* **2004**, *93*, 207204.

[89] B. Bruognolo, Z. Zhu, S. R. White, E. M. Stoudenmire, *arXiv:1705.05578* **2017**.

[90] I. P. McCulloch, *arXiv:0804.2509* **2008**.

[91] F. Verstraete, J. I. Cirac, *Physical Review Letters* **2010**, *104*, 190405.

[92] M. Dolfi, B. Bauer, M. Troyer, Z. Ristivojevic, *Physical Review Letters* **2012**, *109*, 020604.

[93] M. B. Hastings, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, *2007*, P08024.

[94] E. M. Stoudenmire, S. R. White, *Annual Review of Condensed Matter Physics* **2012**, *3*, 111–128.

[95] F. Verstraete, J. I. Cirac, *arXiv:cond-mat/0407066* **2004**.

[96] R. Haghshenas, M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2019**, *100*, 054404.

[97] K. Hyatt, E. M. Stoudenmire, *arXiv:1908.08833* **2019**.

[98] M. Lubasch, J. I. Cirac, M.-C. Banuls, *New Journal of Physics* **2014**, *16*, 033014.

[99] M. Lubasch, J. I. Cirac, M.-C. Banuls, *Physical Review B* **2014**, *90*, 064425.

[100] M. J. O'Rourke, Z. Li, G. K.-L. Chan, *Physical Review B* **2018**, *98*, 205127.

[101] Z. Li, M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2019**, *100*, 155121.

[102] M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2020**, *101*, 205142.

[103] Y.-Y. Shi, L.-M. Duan, G. Vidal, *Physical Review A* **2006**, *74*, 022320.

[104] G. Vidal, *Physical Review Letters* **2007**, *99*, 220405.

[105] G. Evenbly, G. Vidal, *Physical Review Letters* **2009**, *102*, 180406.

[106] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, R. Orús, *Physical Review B* **2015**, *92*, 035142.

[107] E. Ronca, Z. Li, C. A. Jimenez-Hoyos, G. K.-L. Chan, *Journal of Chemical Theory and Computation* **2017**, *13*, 5560–5571.

[108] P. Corboz, *Physical Review B* **2016**, *94*, 035133.

[109] H.-J. Liao, J.-G. Liu, L. Wang, T. Xiang, *Physical Review X* **2019**, *9*, 031041.

[110] Z. Xie, H. Liao, R. Huang, H. Xie, J. Chen, Z. Liu, T. Xiang, *Physical Review B* **2017**, *96*, 045128.

[111] G. M. Crosswhite, D. Bacon, *Physical Review A* **2008**, *78*, 012356.

[112] E. M. Stoudenmire, S. R. White, *Physical Review Letters* **2017**, *119*, 046401.

[113] F. Fröwis, V. Nebendahl, W. Dür, *Physical Review A* **2010**, *81*, 062337.

[114] L. Lin, Y. Tong, *SIAM Journal on Scientific Computing* **2021**, *43*, A164–A192.

[115] M. P. Zaletel, F. Pollmann, *Physical Review Letters* **2020**, *124*, 037201.

[116] A. Kshetrimayum, M. Rizzi, J. Eisert, R. Orús, *Physical Review Letters* **2019**, *122*, 070502.

[117] P. Czarnik, J. Dziarmaga, *Physical Review B* **2015**, *92*, 035120.

[118] R. A. Blythe, M. R. Evans, *Journal of Physics A: Mathematical and Theoretical* **2007**, *40*, R333.

[119] M. C. Bañuls, J. P. Garrahan, *Physical Review Letters* **2019**, *123*, 200601.

[120] L. Causer, M. C. Bañuls, J. P. Garrahan, *arXiv:2103.01265* **2021**.

[121] N. Schuch, D. Perez-Garcia, I. Cirac, *Physical Review B* **2011**, *84*, 165139.

## *Chapter 2*

# DYNAMICAL PHASE BEHAVIOR OF THE SINGLE- AND MULTI-LANE ASYMMETRIC SIMPLE EXCLUSION PROCESS VIA MATRIX PRODUCT STATES

The open asymmetric simple exclusion process (ASEP) has emerged as a paradigmatic model of nonequilibrium behavior, in part due to its complex dynamical behavior and wide physical applicability as a model of driven diffusion. We compare the dynamical phase behavior of the 1D ASEP and the multi-lane ASEP, a previously unstudied extension of the 1D model that may be thought of as a finite-width strip of the fully 2D system. Our characterization employs large deviation theory (LDT), matrix product states (MPS), and the density matrix renormalization group (DMRG) algorithm, to compute the current cumulant generating function and its derivatives, which serve as dynamical order parameters. We use this measure to show that when particles cannot exit or enter the lattice vertically, the phase behavior of the multi-lane ASEP mimics that of its 1D counterpart, exhibiting the macroscopic and microscopic signatures of the maximal current, shock, and high-density/low-density coexistence phases. Conversely, when particles are allowed to freely enter and exit the lattice, no such transition is observed. This contrast emphasizes the complex interplay between latitudinal and longitudinal hopping rates and the effect of current biasing. Our results support the potential of tensor networks as a framework to understand classical nonequilibrium statistical mechanics.

This chapter is adapted from our previously published article:

[1] P. Helms et al., *Physical Review E* **2019**, *100*, 022101,

## 2.1 Introduction

In recent years, the asymmetric simple exclusion process (ASEP) has emerged as a paradigmatic model of nonequilibrium behavior in statistical mechanics [2]. The basis for this popularity resembles that of the Ising model: a simply defined model with contrastingly complex behavior and wide applicability. The ASEP is defined by bulk and boundary hopping rates that govern the stochastic movement of particles between sites on a 1D lattice, limiting each site to an occupancy of one particle at most [3]. This simplistic model, originally used to study protein synthesis [4], has since been applied to understand diverse physical problems such as the transport properties of molecular motors [5], polymer reptation [6], transport through membranes [7], and surface growth [8].

Because of its simplicity and applicability, the open ASEP has been studied extensively, revealing a complex dynamical phase diagram, with both boundary and bulk driven phase transitions existing between many possible phases [9]. Additionally, many exact and semi-analytic results have been derived, making the ASEP a good candidate for benchmarking computational methods [10–13].

Much recent attention on ASEP has centered on more complex realizations of the model, such as multi-species and quantum analogs [14, 15] or studies of the effects of spatial inhomogeneities [16]. In this work, we characterize the behavior of the multi-lane ASEP, which can be thought of as a finite-width strip of the fully 2D model. Previously, analysis of the behavior of the multi-lane ASEP has been limited to specific derivatives of the two-lane model and has focused on the mean behaviors of the relevant observables, potentially missing critical details encoded by fluctuations [17–20]. Alternatively, current fluctuations in the fully 2D ASEP have been explored within macroscopic fluctuation theory, although numerical validation of the resulting expressions has been limited due to the cost of the required computations [21–23].

In this work, we provide a first step towards numerically interpolating between the behaviors of the 1D and fully 2D systems by studying the multi-lane ASEP with up to four lanes. Specifically, we work within the framework established by large deviation theory (LDT) [24–26] and compute the current cumulant generating function (CGF), whose derivatives encode fluctuations of the current and serve as dynamical order parameters.

Because of the difficulties associated with measuring rare events in large or complex systems, significant effort has been devoted to the development of appropriate and robust numerical and analytic approaches for computing large deviation functions such as the CGF. Monte Carlo sampling methods (such as the cloning algorithm and transition path sampling [27–29] augmented with importance sampling [22, 30, 31] and direct rate function evaluation techniques [32]) have been applied to both lattice and continuum nonequilibrium systems [10, 12, 27, 29, 33–42]. Here, we compute these functions using matrix product states (MPS) and the density matrix renormalization group (DMRG) algorithm. This approach is an example of a tensor network (TN) method, where the high-dimensional probability distribution is represented as a contraction of many tensors. The TN approach has been used to study nonequilibrium lattice problems both analytically, via the matrix product ansatz [11, 43], and computationally [10, 36, 44–46]. Recently, it has been used to understand kinetically constrained models of glasses [47].

The remainder of this report proceeds by first providing a brief introduction to large deviation theory, matrix product states, and the density matrix renormalization group algorithm. We then calibrate our implementation on the dynamical phase behavior of the 1D (single-lane) ASEP, where we find excellent agreement between our results and the semi-analytic expressions for the current cumulant generating function in the region of applicability of the functional Bethe ansatz [13, 48, 49]. We note that DMRG has previously been successfully used to validate expressions for high-order current cumulants of the 1D ASEP [10] and to compute critical exponents in the totally asymmetric case [36]. We also report on microscopic observables, such as the local density and activity. While the behavior of these observables has been understood from approximate theories or near various analytically tractable limits [49], the exact numerically computed quantities have typically not been reported.

We use the 1D results as a framework to extend our study to the multi-lane ASEP with up to four lanes. Here, we describe the effect of vertical hopping rates on the longitudinal dynamical phase behavior by comparing the behaviors of the closed multi-lane ASEP, where particle insertion and removal is only permitted at the horizontal boundaries, and the open multi-lane ASEP, where particles freely enter and exit the lattice vertically. The comparison of the behaviors of these models reveals the complexity of the effects of vertical hopping rates on longitudinally biased systems, and serves as a step towards understanding the fully 2D ASEP.

## 2.2 Large Deviation Theory and Matrix Product States

We first briefly summarize some relevant concepts in large deviation theory, the theory of matrix product states and the density matrix renormalization group. A more complete description can be found in recent reviews [26, 29, 50, 51].

In a nonequilibrium system, the state vector $|P_t\rangle$ evolves from an initial state $|P_0\rangle$ according to a master equation with dynamics generated by a non-Hermitian Markov operator $\mathcal{W}$,

$$\partial_t|P_t\rangle = \mathcal{W}|P_t\rangle, \tag{2.1}$$

with the probability of a system configuration $C$ at time $t$ given by $\mathrm{Prob}(C_t) \equiv \langle C|P_t\rangle$. The long-time limit yields the final (steady) state $|P_\infty\rangle$. The probability of observing a given trajectory of configurations $\mathscr{C}(t_N) = \{C_0, C_1, \ldots, C_{t_N}\}$ at times $\{t_0, \cdots, t_N\}$ ($dt = t_N/N$) is

$$\mathrm{Prob}(\mathscr{C}(t_N)) = \mathrm{Prob}(C_0)\prod_{i=0}^{t_{N-1}}\langle C_{i+1}|e^{dt\mathcal{W}}|C_i\rangle. \tag{2.2}$$

We can define dynamical observables along such a trajectory, such as a time-local observable $O = \sum_{i=0}^{t_{N-1}} o(C_{i+1}, C_i)$, with $o$ being an arbitrary function of time-adjacent configurations ($C_{i+1}$ and $C_i$). To characterize the steady-state expectation value and fluctuations of this observable, we define a cumulant generating function

$$\begin{aligned}\psi(\lambda) &= \lim_{t_N\to\infty} t_N^{-1}\ln\left\langle e^{-\lambda O}\right\rangle \\ &= \lim_{t_N\to\infty} t_N^{-1}\ln\sum_{\mathscr{C}(t_N)}\mathrm{Prob}(\mathscr{C}(t_N))e^{-\lambda O},\end{aligned} \tag{2.3}$$

where $\lambda$ is a field conjugate to the observable. At $\lambda = 0$, the first derivative of $\psi$ is the observable's steady-state expectation value $\langle o\rangle$; characterizations of the fluctuations of $o$, via its cumulants, are obtained from higher-order derivatives of $\psi$. A fundamental result in LDT is that $\psi(\lambda)$ is the largest eigenvalue $E_0$, of a tilted operator $\mathcal{W}_\lambda$, i.e.

$$\mathcal{W}_\lambda|P^\lambda\rangle = \psi(\lambda)|P^\lambda\rangle, \tag{2.4}$$

where, for discrete configurations, the tilted operator is

$$\mathcal{W}_\lambda(C, C') = \mathcal{W}(C, C')e^{-\lambda o(C,C')}(1 - \delta_{C,C'}) - R(C)\delta_{C,C'} \tag{2.5}$$

with $R(C) = \sum_{C\neq C'}\mathcal{W}(C, C')$ and with right and left eigenvectors $|P^\lambda\rangle$ and $\langle\bar{P}^\lambda|$. The eigenvectors give the configurational probabilities at initial, intermediate, and final times, respectively being

$$
\begin{aligned}
|P^{\lambda}_{t_0}\rangle &= \frac{\mathrm{diag}\left(|\bar{P}^{\lambda}\rangle \otimes |P^{\lambda=0}\rangle\right)}{\langle \bar{P}^{\lambda}|P^{\lambda=0}\rangle}, \\
|P^{\lambda}_{t_{\mathrm{Int.}}}\rangle &= \frac{\mathrm{diag}\left(|\bar{P}^{\lambda}\rangle \otimes |P^{\lambda}\rangle\right)}{\langle \bar{P}^{\lambda}|P^{\lambda}\rangle}, \\
|P^{\lambda}_{t_f}\rangle &= \frac{|P^{\lambda}\rangle}{\sum_C \langle C|P^{\lambda}\rangle},
\end{aligned}
\tag{2.6}
$$

where the full trajectory satisfies $\langle O\rangle = d\psi(\lambda)/d\lambda$ [52–54].

The computation of $\psi(\lambda)$ and each of the eigenvectors in Eq. (2.4) is achievable via exact diagonalization for only the smallest systems. Alternatively, the equation can be recast as a generalized variational problem

$$
\langle \delta P^{\lambda}|\mathcal{W}_{\lambda}|P^{\lambda}\rangle - \psi(\lambda)\langle \delta P^{\lambda}|P^{\lambda}\rangle = 0, \tag{2.7}
$$

where we seek to make $\psi(\lambda) = \langle P^{\lambda}|\mathcal{W}_{\lambda}|P^{\lambda}\rangle$ stationary with respect to a perturbation from $|P^{\lambda}\rangle$ to $|\delta P^{\lambda}\rangle$. Because $\mathcal{W}_{\lambda}$ is non-Hermitian, $\langle P^{\lambda}|\mathcal{W}_{\lambda}|P^{\lambda}\rangle$ may be above or below the exact $\psi(\lambda)$ for an approximate $|P^{\lambda}\rangle$ or $\langle \bar{P}^{\lambda}|$.

In this work, we use an MPS as an ansatz for $|P^{\lambda}\rangle$ and perform the optimization in Eq. (2.7) using the DMRG algorithm for non-Hermitian operators [44, 55]. To introduce the MPS ansatz, consider a 1D lattice of length $L$ with sites $i = 1 \dots L$. Each site has a local state space $\{\sigma_i\}$ of dimension $d$, with a system configuration $C$ being an ordered list of the local states, $|C\rangle = |\sigma_1, \cdots, \sigma_L\rangle$. The state vector is specified by a tensor of weights

$$
|P^{\lambda}\rangle = \sum_{\{\sigma_1 \cdots \sigma_L\}} \boldsymbol{c}_{\sigma_1,\cdots,\sigma_L} |\sigma_1, \cdots, \sigma_L\rangle, \tag{2.8}
$$

with $\boldsymbol{c}_{\sigma_1,\cdots,\sigma_L}$ specifying the probability of the system being in configuration $|\sigma_1, \cdots, \sigma_L\rangle$.

In this exact representation, arbitrary strong correlations can exist between all sites. However, if the Markov operator $\mathcal{W}_{\lambda}$ only produces transitions between nearby sites, we can expect correlations to decay with distance. An efficient way to represent states with this property is to rewrite $\boldsymbol{c}_{\sigma_1,\cdots,\sigma_L}$ as a product of matrices, i.e. a matrix product state

$$
|P^{\lambda}\rangle = \sum_{\{\sigma_1 \cdots \sigma_L\}} \boldsymbol{M}^{\sigma_1}\boldsymbol{M}^{\sigma_2} \cdots \boldsymbol{M}^{\sigma_{L-1}}\boldsymbol{M}^{\sigma_L} |\sigma_1, \cdots, \sigma_L\rangle\,. \tag{2.9}
$$

where the dimension of the matrices, also called the bond dimension $D$, specifies the amount of correlation that can be transmitted between sites. If we assume that

$D$ saturates with system size, then the representation is asymptotically linear in complexity with respect to system size, i.e. it contains only $O(dD^2L)$ parameters.

Matrix product states with a size-independent $D$ are said to satisfy the 1D area law. In the quantum mechanical setting, the area law states that the entanglement entropy between two partitions of the system is proportional to the length of the boundary between them: in 1D, this is independent of system size. It is known that gapped Hamiltonians in 1D produce ground-states which satisfy this law and thus can be ideally represented by matrix product states [56]. However, note that even when the area law is not satisfied, one can still exactly represent an arbitrary state with an MPS by using a sufficiently large $D$. For example, to satisfy the area law for a multi-lane ASEP, we can use an MPS with a $D$ that grows exponentially with the number of lanes. In the multi-lane case, the representation also depends on the 1D traversal pattern of the sites. Here, we use a zig-zag ordering of sites, shown in Fig. 2.1.

With $|P^\lambda\rangle$ written as an MPS, the DMRG algorithm solves the variational problem posed in Eq. (2.7) optimizing one $\boldsymbol{M}^{\sigma_i}$ at a time by solving an eigenproblem at each site of the form $\mathcal{W}_i^{\text{eff}} \cdot \boldsymbol{M}^{\sigma_i} = \psi(\lambda) N_i \cdot \boldsymbol{M}^{\sigma_i}$, where $\mathcal{W}_i^{\text{eff}}$ describes the action of $\mathcal{W}^\lambda$ in the vector space containing $\boldsymbol{M}^{\sigma_i}$. The metric $N_i$ can be eliminated (i.e. converted to the identity) by using the gauge freedom in the MPS, i.e. a matrix and its inverse may be inserted between any two sites without changing $|P^\lambda\rangle$,

$$\begin{aligned} \text{Prob}(C) =& \boldsymbol{M}^{\sigma_1}\boldsymbol{M}^{\sigma_2}\cdots\boldsymbol{M}^{\sigma_{L-1}}\boldsymbol{M}^{\sigma_L}, \\ =& \boldsymbol{M}^{\sigma_1}\boldsymbol{X}^{-1}\boldsymbol{X}\boldsymbol{M}^{\sigma_2}\cdots\boldsymbol{M}^{\sigma_{L-1}}\boldsymbol{M}^{\sigma_L}, \\ =& \boldsymbol{M}'^{\sigma_1}\boldsymbol{M}'^{\sigma_2}\cdots\boldsymbol{M}^{\sigma_{L-1}}\boldsymbol{M}^{\sigma_L}. \end{aligned} \tag{2.10}$$

Choosing the gauge to eliminate the metric yields the canonical form at the site,

$$|P^\lambda\rangle = \sum_{\{\sigma_1\cdots\sigma_L\}} \boldsymbol{L}^{\sigma_1}\boldsymbol{L}^{\sigma_2}\cdots\boldsymbol{F}^{\sigma_i}\cdots\boldsymbol{R}^{\sigma_{L-1}}\boldsymbol{R}^{\sigma_L}\,|\sigma_1,\cdots,\sigma_L\rangle, \tag{2.11}$$

where $\sum_\sigma \boldsymbol{L}^{\sigma\dagger}\boldsymbol{L}^\sigma = \boldsymbol{I}$ and $\sum_\sigma \boldsymbol{R}^\sigma\boldsymbol{R}^{\sigma\dagger} = \boldsymbol{I}$ and $\boldsymbol{F}^{\sigma_i}$ now denotes the tensor optimized in the local eigenvalue problem. A series of sweeps of optimizations is then performed over the sites, until convergence to the targeted eigenstate of the tilted generator.

The canonical form of Eq. (2.11) is also used to define the bipartite entanglement entropy $S(i)$ at site $i$. Though entanglement is strictly a physical property of quantum systems, here the numerical value of $S(i)$ can still be used to quantify the non-factorizable correlations between the states of sites to the left and right of site $i$,

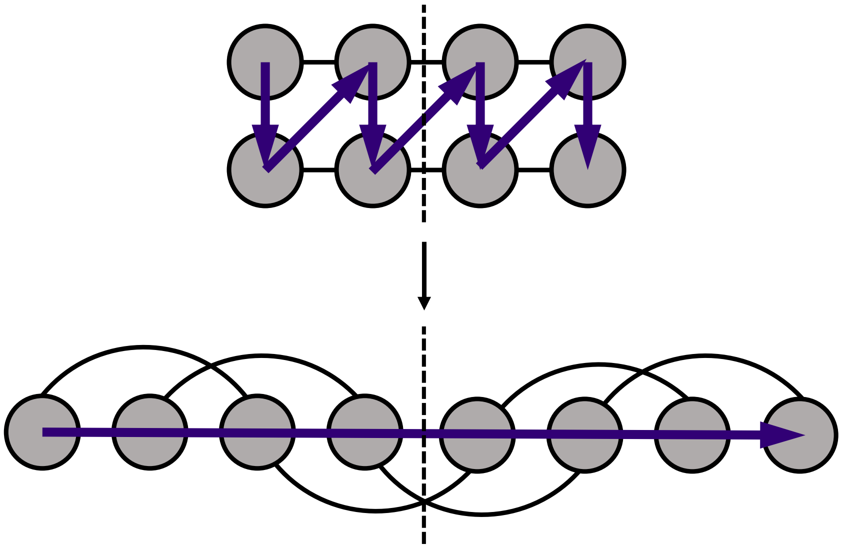

Figure 2.1: A diagrammatic representation of the mapping of a 2D lattice with nearest neighbor interactions onto a 1D lattice with long-range interactions. The arrows indicate how our DMRG optimization traverses the 2D lattice and the dashed line shows the bond over which the numerical entanglement entropy is measured.

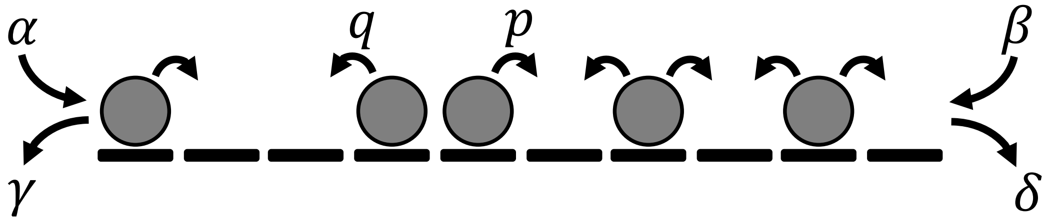

Figure 2.2: The ASEP model where particles on a 1D lattice stochastically hop to a vacant neighboring right (left) site at a rate of $p$ ($q$) and enter (exit) at the left and right boundaries at rates $\alpha$ ($\gamma$) and $\beta$ ($\delta$).

and to bound the maximum bond dimension required to accurately represent the state as an MPS. It can also be used as a generalized order parameter in quantum applications and may thus provide similar insights here [57, 58]. By computing the singular values $\{s_m\}$ of the central $\boldsymbol{F}^{\sigma_i}_{pq}$ reshaped into the matrix $\boldsymbol{G}_{\sigma_i p,q} = \boldsymbol{F}^{\sigma_i}_{pq}$, the numerical entanglement entropy is defined as

$$S(i) = -\sum_m s_m^2 \log_2 s_m^2. \tag{2.12}$$

## 2.3 Model

The 1D ASEP (Fig. 2.2) takes place on a lattice of $L$ sites. Particles hop stochastically to vacant nearest-neighbor sites at the following rates. In the lattice interior, particles hop right (left) with rate $p$ ($q$) with asymmetry enforced via $p \neq q$ (ASEP). At the edges, particles enter (exit) at the left with rate $\alpha$ ($\gamma$) and at the right with rate $\beta$ ($\delta$). In this work, we designate the time-integrated current $J$ for all bonds as the observable $O$ mentioned previously and study phases induced by the current bias $\lambda$ in the parameter regime $\alpha = \beta = \gamma = \delta = 1/2$ and $p + q = 1$. The tilted operator for the current cumulant generating function is

$$\begin{aligned}\mathcal{W}_{\lambda}^{\mathrm{1D}} =&\alpha \left(e^{\lambda}\boldsymbol{a}_1^{\dagger} - \boldsymbol{v}_1\right) + \gamma \left(e^{-\lambda}\boldsymbol{a}_1 - \boldsymbol{n}_1\right) \\ &+ \sum_{i=1}^{L-1} p \left(e^{\lambda}\boldsymbol{a}_i\boldsymbol{a}_{i+1}^{\dagger} - \boldsymbol{n}_i\boldsymbol{v}_{i+1}\right) \\ &+ \sum_{i=1}^{L-1} q \left(e^{-\lambda}\boldsymbol{a}_i^{\dagger}\boldsymbol{a}_{i+1} - \boldsymbol{v}_i\boldsymbol{n}_{i+1}\right) \\ &+ \beta \left(e^{-\lambda}\boldsymbol{a}_L^{\dagger} - \boldsymbol{v}_L\right) + \delta \left(e^{\lambda}\boldsymbol{a}_L - \boldsymbol{n}_L\right),\end{aligned} \tag{2.13}$$

where $\boldsymbol{a}_i$, $\boldsymbol{a}_i^{\dagger}$, $\boldsymbol{n}_i$, and $\boldsymbol{v}_i$ are annihilation, creation, particle number, and vacancy number operators. Note that the tilted operator is invariant with respect to the combined operation of particle-hole transformation/inversion ($\boldsymbol{a}^{\dagger} \leftrightarrow \boldsymbol{a}$ and $\{..., i, i+1, ...\} \leftrightarrow \{..., i+1, i, ...\}$). The eigenvalues of $\mathcal{W}_{\lambda}^{\mathrm{1D}}$ also exhibit a Gallavotti-Cohen (GC) symmetry [52, 59] of the form $\psi(\lambda) = \psi(\lambda^*)$ where, for the specified ASEP parameters, $\lambda^* = -\frac{L-1}{L+1}\ln(p/q) - \lambda$.

The multi-lane ASEP is defined on a 2D lattice of $L_y \times L_x$ sites. It augments the 1D ASEP with bulk hopping in the vertical (transverse) direction (at rates $p_y$, $q_y$) and particles inserted and removed at the vertical boundaries (at rates $\alpha_y, \beta_y, \gamma_y, \delta_y$). We apply the current bias in the (longitudinal) $x$-direction, with a tilted operator that takes the form,

$$\mathcal{W}_{\lambda}^{\mathrm{2D}} = \mathcal{W}_{\lambda}^{\mathrm{1D_x}} + \mathcal{W}_{0}^{\mathrm{1D_y}}, \tag{2.14}$$

and retains the above GC and particle-hole/inversion symmetries. To understand the effects of the transverse parameters on the longitudinal system's phase behavior, we focus on two multi-lane parameter sets, namely open and closed vertical boundaries. Both require $p_x + q_x = 1$, $p_y = q_y = 1/2$, and $\alpha_x = \beta_x = \gamma_x = \delta_x = 1/2$, while the open (closed) case specifies $\alpha_y = \beta_y = \gamma_y = \delta_y = 1/2$ ($\alpha_y = \beta_y = \gamma_y = \delta_y = 0$).

To characterize the system, the DMRG algorithm is used to determine the largest eigenvalue of each tilted operator, through which the steady-state total current and current susceptibility are computed as $J = \partial\psi(\lambda)/d\lambda$ and $\chi = \partial^2\psi(\lambda)/d\lambda^2$. Local densities, currents, and activities may also be computed by contracting the resulting left and right eigenvector with the appropriate operator, i.e.,

$$\begin{aligned} \rho_i &= \left\langle P^\lambda \right| \boldsymbol{n}_i \left| P^\lambda \right\rangle, \\ J_i &= \left\langle P^\lambda \right| p e^{\lambda} \boldsymbol{a}_i \boldsymbol{a}_{i+1}^{\dagger} - q e^{-\lambda} \boldsymbol{a}_i^{\dagger} \boldsymbol{a}_{i+1} \left| P^\lambda \right\rangle, \\ K_i &= \left\langle P^\lambda \right| p e^{\lambda} \boldsymbol{a}_i \boldsymbol{a}_{i+1}^{\dagger} + q e^{-\lambda} \boldsymbol{a}_i^{\dagger} \boldsymbol{a}_{i+1} \left| P^\lambda \right\rangle, \end{aligned} \tag{2.15}$$

assuming $\left\langle P^\lambda \middle| P^\lambda \right\rangle = 1$.

## 2.4 Results

### Benchmark MPS calculations of the 1D ASEP

We begin by using MPS and DMRG to characterize the phase behavior in the aforementioned parameter space and benchmark this approach against earlier results from the semi-analytical functional Bethe ansatz and approximate results from macroscopic fluctuation theory [9]. In this space, there are three expected phases, which are described in Fig. 2.3(a) via rudimentary sketches of both the steady-state density profile and the most probable particle configurations. These are the Maximal Current (MC) phase, where, in the most probable microscopic configurations, particles are evenly spaced throughout the lattice, allowing a maximal amount of biased hopping, the Shock (S) phase, where particles conglomerate on one side of the lattice to form a shock that, in path-space simulations, performs a Brownian walk on the lattice, and the High-Density/Low-Density Coexistence (HD+LD) phase, where the entirely filled and empty states (with some boundary effects) are degenerate in the thermodynamic limit and correspond to a steady-state density profile of $\rho = 1/2$.

The predicted phase diagram is mapped in Fig. 2.3(b) where the lines indicate the line of GC symmetry (solid), the boundary between the MC and S phases (dotted, via macroscopic fluctuation theory), and the boundary between the S and HD+LD phases (dashed, via functional Bethe ansatz). The steady-state current is also shown, computed via DMRG for an $L = 20$ ASEP, showing that current functions as a dynamical order parameter for the transition from S to HD+LD, going effectively to zero in the HD+LD phase. While the boundary between the MC and S phases is commonly defined as the point where the per site current is $J = (p - q)/4$, we are not aware of an order parameter for this transition, which instead appears as a smooth crossover in the current rather than a true phase boundary. Also note that because of the symmetries of the system, the remaining analysis can be limited to the lower left region of the parameter space ($p < 1/2$ to the left of the line of GC symmetry), with the rest of the diagram mapped out by symmetry.

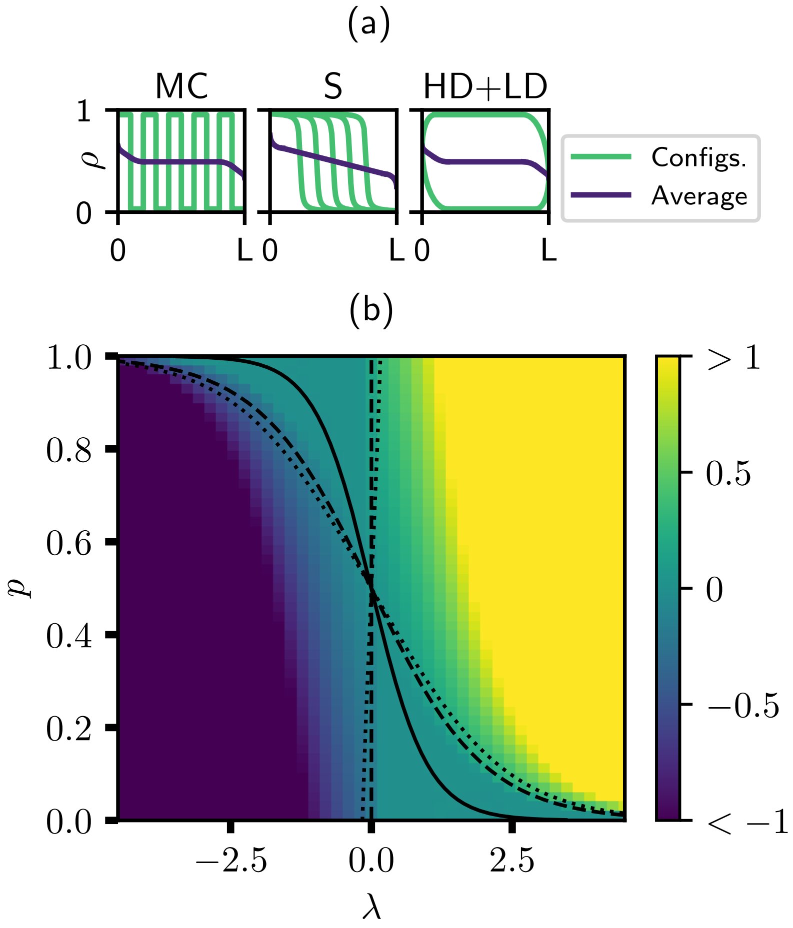

Figure 2.3: (a) Rudimentary sketches of the density profiles in the three possible phases. Blue curves represent approximate steady-state density profiles while green curves depict typical particle configurations. (b) A map of the dynamical phase behavior of the ASEP showing the steady-state current $J$ as a function of $p$ and $\lambda$ for a length $L = 20$ lattice as determined via DMRG. Additionally shown in black are lines indicating the center of the GC symmetry (solid) and the predicted boundaries between the MC and shock phases (dotted, via macroscopic fluctuation theory [9, 60, 61]) and the shock and HD+LD phases (dashed, via functional Bethe ansatz [9, 62]).

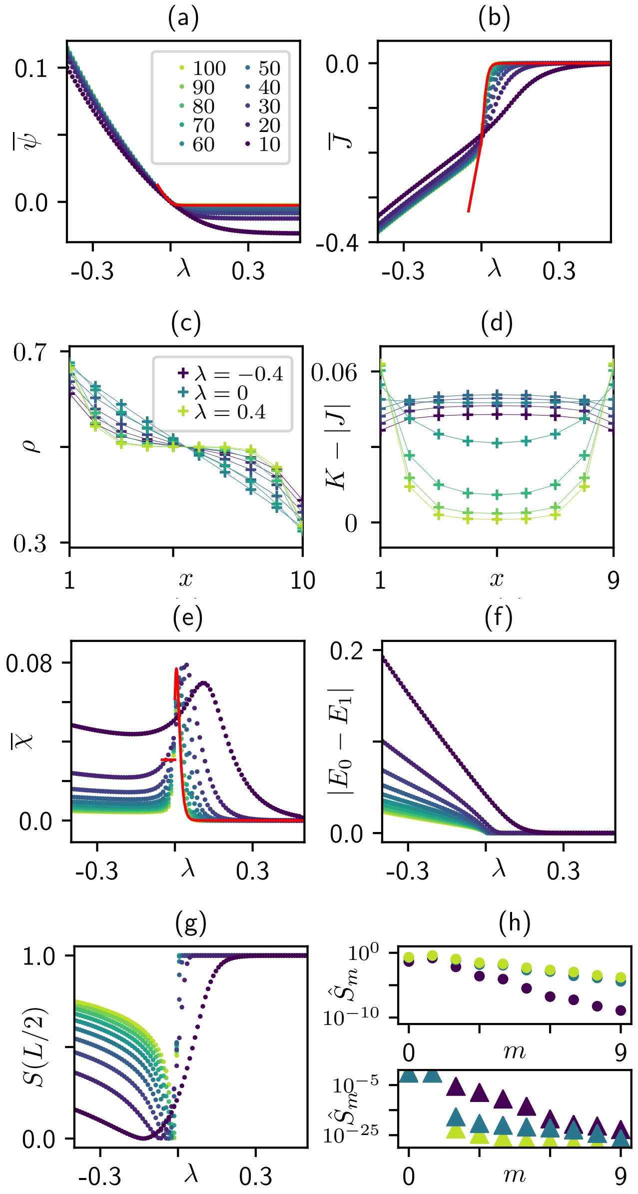

Figure 2.4: The behavior of the 1D ASEP with lattice lengths of $L = [10, 100]$. The DMRG results for the normalized (a) CGF. $\overline{\psi} = E_0/L$, (b), current $\overline{J} = \partial_\lambda \psi / L$, and (e) scaled current susceptibility $\overline{\chi} = \partial_\lambda^2 \psi / L^2$ compared with the analytic functional Bethe ansatz expressions (red), valid for $\lambda \to 0^-$ and $\lambda > 0$; additionally (f) shows the gap between the first and and second largest eigenvalues $E_0$ and $E_1$. Plots (c) and (d) show the density $\rho$ and recurrent hopping $K - |J|$ as a function of position in a $L = 10$ lattice, $x$, and $\lambda$. (g) shows the numerical entanglement entropy $S$ of a bipartition at the center bond as a function of $\lambda$ with the upper (lower) subfigures in (h) showing the corresponding ordered numerical entanglement spectrum, with $\hat{S}_m = -s_m^2 \log_2 s_m^2$, at $\lambda = -0.3$ ($\lambda = 0.3$).

Finite size errors can be converged rapidly by increasing the lattice size. In Fig. 2.4, we characterize this behavior using system properties such as the cumulant generating function, current, current susceptibility, and excited state gap for a range of $\lambda$ near $\lambda = 0$ with $p = 0.1$ and for lattice sizes up to $L = 100$ via DMRG with bond dimension $D$ between 50 and 300. The relative errors in the energy, current, and numerical entanglement entropy in the MC phase for a bond dimension of $D = 10$ are approximately, $\text{Err}_E = 0.01\%$, $\text{Err}_J = 0.02\%$, and $\text{Err}_S = 10\%$, respectively. Increasing to a bond dimension of $D = 150$ improves these relative errors to $\text{Err}_E = 0.0001\%$, $\text{Err}_J = 0.001\%$, and $\text{Err}_S = 0.1\%$. As a benchmark, the solid red line in Fig. 2.4 (a), (b), and (e) corresponds to the functional Bethe ansatz result, which is valid only in the HD+LD phase and near $\lambda = 0$ in the S phase.

As $L \rightarrow \infty$, a number of interesting behaviors are observed, particularly at the interface between the S and HD+LD phases. In this region, the cumulant generating function transitions from having a finite negative slope to become nearly flat, signifying a transition into a low-current regime. This transition is marked by a continuous change in the current and an abrupt change in the current susceptibility, as shown in Fig. 2.4 (e). Note that here the current susceptibility has been scaled by $L^2$, instead of $L$, to show that this scaled measure does not diverge at $\lambda = 0$. We also see that the system becomes gapless due to the degeneracy of the high-density and low-density configurations. This degeneracy does not cause a spike in the susceptibility because the high-density and low-density states are of the same particle-hole/inversion symmetry while $\partial_\lambda \mathcal{W}_\lambda$ is odd under this symmetry. Instead, the growing susceptibility is controlled by the second gap (between $E_2$ and $(E_0, E_1)$) which also closes at this point.

The MPS representation also provides the state's full configurational information, enabling us to study the microscopic structure of the phases and quantities that are not derivatives of the cumulant generating function. Fig. 2.4 (c) and (d) show the steady-state density, $\rho$, and recurrent hopping, $K - |J|$, computed as specified in Eq. (2.15), as a function of the position in the lattice $x$ and the current bias $\lambda$. These density profiles correspond to those shown in Fig. 2.3(a), with the linear profile near $\lambda = 0$ corresponding to the shock phase. The HD+LD and MC phases can here be distinguished via the rate of recurrent hopping; particles and holes are spatially dispersed in the MC phase, allowing frequent opportunities to hop back and forth, as indicated by the finite observed recurrent hopping at $\lambda < 0$. When the transition is made into the HD+LD phase, the recurrent hopping drops to nearly zero in the lattice bulk, attributable to the lattice being nearly entirely filled or empty in this phase and thus providing few opportunities for recurrent hops.

An additional way to summarize the microscopic information (and the associated correlations in the system) is via the numerical entanglement entropy and spectrum ($S(i)$ and $\{s_m\}$ in Eq. (2.12)) which we measure at the middle of the lattice. These are plotted for the right eigenvector $|P^{\lambda}\rangle$ in Fig. 2.4 (g). There are two clear regions present in the numerical entanglement entropy, one corresponding to the MC phase, the other to the HD+LD phase. For the MC phase, the spectrum decays slowly, indicating that a relatively large bond dimension is required to accurately represent the given state. In the HD+LD phase, the numerical entanglement entropy is larger and appears to be exactly 1 ($\log_2 2$). The numerical entanglement spectrum shows that only two modes contribute, arising from the filled and empty configurations, indicating the state can be represented exactly by an MPS of bond dimension 2.

### Multi-lane ASEP model

We now consider a system comprised of multiple ASEP lanes, with particles that may hop vertically ($y$-direction) or horizontally ($x$-direction), where we will examine the unexplored interplay between vertical and horizontal currents that can generate new phase behaviour.

### Closed Multi-lane ASEP

A simple, but nontrivial, extension of the 1D ASEP into multiple lanes, as specified in Sec. 2.3, is to augment horizontal hopping and entry/exit parameters with equal vertical hopping rates $p_y = q_y = 1/2$ and no entry/exit at the vertical bounds, i.e. closed boundary conditions. To understand the phase behavior here, we again carried out DMRG calculations mapping out the behavior as a function of the longitudinal current bias $\lambda_x$ for fixed $p_x = 0.1$, with bond dimensions $D$ between 50 and 300 and with system widths and lengths of up to $L_y = 4$ and $L_x = 50$.

The resulting cumulant generating function, current, current susceptibility, and first excited state gap are displayed respectively in Fig. 2.5 (a), (b), (d) and (e) for the $L_y = 4$ ASEP (with the $L_y = [2, 3]$ results being essentially indistinguishable from these). A comparison between this figure and Fig. 2.4 shows no qualitative difference between the single lane and closed multi-lane ASEP. We can analyze the ground state MPS to confirm whether the microscopic configurations in the multi-lane system correspond to those seen in 1D.

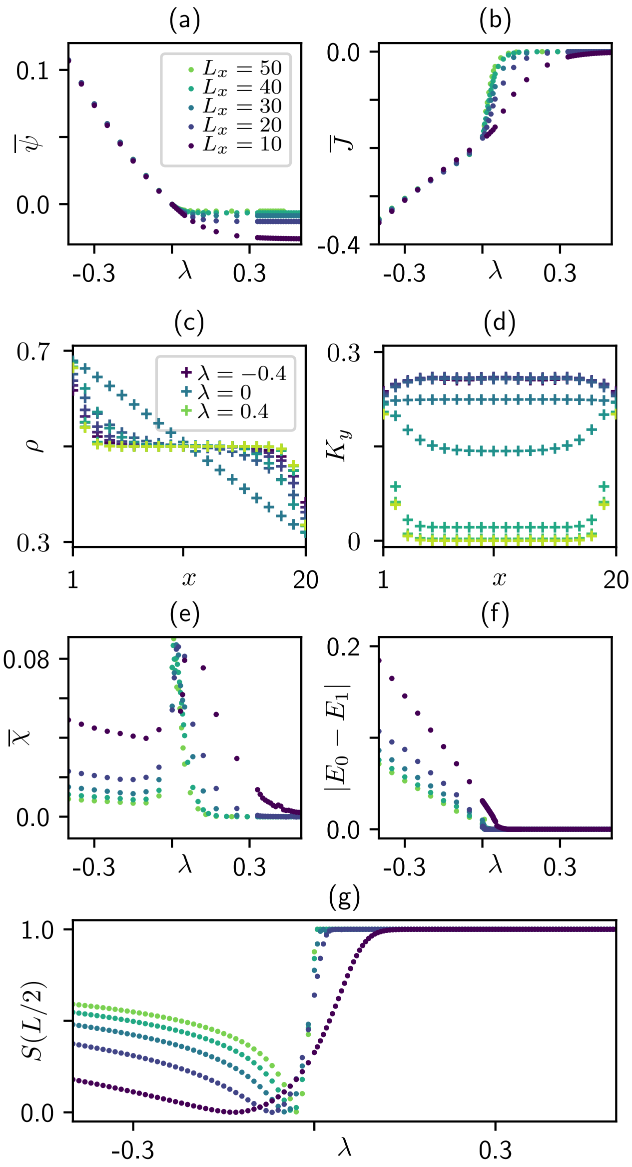

Figure 2.5: The behavior of the closed multi-lane ASEP showing the DMRG results for the normalized (a) CGF $\overline{\psi} = E_0/(L_x L_y)$, (b) current $\overline{J} = \partial_\lambda \psi/(L_x L_y)$, and (e) scaled current susceptibility $\overline{\chi} = \partial_\lambda^2 \psi/(L_x^2 L_y)$ as well as (f) the gap between the first and second largest eigenvalues $E_0$ and $E_1$ for the four lane systems with lengths up to $L_x = 50$. Plots (c) and (d) show the density $\rho$ and vertical hopping activity $K_y$ between lanes for a two-lane ASEP with $L_x = 20$. (g) Shows the numerical entanglement entropy $S$ of a bipartition of the system at the center bond as a function of $\lambda$.

Fig. 2.5 (c) and (d) show the behaviors of key observables as a function of $\lambda$. Using results from a two lane calculation, Fig. 2.5 (c) shows the density profile in one of the lanes as a function of $\lambda$, with the most notable point being the linear profile near $\lambda = 0$, indicative of a shock phase. The MC and HD+LD phases are again indistinguishable by their density profiles. As a means of distinguishing the two phases, we can use either the horizontal recurrent hopping rate profile (as done in 1D and not shown here) or the vertical activities between the two lanes as demonstrated in Fig. 2.5 (d). Here, the bulk vertical activity is near $K_y = 1/4$ per site when in the MC phase, supporting a microscopic structure where particles neighbor holes with probability $1/2$ and the probability of a vertical hop when such a configuration occurs is $p_y = q_y = 1/2$. After crossing the 1D ASEP phase boundary at $\lambda = 0$, the bulk vertical activity approaches zero, indicating that hops are prevented by an entirely full or empty lattice.

This picture is further supported by the profile of the numerical entanglement entropy for the two-lane ASEP shown in Fig. 2.5 (g), which again mimics the behavior seen for the 1D ASEP. We would usually expect the numerical entanglement entropy across the central cut to grow with the width of the system, which it appears to do in the MC phase. In the HD+LD, however, the numerical entanglement entropy is independent of the lattice width because the phase results from entirely empty and full configurations (where particle occupancy is perfectly correlated between the lanes in both configurations).

### Open Multi-lane ASEP

To quantify the effects of vertical boundaries on the horizontally biased dynamical phase behavior of this multi-lane ASEP, we further consider a vertically open multi-lane ASEP, where vertical entry/exit rates are $1/2$, as specified in Sec. 2.3. In these calculations, we employed DMRG to study the ASEP behavior as a function of the horizontal bias, $\lambda_x$, near $\lambda_x = 0$, with $p_x = 0.1$ for systems of up to length $L_x = 50$ with up to three lanes ($L_y = 3$) using a maximum bond dimension of $D = 50$.

The results are displayed in Fig. 2.6, with the cumulant generating function, current, current susceptibility, and first excited state gap being shown in subfigures (a), (b), (e), and (f). The per site macroscopic observables are nearly indistinguishable for the various system sizes, with the only noticeable difference caused by the requisite shifting of the point of GC symmetry as a function of system length. While in the closed multi-lane model the current detected a transition into the HD+LD phase, no

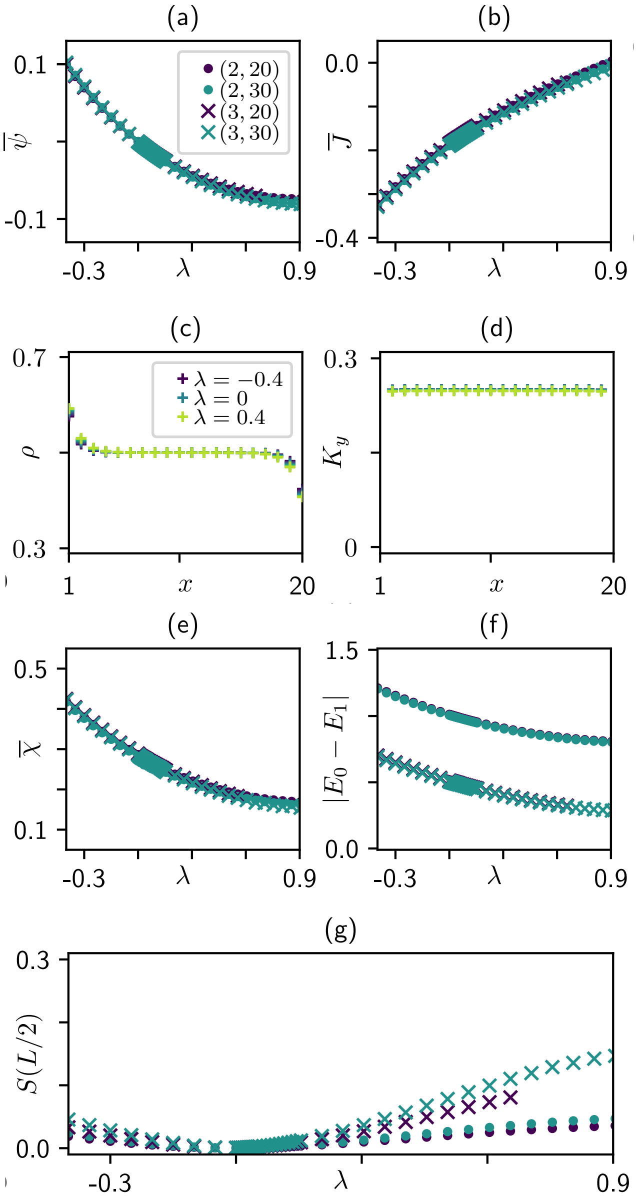

Figure 2.6: The behavior of the open multi-lane ASEP showing the DMRG results for the normalized (a) CGF $\overline{\psi} = E_0/(L_x L_y)$, (b) current $\overline{J} = \partial_\lambda \psi/(L_x L_y)$, and (e) current susceptibility $\overline{\chi} = \partial_\lambda^2 \psi/(L_x L_y)$ as well as (f) the gap between the first and second largest eigenvalues $E_0$ and $E_1$ for the two- and three-lane systems with lengths up to $L_x = 30$. Plots (c) and (d) show the density $\rho$ and vertical hopping activity $K_y$ between lanes for a two-lane ASEP with $L_x = 20$. (g) shows the numerical entanglement entropy $S$ of a bipartition of the system at the center bond as a function of $\lambda$.

such transition is apparent here.

This is further supported by a microscopic analysis for a lattice of size $2 \times 20$. The density and activity profiles are shown in Fig. 2.6 (c) and (d) as a function of $\lambda$. The $\lambda$ sweep show no changes in the behavior of the density and vertical activity. This is also true at $\lambda = 0$, where the phase transition would be expected to occur. While the steady-state number of hops between lattice sites does not seem to indicate any phase transition, we note that the desired low current behavior is created in a MC-like density profile by causing a small current to flow to the left in the bulk to counter the large current flowing to the right at the boundaries. This also illustrates a significant difference between the single-lane and multi-lane systems, namely that the steady-state current need not be spatially homogenous.

The lack of the phase transition in the open multi-lane system contrasts with the behavior of the closed multi-lane system. The behavior of the open model likely arises due to the availability of a vertical particle bath that enables rapid relaxation when jammed phases begin to form.

### 2.5 Conclusions

In conclusion, we have used MPS and DMRG to conduct a systematic study of the 1D and multi-lane ASEP with open horizontal boundary conditions under a current bias. With regards to the physics of the ASEP, we characterized the phase transition between the MC and HD+LD phases in the 1D system, showing agreement for the current cumulant generating function with known semi-analytic expressions and demonstrating changes in underlying microscopic structure via the steady-state density and activity profiles. We additionally found that the numerical entanglement entropy and spectrum, though different in physical meaning to their quantum counterparts, provide a global summary of the correlations in the system, identifying the sharp structure of the transition into the HD+LD phase. In the case of the multi-lane ASEP, we demonstrated that when biasing the current longitudinally with a simple choice of vertical hopping rates the development of the HD+LD phase occurs when vertical particle entry/exit is prohibited, but the phase boundary disappears entirely when this is reintroduced. This emphasizes the complex interplay between vertical and horizontal hopping parameters in this class of boundary driven processes and calls for the development of a more complete understanding of the multi-lane and fully 2D ASEP.

The TN methods used in this work are numerical realizations of the matrix product ansatz, which has long been used to produce semi-analytical solutions in driven lattice models. As we have shown, the flexibility of the numerical approach allows this framework to be applied to problems where analytical techniques are difficult to use, as demonstrated here with the multi-lane ASEP. In addition to providing a simple numerical route to compute large deviation functions, this approach also provide access to details of the underlying microscopic configurations, all without encountering the limitations of analytic methods or the sampling difficulties that plague Monte Carlo techniques. In addition, the success seen here and in other recent work [47] indicates the potential of more general tensor network approaches, which allow for a natural treatment of two-dimensional, three-dimensional, and thermodynamic lattice systems [63–65]. Currently, the TN methods remain challenging to apply to continuum systems. Continuing to further the application of tensor networks to nonequilibrium statistical models thus remains an exciting area of ongoing research.

# BIBLIOGRAPHY

[1] P. Helms, U. Ray, G. K.-L. Chan, *Physical Review E* **2019**, *100*, 022101,

[2] T. Chou, K. Mallick, R. Zia, *Reports on Progress in Physics* **2011**, *74*, 116601.

[3] B. Schmittmann, R. K. Zia, *Statistical Mechanics of Driven Diffusive Systems*, *Vol. 17*, Elsevier, **1995**, pages 3–214.

[4] F. Spitzer, *Advances in Mathematics* **1970**, *5*, 246–290.

[5] S. Klumpp, R. Lipowsky, *Journal of Statistical Physics* **2003**, *113*, 233–268.

[6] B. Widom, J. Viovy, A. Defontaines, *Journal de Physique I* **1991**, *1*, 1759–1784.

[7] K. Heckmann in *Biomembranes : Passive Permeability of Cell Membranes*, (Eds.: F. Kreuzer, J. F. G. Slegers), Springer US, Boston, MA, **1972**, pages 127–153.

[8] J. Krug, *Advances in Physics* **1997**, *46*, 139–282.

[9] A. Lazarescu, *Journal of Physics A: Mathematical and Theoretical* **2015**, *48*, 503001.

[10] M. Gorissen, A. Lazarescu, K. Mallick, C. Vanderzande, *Physical Review Letters* **2012**, *109*, 170601.

[11] B. Derrida, M. R. Evans, V. Hakim, V. Pasquier, *Journal of Physics A: Mathematical and Theoretical* **1993**, *26*, 1493.

[12] B. Derrida, J. L. Lebowitz, E. R. Speer, *Journal of Statistical Physics* **2003**, *110*, 775–810.

[13] B. Derrida, J. L. Lebowitz, *Physical Review Letters* **1998**, *80*, 209.

[14] N. Crampe, C. Finn, E. Ragoucy, M. Vanicat, *Journal of Physics A: Mathematical and Theoretical* **2016**, *49*, 375201.

[15] D. Bernard, T. Jin, *arXiv e-prints* **2019**, *arXiv:1904.01406*.

[16] A. Lazarescu, *Journal of Physics A: Mathematical and Theoretical* **2017**, *50*, 254004.

[17] A. K. Verma, A. K. Gupta, *Journal of Physics Communications* **2018**, *2*, 045020.

[18] Z. Ding, T. Liu, X. Lou, Z. Shen, K. Zhu, R. Jiang, B. Wang, B. Chen, *Physica A: Statistical Mechanics and its Applications* **2019**, *516*, 317–326.

[19] H. Yamamoto, D. Yanagisawa, K. Nishinari, *arXiv:1904.03988* **2019**.

[20] I. Dhiman, A. K. Gupta, *International Journal of Modern Physics C* **2018**, *29*, 1850037.

[21] Z.-J. Ding, S.-L. Yu, K. Zhu, J.-X. Ding, B. Chen, Q. Shi, X.-S. Lu, R. Jiang, B.-H. Wang, *Physica A: Statistical Mechanics and its Applications* **2018**, *492*, 1700–1714.

[22] U. Ray, G. K.-L. Chan, D. T. Limmer, *Physical Review Letters* **2018**, *120*, 210602.

[23] C. Pérez-Espigares, P. I. Hurtado, *arXiv:1902.01276* **2019**.

[24] M. Donsker, S. Varadhan, *Communications on Pure and Applied Mathematics* **1975**, *28*, 1–47.

[25] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer, **1985**.

[26] H. Touchette, *Physics Reports* **2009**, *478*, 1–69.

[27] Lasanta, A., Hurtado, Pablo I., Prados, A., *European Physical Journal E* **2016**, *39*, 35.

[28] T. Nemoto, R. L. Jack, V. Lecomte, *Physical Review Letters* **2017**, *118*, 115702.

[29] U. Ray, G. K.-L. Chan, D. T. Limmer, *The Journal of Chemical Physics* **2018**, *148*, 124120.

[30] T. Nemoto, F. Bouchet, R. L. Jack, V. Lecomte, *Physical Review E* **2016**, *93*, 062123.

[31] K. Klymko, P. L. Geissler, J. P. Garrahan, S. Whitelam, *Physical Review E* **2018**, *97*, 032123.

[32] D. Jacobson, S. Whitelam, *arXiv e-prints* **2019**, arXiv:1903.06098.

[33] T. Bodineau, B. Derrida, *Physical Review E* **2005**, *72*, 066110–8.

[34] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, *Journal of Physics A: Mathematical and Theoretical* **2009**, *42*, 075007.

[35] J. de Gier, F. H. L. Essler, *Physical Review Letters* **2011**, *107*, 010602.

[36] M. Gorissen, C. Vanderzande, *Journal of Physics A: Mathematical and Theoretical* **2011**, *44*, 115005.

[37] A. Prados, A. Lasanta, P. I. Hurtado, *Physical Review E* **2012**, *86*, 355–16.

[38] F. Carollo, J. P. Garrahan, I. Lesanovsky, *Physical Review B* **2018**, *98*, 094301.

[39] T. Brewer, S. R. Clark, R. Bradford, R. L. Jack, *Journal of Statistical Mechanics: Theory and Experiment* **2018**, *2018*, 053204.

[40] A. J. Schile, D. T. Limmer, *The Journal of Chemical Physics* **2018**, *149*, 214109.

[41] T. GrandPre, D. T. Limmer, *Physical Review E* **2018**, *98*, 060601.

[42] C. Y. Gao, D. T. Limmer, *arXiv e-prints* **2018**, arXiv:1812.01470.

[43] R. A. Blythe, M. R. Evans, *Journal of Physics A: Mathematical and Theoretical* **2007**, *40*, R333.

[44] E. Carlon, M. Henkel, U. Schollwöck, *The European Physical Journal B-Condensed Matter and Complex Systems* **1999**, *12*, 99–114.

[45] M. Gorissen, J. Hooyberghs, C. Vanderzande, *Physical Review E* **2009**, *79*, 020101.

[46] T. H. Johnson, S. R. Clark, D. Jaksch, *Physical Review E* **2010**, *82*, 036702.

[47] M. C. Bañuls, J. P. Garrahan, *Physical Review Letters* **2019**, *123*, 200601.

[48] S. Prolhac, *arXiv:0904.2356* **2009**.

[49] A. Lazarescu, V. Pasquier, *Journal of Physics A: Mathematical and Theoretical* **2014**, *47*, 295202.

[50] H. Touchette, *arXiv:1106.4146* **2011**.

[51] U. Schollwöck, *Annals of Physics* **2011**, *326*, 96–192.

[52] J. L. Lebowitz, H. Spohn, *Journal of Statistical Physics* **1999**, *95*, 333–365.

[53] B. Derrida, T. Sadhu, *arXiv:1807.06543* **2018**.

[54] B. Derrida, T. Sadhu, *arXiv:1905.07175* **2019**.

[55] G. K.-L. Chan, T. Van Voorhis, *The Journal of Chemical Physics* **2005**, *122*, 204101.

[56] M. B. Hastings, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, *2007*, P08024.

[57] A. Kitaev, J. Preskill, *Physical Review Letters* **2006**, *96*, 110404.

[58] F. Pollmann, A. M. Turner, E. Berg, M. Oshikawa, *Physical Review B* **2010**, *81*, 064439.

[59] G. Gallavotti, E. G. D. Cohen, *Physical Review Letters* **1995**, *74*, 2694.

[60] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *Journal of Statistical Physics* **2002**, *107*, 635–675.

[61] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, C. Landim, *Reviews of Modern Physics* **2015**, *87*, 593.

[62] S. Prolhac, K. Mallick, *Journal of Physics A: Mathematical and Theoretical* **2008**, *41*, 175002.

[63] F. Verstraete, V. Murg, J. I. Cirac, *Advances in Physics* **2008**, *57*, 143–224.

[64] R. Orús, *Annals of Physics* **2014**, *349*, 117–158.

[65] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, R. Orús, *Physical Review B* **2015**, *92*, 035142.

## *Chapter 3*

# DYNAMICAL PHASE TRANSITIONS IN A 2D CLASSICAL NONEQUILIBRIUM MODEL VIA 2D TENSOR NETWORKS

We demonstrate the power of two-dimensional tensor networks for obtaining large deviation functions of dynamical observables in a classical nonequilibrium setting. Using these methods, we analyze the previously unstudied dynamical phase behavior of the fully two-dimensional asymmetric simple exclusion process with biases in both the $x$ and $y$ directions. We identify a dynamical phase transition, from a jammed to a flowing phase, and characterize the phases and the transition, with an estimate of the critical point and exponents.

This chapter is adapted from our previously published article:

[1] P. Helms, G. K.-L. Chan, *Physical Review Letters* **2020**, *125*, 140601,

## 3.1 Introduction

Large deviation theory (LDT) has emerged as a powerful framework for studying the fluctuations of macroscopic dynamical observables in classical nonequilibrium systems [2–6]. Reminiscent of equilibrium statistical mechanics, where ensembles of configurations are organized by their macroscopic properties, such as temperature or energy, LDT prescribes the grouping of trajectories into ensembles based on their dynamical or static macroscopic properties, such as current or density. This approach allows for the definition of dynamical partition functions, derivatives of which are the mathematical analogs to entropy and free energy, named large deviation functions (LDFs), which encode the statistics of dynamical observable fluctuations. As in equilibrium systems, these are critical for identifying and characterizing phase transitions, particularly those which occur in the space of trajectories, called dynamical phase transitions (DPTs) [6].

The success of LDT has been accompanied by the development of numerical methods for computing LDFs, with significant emphasis and progress centered in sophisticated sampling techniques [5, 7–13]. Alternatively, the matrix product ansatz, a powerful analytical representation of nonequilibrium steady states [14–16], foreshadowed the recent success of numerical tensor network (TN) algorithms. In particular, calculations using matrix product states (MPS), the one-dimensional TN that underpins the density matrix renormalization group (DMRG) algorithm [17], provide a noiseless alternative to sampling methods. As demonstrated in the recent applications to DPTs in kinetically constrained and driven diffusive models [18–22], the MPS provides a remarkably compact representation of nonequilibrium steady states.

While the TN approach is promising, the use of the MPS, which only efficiently encodes correlations in one dimension, limits the study of higher-dimensional problems [23], Consequently, LDF computations beyond one dimension have relied on Monte Carlo methods [24–27]. In this letter, we demonstrate how an inherently two-dimensional TN, the projected entangled pair state (PEPS) [28–31], serves as an efficient ansatz to determine LDFs in two-dimensional nonequilibrium lattice problems.

We use this approach to obtain new insights into the fully two-dimensional asymmetric simple exclusion process (ASEP). In one-dimensional, the ASEP has become a paradigmatic model of nonequilibrium behavior frequently employed to understand important physical systems and phenomena including surface growth [32, 33],

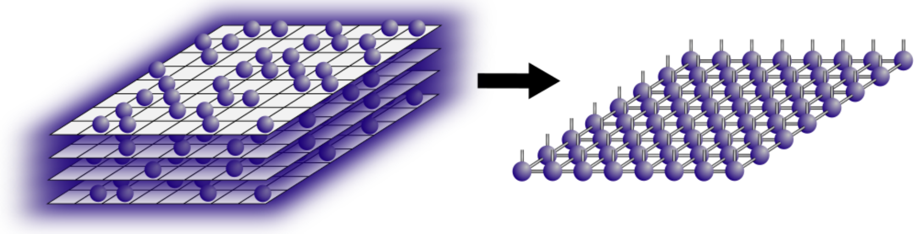

Figure 3.1: A stack of possible configurations of the two-dimensional ASEP (left), representing all possible configuration probabilities, is stored as a two-dimensional PEPS, whose TN diagram is shown on the right. Contracting all auxiliary bonds gives the probability of all possible lattice configurations.

molecular motors [34–36], and traffic flow [37]. The two-dimensional ASEP is of similarly wide interest, but it has remained poorly characterized [38–43], especially with regards to its dynamical phase behavior, which is unknown except in the periodic, weakly asymmetric limit [24]. We show that two-dimensional TN now allow us to shed light on the general two-dimensional ASEP, by computing detailed observables along a line in the dynamical phase diagram. In so doing, we find and characterize a hitherto unobserved DPT between jammed and flowing phases.

### 3.2 Large Deviation Theory and Projected Entangled Pair States

We begin with a short overview of relevant theory and methods associated with LDT, TNs, and PEPS. More comprehensive treatments of all three topics are provided in recent reviews and methodological papers [6, 29, 31].

A Markovian nonequilibrium system's time evolution is governed by a master equation, $\partial_t|P_t\rangle = \mathcal{W}|P_t\rangle$, where vector $|P_t\rangle$ represents the configurational probabilities at time $t$ and the generator, $\mathcal{W}$, dictates the transition rates between configurations. At steady-state, the time-averaged current vector, $\bar{\boldsymbol{J}} = \boldsymbol{J}/t$ obeys a large deviation principle, $P(\bar{\boldsymbol{J}}) \approx e^{-t\phi(\bar{\boldsymbol{J}})}$, as does its moment generating function, $Z(\boldsymbol{\lambda}) = \langle e^{-\boldsymbol{\lambda}\bar{\boldsymbol{J}}}\rangle \approx e^{-t\psi(\boldsymbol{\lambda})}$, indicating that the probability of observing all but the most likely current decays exponentially with averaging time. The rate function (RF), $\phi(\bar{\boldsymbol{J}})$, defines the probability of a given current, and $\psi(\boldsymbol{\lambda})$ is the scaled cumulant generating function (SCGF), whose derivatives at $\boldsymbol{\lambda} = 0$ give the cumulants of the current.

Performing a tilting of the generator, $\mathcal{W} \rightarrow \mathcal{W}^{(\lambda)}$, effectively weights trajectories according to their currents, by scaling all forward (backward) hopping terms by

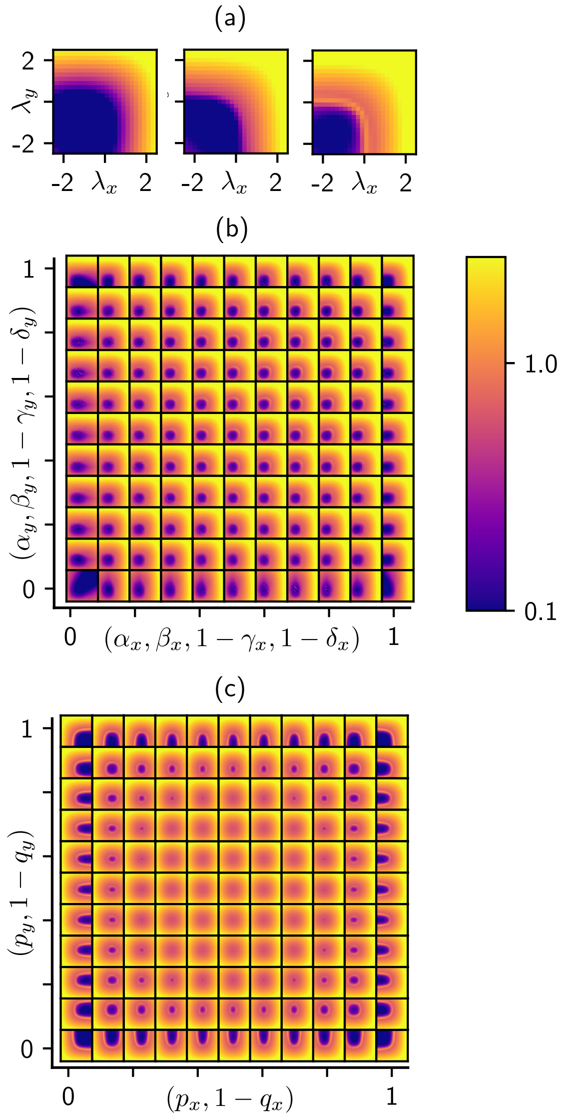

Figure 3.2: A mapping of the mean field dynamical phase diagram of the two-dimensional ASEP with (a) showing the SCGF (top), current (middle), and current susceptibility (bottom) as a function of bias at one point in the physical phase space, while (b) and (c) respectively show plots of the current susceptibility as a function of bias for a bulk biased and a boundary biased two-dimensional ASEP. For (a), $p_{x,y} = 1 - q_{x,y} = 1$ with boundary terms at $1/2$ and current biases, $\lambda_x, \lambda_y \in [-2.5, 2.5]$; we can see the transition between the jammed (dark) and flowing (bright) phases. In (b), bulk rates are fixed at $p_{x,y} = 1 - q_{x,y} = 0.9$ while sweeping over a subset of boundary rates ($\alpha_{x,y} = \beta_{x,y} = 1 - \gamma_{x,y} = 1 - \delta_{x,y}$). In (c), all boundary terms are set to $1/2$ and we sweep over bulk hopping rates ($p_{x,y}$,$q_{x,y}$). Each subplot in (b) and (c) sweeps over current biases $\lambda_x, \lambda_y \in [-2.5, 2.5]$.

$e^{-\lambda}$ ($e^{\lambda}$), making $\mathcal{W}^{(\lambda)}$ non-Markovian and non-Hermitian. A central finding in LDT dictates that the largest eigenvalue of the tilted generator is the SCGF, i.e. $\mathcal{W}^{(\lambda)}|P^{(\lambda)}\rangle = \psi(\lambda)|P^{(\lambda)}\rangle$. Furthermore, the corresponding left and right eigenvectors detail trajectory characteristics associated with particular fluctuations. For example, the time averaged local density associated with a fluctuation is $\rho_i = \langle P^{(\lambda)}|n_i|P^{(\lambda)}\rangle/\langle P^{(\lambda)}|P^{(\lambda)}\rangle$, where $n_i$ is the particle number operator acting on site $i$ and $\langle P^{(\lambda)}|$ and $|P^{(\lambda)}\rangle$ are the left and right eigenvectors.

The PEPS TN ansatz is a intuitive representation of the approximate eigenstates of the tilted generator and a diagrammatic representation of this ansatz is shown on the right side of Figure 3.1, where a tensor is allocated for each lattice site. Diagrammatically, each tensor is represented as a ball with tensor indices corresponding to lines connected to the ball. The vertical indices, called the physical bonds, correspond to the local state space of the system and are of size $d$, which is the local state dimension (for hard core particles $d = 2$, corresponding to an empty or occupied site). Additionally, nearest neighbor tensors are connected by indices, called auxiliary bonds, of size $D$, enabling information transfer between sites. This results in a lattice of rank five bulk tensors $\mathcal{T}^{[x,y]}_{ijklm}$ of size $(d, D, D, D, D)$. The size of the auxiliary bonds, called the bond dimension, controls the accuracy of the ansatz by truncating the considered Hilbert space and for sufficiently large $D$ the ansatz is exact. While $D$ must grow exponentially with the size of the lattice to accurately represent arbitrary states, in practice, many states are accurately captured by a PEPS with finite $D$ even as the lattice grows. By contracting over all auxiliary bonds, the eigenstate of the tilted generator is recovered, thus the mapping in Figure 3.1 roughly illustrates how the set of all configurational probabilities are stored as a PEPS.

The development of appropriate PEPS optimization methods for quantum many body problems is an active area of research [44–47]. For this work, we simply adapt many of the most successful standard techniques to the non-equilibrium master equation setting. Using the time-evolving block decimation approach [29, 48], we integrate the tilted master equation forwards in time, giving $|P_t^{(\lambda)}\rangle = e^{t\mathcal{W}^{(\lambda)}}|P_0^{(\lambda)}\rangle$. We apply the time evolution operator to the initial PEPS via its Suzuki-Trotter decomposition into local gates, $e^{t\mathcal{W}^{(\lambda)}} \approx \left(e^{\delta t\mathcal{W}^{(\lambda)}_{i,i+1}}\right)^{t/\delta t}$, and iterate this application until convergence to the steady-state. The bond dimension between two sites grows after the application of the gate, thus an alternating least squares approach is used to compress the tensors back to dimension $D$ [31]. The alternating least squares algorithm uses information from all the other tensors which are contracted into an

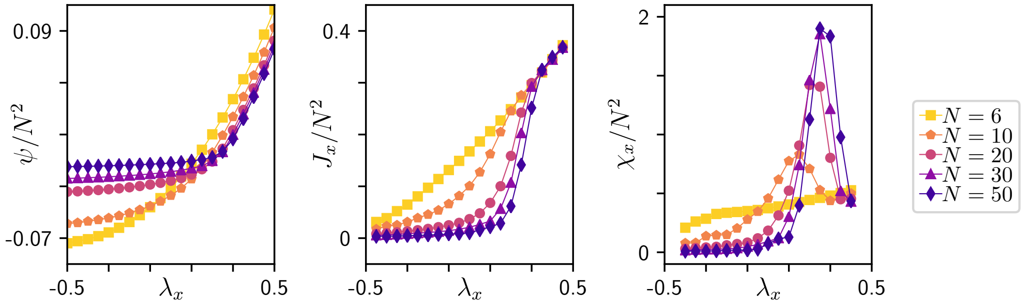

Figure 3.3: PEPS calculation results analyzing the phase transition along a line in the dynamical phase space of the two-dimensional ASEP. From left to right, we show the per site SCGF $\psi(\lambda_x, \lambda_y)/N^2$, horizontal current $J_x/N^2$, and horizontal current susceptibility $\chi_x/N^2$ at $\lambda_y = -1/2$ with $\lambda_x \in [-1/2, 1]$. Each line corresponds to a system size $N \in [6, 10, 20, 30, 50]$.

approximate environment using the single-layer boundary method [49] and tensor reduction [50, 51]. The accuracy of the environment is then determined by an additional parameter, $\chi$, which corresponds to the bond dimension of a boundary MPS. Like $D$, $\chi$ must also be increased to converge to the exact stationary state. In practice, because the environment computation is expensive, we can first determine an approximate stationary state via the "simple update" algorithm where no environment is used [52]; then $D$ and $\chi$ are increased in subsequent time evolution steps using the full environment information ("full update" algorithm [48]) while $\delta t$ is also decreased to reduce the Suzuki-Trotter error.

### 3.3 Model: Two-Dimensional Asymmetric Simple Exclusion Process

The two-dimensional ASEP, Figure 3.1 (left), takes place on a square $N \times N$ lattice, where each site may be occupied by a particle or empty. Particles stochastically hop into vacant nearest-neighbor lattice sites in the right (up) and left (down) directions at rates $p_x$ ($p_y$) and $q_x$ ($q_y$) respectively. At the {left, bottom, right, top} boundaries, particles are inserted at rates $\{\alpha_x, \alpha_y, \delta_x, \delta_y\}$, and removed at rates $\{\gamma_x, \gamma_y, \beta_x, \beta_y\}$. Additionally, as detailed in the previous section, we utilize a current bias in both directions, $\boldsymbol{\lambda} = (\lambda_x, \lambda_y)$, to probe the trajectory phase space. The tilted generator is built from hopping operators $\boldsymbol{o}^{\text{hop}}_{i,j} = r_{i,j}(e^{\lambda_{i,j}}\boldsymbol{a}_i\boldsymbol{a}^\dagger_j - \boldsymbol{n}_i\boldsymbol{v}_j)$ and similarly defined insertion and removal operators, where $r_{i,j}$ is the hopping rate from site $i$ to $j$ and $\boldsymbol{a}_i$, $\boldsymbol{a}^\dagger_i$, $\boldsymbol{n}_i$, and $\boldsymbol{v}_i$ are respectively annihilation, creation, particle number and vacancy operators. Because hopping occurs only between nearest neighbor sites, the full tilted generator, $\mathcal{W}^{(\lambda)}$, then decomposes naturally into nearest neighbor gates. At

$\lambda_{i,j} = 0, \forall(i,j)$, the system undergoes its typical dynamics, otherwise the biasing allows for probing of rare trajectories.

## 3.4 Results

We first probed for the existence of a DPT in the two-dimensional ASEP by performing mean field (MF) computations of the SCGF on an $8 \times 8$ lattice in two subsets of the phase space, with results shown in Figure 3.2. In Figure 3.2(a) we show, from top to bottom, the per site SCGF, total current, and current susceptibility at $p_{x,y} = 1 - q_{x,y} = 1$ with $\alpha_{x,y} = \beta_{x,y} = \gamma_{x,y} = \delta_{x,y} = 1/2$. and current biases sweeping over $\lambda_x, \lambda_y \in [-2.5, 2.5]$. In the bottom left of these plots, we see a low-current regime materialize, where the SCGF and current flattens, bounded by a small peak in the susceptibility (the thin bright line between the purple and orange regions).

To further explore where this low-current phase materializes, Figure 3.2 (b) and (c) contain subplots at various points in the rate parameter space, each showing the per site current susceptibility as a function of $\lambda_{x,y} \in [-2.5, 2.5]$. (b) explores boundary effects, sweeping boundary terms with $\alpha_{x,y} = \beta_{x,y} = 1 - \gamma_{x,y} = 1 - \delta_{x,y}$ and maintaining asymmetric interior rates $p_{x,y} = 1 - q_{x,y} = 0.9$ while (c) probes the effect of bulk hopping rates, sweeping interior hopping rates while holding boundary terms at $\alpha_{x,y} = \beta_{x,y} = \gamma_{x,y} = \delta_{x,y} = 1/2$.

Phase transitions can be marked by a peak in the current susceptibility, as seen in Figure 3.2 (a). In Figure 3.2 (c) this becomes visible at sufficiently high biases ($\approx p_x > 0.8$), again accompanied by a region of distinctly low current. This aligns with the known behavior of the one-dimensional ASEP, where a DPT is observed except when $p_x = q_x = 1/2$, which corresponds to the Symmetric Simple Exclusion Process (SSEP). Furthermore, intuition from the one-dimensional ASEP would further predict a DPT to appear for low biases in the thermodynamic limit. For the boundary biased results, Figure 3.2 (b), we observe the boundary rates to have little effect, except at extreme values, where the location of the DPT becomes distorted due to no insertion or removal at a boundary.

Selecting a line within the phase space covered in Figure 3.2 (c) at $p_{x,y} = 1 - q_{x,y} = 0.9$ $\lambda_y = -1/2$ with $\lambda_x \in [-1/2, 1/2]$, we carried out PEPS calculations on $N \times N$ lattices with $N \in \{6, 10, 20, 30, 50\}$ to probe the DPT's finite size behavior. Here, we used $D \in [2, 8]$ and $\chi = 80$ while systematically reducing $\delta t \in [10^{-1}, 10^{-4}]$. Figure 3.3 displays key results from these calculations in support of the existence of a DPT.

There, the left plot shows the SCGF for the $\lambda_x$ sweep, with the flattening of the curve for large systems on the left side of the plot indicating a low-current region. The horizontal current $J_x$ and current susceptibility $\chi_x$, shown in the center and right plots, are computed via central difference numerical differentiation with respect to $\lambda_x$; while they can also be computed via contractions with the left and right PEPS eigenstates of $\mathcal{W}^{(\lambda)}$, for the largest systems this can be numerically challenging and requires well-converged left and right states.

In all plots, we see two distinct regions, indicative of a DPT. Moving from right to left, we see the emergence of a low-current phase at $\approx \lambda_x = 1/4$, where both $J_x$ and $J_y$ (not shown) are small. The transition becomes sharper as the size of the lattice increases, as seen by the increasingly large peaks in current susceptibility, substantiating the existence of a second-order DPT between the jammed and flowing phases. Furthermore, the most likely configurations in the flowing phase are those where particles are evenly distributed throughout the lattice, while in the low-current phase, those most likely are entirely filled, jamming flow in the bulk.

To gauge the accuracy of these results, Figure 3.4 displays the convergence of the SCGF for calculations with $N = 20$. Here, the SCGF is computed from the right and left eigenstates, $\psi_R$ and $\psi_L$, in the jammed (top) and flowing (bottom) phases with $\lambda = -0.5$ and $\lambda = 0.5$ respectively. Shaded regions correspond to $D$, starting with mean field results on the left and increasing to the right, where within each shaded region, the accuracy is improved by increasing $\chi$. Each computation was performed independently, doing the "full update" procedure from a random initial state, decreasing the time step sizes from $\delta t = 0.5$ to $\delta t = 0.01$. In addition to the convergence with bond dimension, the difference between the estimate of the eigenvalue from the left and right eigenvectors serves as an additional check on accuracy.

We find that with very modest computational resources ($D = 3$, $\chi = 100$), the SCGF easily converges to approximately three significant digits, significantly greater than MF results. It is also clear that unlike in quantum systems, where the variational principle prevents the ground state energy from going below the exact ground state energy, our computed SCGF can go above and below the exact value. Also notable is that calculations in the jammed regime converge to more accurate results at a low bond dimension than those in the flowing region. Without an initial set of sufficiently large time steps, we found that calculations in the jammed phase tend to converge to local minima.

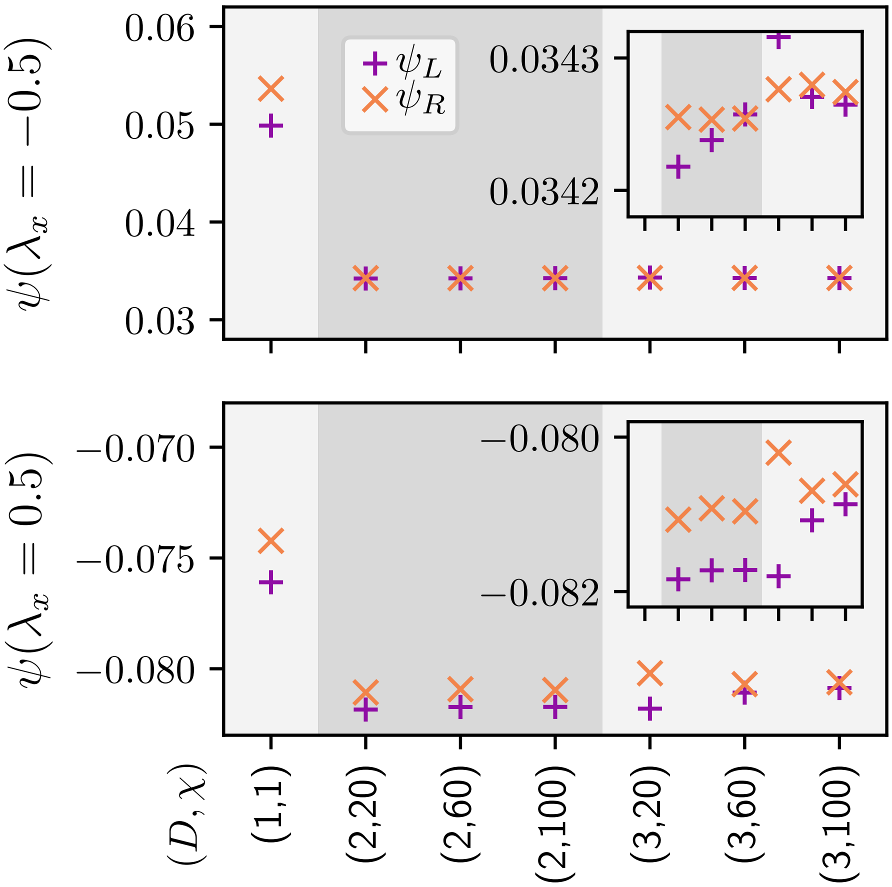

Figure 3.4: The convergence of PEPS calculations, showing the SCGF, computed as the left and right eigenvalue of the tilted generator, $\psi_L$ and $\psi_R$, for a 20 lattice as a function of the bond dimension $D$ (shaded) and the boundary bond dimension $\chi$ (labeled as $(D, \chi)$). The top (bottom) plot corresponds to results in the jammed (flowing) phase at $\lambda = -0.5$ ($\lambda = 0.5$). The insets provide magnified results to illustrate the extent of convergence.

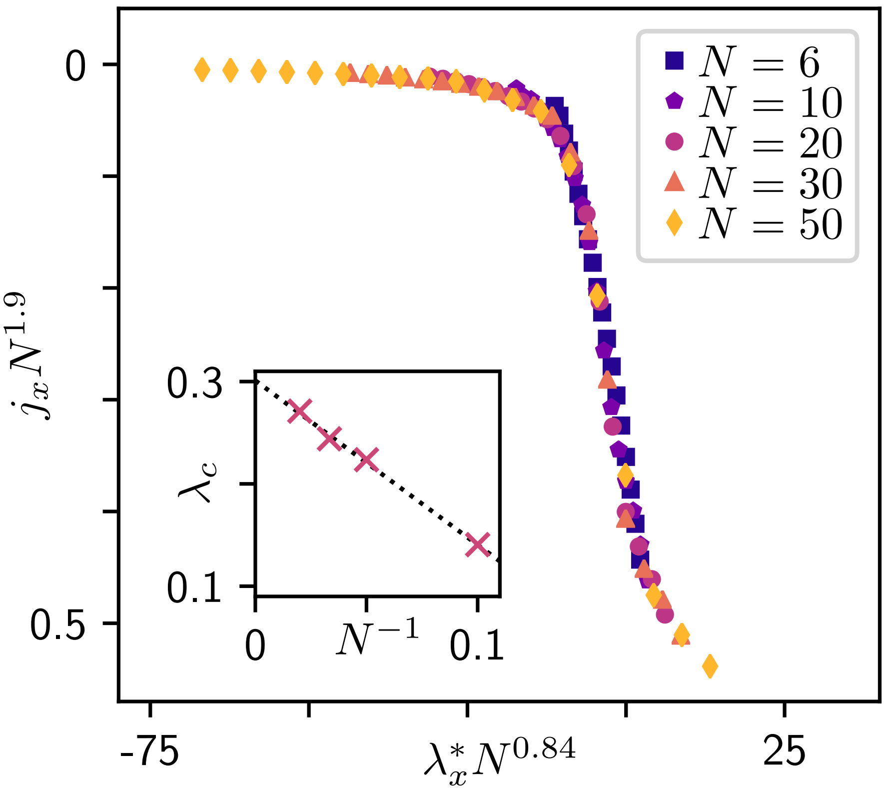

Figure 3.5: Scaling plot of the transition between the flowing and jammed phases, showing the collapse of the per site horizontal current as a function of the reduced horizontal bias, $\lambda_x^*$. The inset plot shows a finite size extrapolation to estimate of the critical point $\lim_{N\to\infty} \lambda_c = 0.30$, with $\lambda_c(N)$ by fitting a quadratic function to the three largest points in the susceptibility peaks for each $N$ in Figure 3.3

Last, we can perform a finite size scaling analysis of the observed transition to extract the critical exponents in the thermodynamic limit. Because the system sizes studied are limited to a linear dimension of $N \leq 50$, the results retain some finite-size error, though we expect that future work performing PEPS calculations on larger lattices, possible because PEPS calculation costs grow linearly with system size, or adapting infinite PEPS algorithms [31] could further refine these estimates. The scaling relation for the per site horizontal current is $j_x(\lambda_x^*, N) = N^d f(\lambda_x^* N^c)$, where $d$ and $c$ are critical exponents, $f$ is the scaling function, and $\lambda_x^*$ is analogous to a reduced temperature, i.e. $\lambda_x^* = (\lambda - \lambda_c)/\lambda_c$. The inset of Figure 3.5 shows a linear extrapolation of the location of the susceptibility peaks in Figure 3.3 to determine the critical point to be $\lim_{N\to\infty} \lambda_c = 0.30$. The critical parameters are then computed via numerical data collapse [53], giving $d = -1.9 \pm 0.1$ and $c = 0.84 \pm 0.1$ with

Figure 3.5 showing the resulting scaling plot, which displays good data collapse.

## 3.5 Conclusion

We have provided the first insights into the dynamical phase behavior of the fully two-dimensional ASEP, finding evidence for a dynamical phase transition between a flowing and a jammed phase, as detected by a sharp change in the current in the horizontal and vertical directions. We have also demonstrated how two-dimensional tensor networks, in particular the PEPS ansatz, can be used to compute large deviation functions in classical nonequilibrium systems, characterize nonequilibrium phases, and obtain critical exponents. This is a natural extension of the success of one-dimensional tensor network methods in this field and provides significant promise for the future use of TNs in coordination with LDT. Because numerical methods based on PEPS are relatively young, continued progress is likely, and we expect such higher-dimensional TNs to become standard tools in the study of nonequilibrium classical statistical mechanics.

# BIBLIOGRAPHY

[1] P. Helms, G. K.-L. Chan, *Physical Review Letters* **2020**, *125*, 140601,

[2] B. Derrida, *Journal of Statistical Mechanics: Theory and Experiment* **2007**, *2007*, P07023.

[3] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, *Journal of Physics A: Mathematical and Theoretical* **2009**, *42*, 075007.

[4] A. Prados, A. Lasanta, P. I. Hurtado, *Physical Review Letters* **2011**, *107*, 140601.

[5] U. Ray, G. K.-L. Chan, D. T. Limmer, *The Journal of Chemical Physics* **2018**, *148*, 124120.

[6] H. Touchette, *Physics Reports* **2009**, *478*, 1–69.

[7] U. Ray, G. K.-L. Chan, D. T. Limmer, *Physical Review Letters* **2018**, *120*, 210602.

[8] T. Nemoto, F. Bouchet, R. L. Jack, V. Lecomte, *Physical Review E* **2016**, *93*, 062123.

[9] K. Klymko, P. L. Geissler, J. P. Garrahan, S. Whitelam, *Physical Review E* **2018**, *97*, 032123.

[10] T. Nemoto, R. L. Jack, V. Lecomte, *Physical Review Letters* **2017**, *118*, 115702.

[11] U. Ray, G. K. Chan, *arXiv:1909.11283* **2019**.

[12] G. Margazoglou, L. Biferale, R. Grauer, K. Jansen, D. Mesterházy, T. Rosenow, R. Tripiccione, *Physical Review E* **2019**, *99*, 053303.

[13] A. Das, D. T. Limmer, *The Journal of Chemical Physics* **2019**, *151*, 244123.

[14] R. A. Blythe, M. R. Evans, *Journal of Physics A: Mathematical and Theoretical* **2007**, *40*, R333.

[15] S. Prolhac, M. R. Evans, K. Mallick, *Journal of Physics A: Mathematical and Theoretical* **2009**, *42*, 165004.

[16] B. Derrida, M. R. Evans, V. Hakim, V. Pasquier, *Journal of Physics A: Mathematical and Theoretical* **1993**, *26*, 1493.

[17] U. Schollwöck, *Annals of Physics* **2011**, *326*, 96–192.

[18] M. Gorissen, A. Lazarescu, K. Mallick, C. Vanderzande, *Physical Review Letters* **2012**, *109*, 170601.

[19] M. Gorissen, C. Vanderzande, *Journal of Physics A: Mathematical and Theoretical* **2011**, *44*, 115005.

[20] M. Gorissen, J. Hooyberghs, C. Vanderzande, *Physical Review E* **2009**, *79*, 020101.

[21] M. C. Bañuls, J. P. Garrahan, *Physical Review Letters* **2019**, *123*, 200601.

[22] P. Helms, U. Ray, G. K.-L. Chan, *Physical Review E* **2019**, *100*, 022101,

[23] E. M. Stoudenmire, S. R. White, *Annual Review of Condensed Matter Physics* **2012**, *3*, 111–128.

[24] N. Tizón-Escamilla, C. Pérez-Espigares, P. L. Garrido, P. I. Hurtado, *Physical Review Letters* **2017**, *119*, 090602.

[25] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, F. van Wijland, *Physical Review Letters* **2007**, *98*, 195702.

[26] L. O. Hedges, R. L. Jack, J. P. Garrahan, D. Chandler, *Science* **2009**, *323*, 1309–1313.

[27] D. Chandler, J. P. Garrahan, *Annual Review of Physical Chemistry* **2010**, *61*, 191–217.

[28] F. Verstraete, M. M. Wolf, D. Perez-Garcia, J. I. Cirac, *Physical Review Letters* **2006**, *96*, 220601.

[29] M. Lubasch, J. I. Cirac, M.-C. Banuls, *New Journal of Physics* **2014**, *16*, 033014.

[30] R. Orús, *Annals of Physics* **2014**, *349*, 117–158.

[31] H. N. Phien, J. A. Bengua, H. D. Tuan, P. Corboz, R. Orús, *Physical Review B* **2015**, *92*, 035142.

[32] J. Krug, *Advances in Physics* **1997**, *46*, 139–282.

[33] G. Ódor, B. Liedke, K.-H. Heinig, *Physical Review E* **2009**, *79*, 021125.

[34] S. Klumpp, R. Lipowsky, *Journal of Statistical Physics* **2003**, *113*, 233–268.

[35] T. Chou, G. Lakatos, *Physical Review Letters* **2004**, *93*, 198101.

[36] R. Lipowsky, Y. Chai, S. Klumpp, S. Liepelt, M. J. Müller, *Physica A* **2006**, *372*, 34–51.

[37] A. Schadschneider, *Physica A* **2000**, *285*, 101–120.

[38] F. J. Alexander, Z. Cheng, S. A. Janowsky, J. L. Lebowitz, *Journal of Statistical Physics* **1992**, *68*, 761–785.

[39] Z.-J. Ding, S.-L. Yu, K. Zhu, J.-X. Ding, B. Chen, Q. Shi, X.-S. Lu, R. Jiang, B.-H. Wang, *Physica A: Statistical Mechanics and its Applications* **2018**, *492*, 1700–1714.

[40] N. Singh, S. M. Bhattacharjee, *Physical Letters A* **2009**, *373*, 3113–3117.

[41] H.-T. Yau, *Annals of Mathematics* **2004**, 377–405.

[42] M. Tamm, S. Nechaev, S. N. Majumdar, *Journal of Physics A: Mathematical and Theoretical* **2010**, *44*, 012002.

[43] B. Schmittmann, K. Hwang, R. Zia, *EPL (Europhysics Letters)* **1992**, *19*, 19.

[44] P. Corboz, *Physical Review B* **2016**, *94*, 035133.

[45] L. Vanderstraeten, J. Haegeman, F. Verstraete, *Physical Review B* **2019**, *99*, 165121.

[46] M. J. O'Rourke, G. K. Chan, *Physical Review B* **2020**, *101*, 205142.

[47] R. Haghshenas, M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2019**, *100*, 054404.

[48] M. Lubasch, J. I. Cirac, M.-C. Banuls, *Physical Review B* **2014**, *90*, 064425.

[49] Z. Y. Xie, H. J. Liao, R. Z. Huang, H. D. Xie, J. Chen, Z. Y. Liu, T. Xiang, *Physical Review B* **2017**, *96*, 045128.

[50] P. Corboz, R. Orús, B. Bauer, G. Vidal, *Physical Review B* **2010**, *81*, 165104.

[51] C. Pineda, T. Barthel, J. Eisert, *Physical Review A* **2010**, *81*, 050303.

[52] H.-C. Jiang, Z.-Y. Weng, T. Xiang, *Physical Review Letters* **2008**, *101*, 090603.

[53] S. M. Bhattacharjee, F. Seno, *Journal of Physics A: Mathematical and Theoretical* **2001**, *34*, 6375.

## *Chapter 4*

# DMRG-STYLE OPTIMIZATION OF TWO-DIMENSIONAL TENSOR NETWORKS WITH LONG-RANGE INTERACTIONS

The success of the density matrix renormalization group (DMRG), underpinned by the matrix product state (MPS) ansatz, triggered rapid growth in the development and application of tensor networks for the simulation of both quantum and classical systems. In two dimensions, the projected entangled pair state (PEPS), a generalization of the MPS, has proven widely effective in various contexts. A significant impediment to continued progress, however, is the difficulty of creating compact representations of operators with long-range interactions. While a handful of proposals have emerged, it remains unclear how well these operator representations work in a practical computation. Here, we present three representations of long-range operators for two-dimensional tensor networks, namely comb operators, comb-like projected entangled pair operators (PEPOs), and sums of matrix product operators (MPOs). These operator representations are tailored for use in DMRG-style optimizations. We evaluate their efficiency and stability in practice by computing the ground state of a coulombic Heisenberg model. We find that while the comb operators and sums of MPOs are competitive with one another and provide viable operator representations, the comb-like PEPO, though most efficient in representation, is the least stable and efficient in practice.

## 4.1 Introduction

During the past two decades, tensors networks (TNs) have been proven to be powerful representations of complex states in both quantum [1–3] and classical systems [4–7], with recent work extending their utility into machine learning algorithms [8–12] and quantum computer simulation [13, 14] [15]. In one dimension, TN algorithm development has converged upon several widely successful routines such as the density matrix renormalization group (DMRG) algorithm for determining quantum ground states [1] and extensions to time evolution [16, 17] and thermal states [18][19], which rely heavily upon the one-dimensional tensor network state and operator forms, the matrix product state (MPS) and matrix product operator (MPO). Properties of these one-dimensional TNs such as canonical forms, compact long-range operator representations, and the ability to contract the norm and energy exactly underpin the success of these methods.

The utility of one-dimensional TNs is limited by the amount of entanglement between the two halves of a bipartitioned state. To rigorously address the need for a TN class for two-dimensional systems, the projected entangled pair state (PEPS) was introduced and significant progress has resulted in algorithms that have enabled accurate simulation of complex quantum behaviors [2, 20]. For PEPS methods to become as robust and widely used as their one-dimensional counterparts, several algorithmic challenges are being addressed. For example, while the cost of contracting the energy or norm of a PEPS scales exponentially with the lattice size, the boundary contraction method has proved widely accurate for performing efficient approximate contractions. Unfortunately, other issues, such as the canonicalization of PEPS and compact long-range operator representations, have been more challenging and remain open. A significant amount of effort has been put into the former of these, with several proposals for creating canonical PEPS representations approximately having been introduced and used in the past few years [21–23], representing meaningful progress in this area.

In this work, we focus on improving ground state optimization approaches for Hamiltonians with long-range interactions in two dimensions. Since their introduction, nearly all PEPS calculations have been for Hamiltonians with local terms, primarily limited to nearest-neighbor interactions. Simulation of many physically important systems, however, requires the inclusion of long-range interactions, particularly potentials that decay uniformly as a function of the linear distance between sites. Of particular importance in quantum chemistry is the coloumbic interaction between electrons. For classical systems, as TN methods are extended to the continuum limit [24–27], many relevant models require particles to interact via long-range terms such as the Lennard-Jones or Weeks-Chandler-Anderson (WCA) potentials.

In one-dimensional systems, a variety of approaches are used depending on the type of interaction considered. These include using a sum of operators [28], linear combinations of exponential interactions [1], and algorithms to accurately compress MPOs for arbitrary interactions [29]. When moving to two dimensions, two complications arise. First, creating compact representations of long-range operators in two dimensions is significantly more complicated than doing so in one dimension. Second, assuming we construct an accurate representation for a given long-range Hamiltonian, it is not clear that it will be useful for performing a DMRG-style ground state search.

In recent years, two preliminary proposals have arisen for overcoming these issues [30, 31]. In the first approach, the long-range Hamiltonian is approximated by a Hamiltonian with nearest-neighbor terms that act on an expanded system with auxiliary sites inserted between all physical sites and traced out in the end [31]. While this is the only work up until now that has been used for ground state optimizations, the size of possible systems is limited because of the additional sites. The second technique utilizes a series of long-range comb-shaped operator TNs that are constructed using both the sum of exponential and MPO compression approximations mentioned previously [30]. These operators are designed such that no operator bonds are compressed to improve accuracy and stability during optimizations. With the preliminary nature of both of these methods, there is significant need for an analysis of the accuracy and efficiency of numerical optimizations using various operator representations.

In this work, we build upon the latter of these approaches and introduce three long-range operator representations, each tailored for a DMRG-style PEPS algorithm [22, 32], and evaluate their effectiveness in ground state calculations. These include two operator representations inspired by previously used comb operators [30], a rerouted comb and a comb-like PEPO, and a naive representation as a sum of MPOs. To contextualize each of these representations, the following section gives an overview of DMRG-inspired PEPS optimization algorithms [22, 32]. Afterward, we introduce each operator representation in Section 4.3 and present the computational scaling of each in Section 4.4. We then evaluate each of these representations by computing the ground state energy of a long-range Heisenberg model with all-to-all coulombic interactions. In this application, we find that the PEPO performs significantly worse than the other approaches, which provide comparable cost and accuracy. It becomes clear that the optimal operator representation is a non-trivial function of the system size, bond dimension, and Hamiltonian. We conclude in 5.1, discussing possible improvements and the outlook for the use of these operator representations for more complex quantum and classical TN calculations.

## 4.2 Two-Dimensional DMRG Optimization

Because of the indisputable success of the DMRG algorithm, initially introduced without the TN formalism as a renormalization group method for one-dimensional lattice systems [33], it is natural to use a similar approach for higher-dimensional TNs. The fundamental element of the DMRG algorithm is to iteratively optimize one or two MPS tensors at a time using a local eigenproblem where all system degrees of freedom are projected into a renormalized local basis. The algorithm then proceeds by sweeping through all sites, updating one or two at a time, until the system's energy converges [1]

While this is conceptually simple to extend to higher dimensions, important properties of the underlying MPS ansatz are not obvious in higher dimensions, which prevents DMRG's immediate success there. Thus most calculations have been limited to nearest-neighbor Hamiltonian terms optimized by the imaginary time evolution via block decimation (TEBD) algorithm [34, 35]. Recent work, however, has shown that not only have two-dimensional PEPS algorithms evolved such that DMRG-style calculations are possible for some finite and infinite systems, but that they also provide improved accuracy over TEBD optimizations [22, 32].

To match the rigor and stability of the DMRG framework in one dimension, a two-dimensional implementation would first need accurate representations of long-range Hamiltonians in a local basis for the iterative eigenproblems. Ideally, these should be built iteratively such that the full algorithmic cost is linear with the number of lattice sites. Figure 4.1(b) illustrates how a long-range Hamiltonian (blue) sandwiched between a bra and a ket PEPS (pink) can be approximately contracted around the optimization site (red). While we represent the Hamiltonian here abstractly to emphasize the generality of this local construction, the Hamiltonian is normally decomposed into tensors that act on a single site with auxiliary bonds that mimic the structure of the PEPS. Second, ideally, an orthonormal basis representation would be provided by a canonical form of the TN [21–23]. This is illustrated

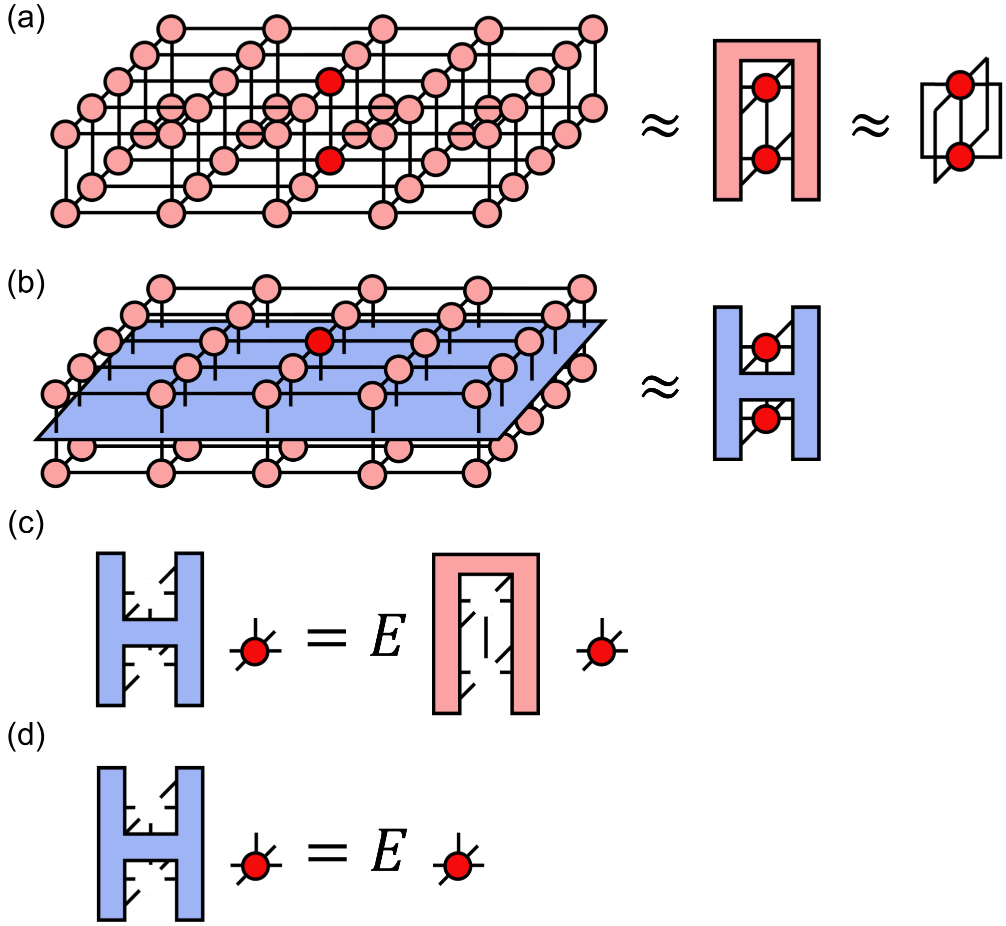

Figure 4.1: A diagrammatic representation of the fundamentals for a DMRG-style optimization for a two-dimensional PEPS. The red site indicates the current optimization site. In (a) the norm environment is approximately contracted into a rank 8 tensor around the optimization site and, if the PEPS is in a canonical form, the norm environment is shown to trace to an identity. (b) illustrates how a local Hamiltonian can be contracted using an abstract representation of the Hamiltonian as a dense tensor instead of the specific representations discussed later. In (c) and (d) we show how the local optimization can be cast as an eigenproblem. Without canonicalization this is a generalized eigenproblem (c) versus the standard eigenproblem in (d).

in Figure 4.1(a), where we show the contraction of a PEPS norm (pink) around the optimization sites (red). While with no conditioning of the PEPS, this norm environment is some tensor, shown in the middle of the figure, if the PEPS has been canonicalized, the environment tensor becomes an identity, as shown on the right. Without canonicalization, the tensor of interest can be optimized using a generalized eigenproblem, expressed as a TN diagram in Figure 4.1(c), while a tensor in a canonicalized PEPS is optimized via a standard local eigenproblem. While one can get around the latter requirement by using a generalized eigensolver with increased instability and cost, the former requirement is unavoidable, explaining current limitations to nearest-neighbor Hamiltonians.

Regardless of the Hamiltonian representation or use of a canonical form, computing the local Hamiltonian or norm is usually the most expensive part of a PEPS optimization, involving contracting all operator tensors with all PEPS tensors except those being optimized. While the structure of an MPS allows this to be determined exactly with a cost that scales linearly with the number of lattice sites, the corresponding contraction for a PEPS scales exponentially with system size and is prohibitive for all but the smallest lattices. For this reason, the approximate boundary contraction method has become nearly universally used for computing PEPS norms, expectation values, and local Hamiltonians for short and long-ranged operators [35, 36].

This approach, sketched using a bird's eye view of the TN in Figure 4.2, exploits the similarities between contracting a boundary PEPS column with its neighboring column and a contraction between an MPS and MPO. In both cases, as more columns or MPOs are contracted, the number and size of uncontracted bonds remain constant while the auxiliary bond of the MPS grows quickly, shown in the left two diagrams in Figure 4.2(b). For this reason, after each column or MPO is contracted into the boundary, the MPS is canonicalized and the auxiliary bonds are subsequently truncated, shown from the second to the third diagram of Figure 4.2(b), to a constant $\chi$, called the boundary bond dimension. The approximate Hamiltonian is created by repeatedly doing boundary MPS contractions and truncations from both sides (including operator terms where necessary) until a left and right environment can be contracted exactly with the PEPS tensors in the column that is being optimized, the TN for which is shown in Figure 4.2(c). While the nearest-neighbor TEBD algorithm avoids needing operator tensors in the boundary contraction, calculations with long-range operators inevitably must include these and the operator representation significantly affects the accuracy of the boundary contraction and consequently both

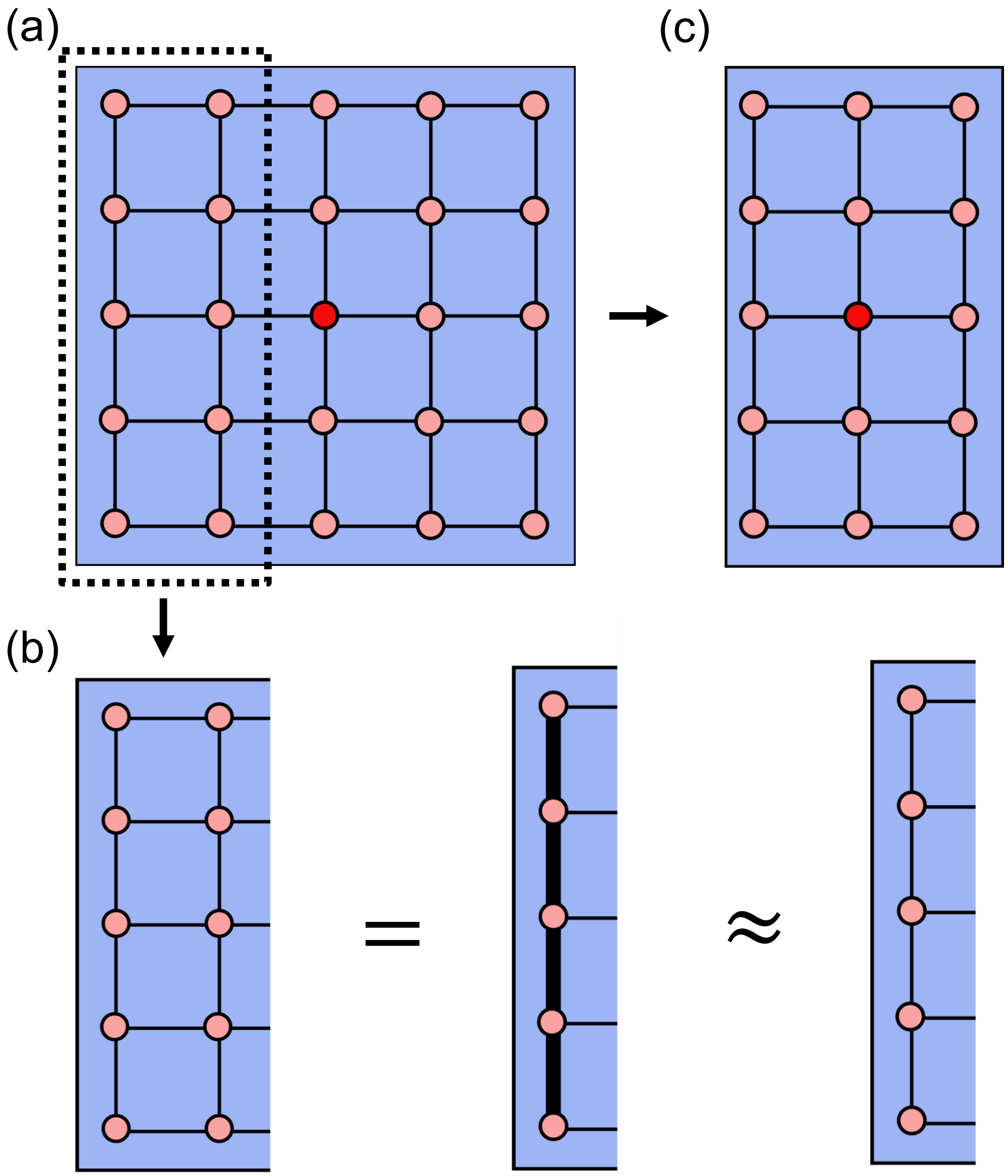

Figure 4.2: A diagrammatic description of the boundary contraction method. (a) shows how we seek to combine the left two columns to create a left environment. In (b) this is done by first contracting the two columns to create an MPS with an enlarged auxiliary bond. This bond is then truncated using canonicalization via QR decompositions and truncation via singular value decompositions. By doing this for the left and right boundaries (iteratively for systems with more columns) we get what looks like an MPO sandwiched between two MPSs, which can be contracted exactly to get the local Hamiltonian or norm.

the local eigenproblems and the overall DMRG optimization.

A generic DMRG-style algorithm for a ground state PEPS calculation can then proceed by repeating the following until the energy converges:

1. Compute and store the right boundary MPSs for the norm and Hamiltonian environments, starting from column $N_y$ and iteratively building the right boundary MPSs to column 2.
2. Loop through all of the columns in the lattice from left to right, doing the following for each column:
    a) Compute and store the top environments for the Hamiltonian and norm environments, starting from the top row $N_x$ and iteratively building towards the bottom.
    b) Loop through each row in the column doing the following:
        i. Construct the local norm and Hamiltonian using the left, right, top, and bottom environment tensors.
        ii. Solve the local eigenproblem and update the optimization site as the eigenstate that corresponds to the largest eigenvalue.
        iii. Update the bottom environment tensors
    c) Update and store the left boundary MPSs for the current column for the norm and Hamiltonian.
3. All of the left boundary MPSs have now been constructed and we can perform a backward sweep by repeating an analog of step 2, moving from right to left through the columns of the PEPS.

### 4.3 Long-Range Operator Representations

Knowing that the computational cost and stability of a DMRG-style PEPS optimization is highly dependent on the accuracy of the local Hamiltonian, the challenge is to now design a representation of long-range Hamiltonians that provides a balance between contraction cost and accuracy. In this section, we will introduce three representations for long-range Hamiltonians $\mathcal{H}$ of the form

$$\mathcal{H} = \sum_{i<j} V_{ij} \left( \mathcal{A}_i \mathcal{B}_j + \mathcal{C}_i \mathcal{D}_j + \mathcal{E}_i \mathcal{F}_j + \cdots \right) \tag{4.1}$$

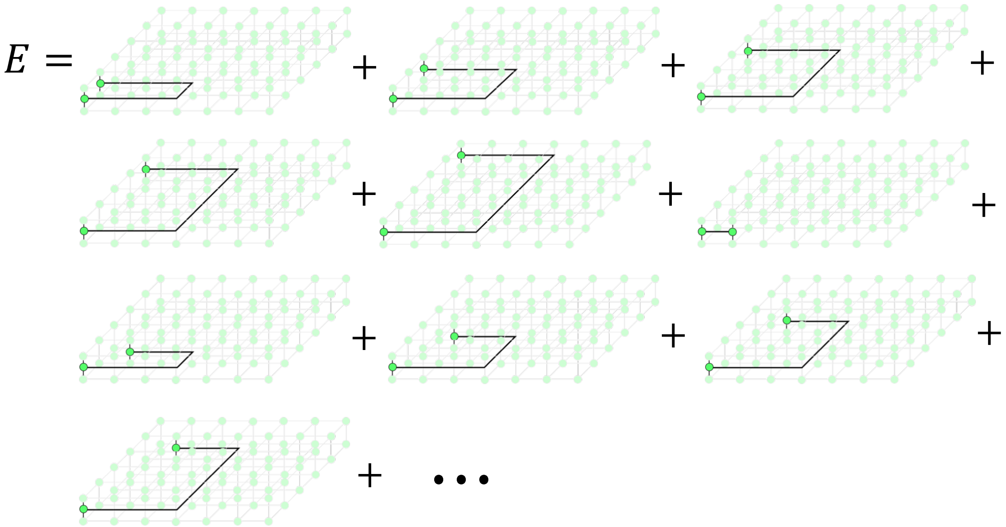

Figure 4.3: The sum of MPOs long-range operator representation, showing how the energy is given by a sum over PEPS contractions with two-site MPOs acting between all pairs of lattice sites.

where $i$ and $j$ loop over the sites in the PEPS, $\{\mathcal{A}_i, \mathcal{B}_i, C_i, etc.\}$ are operators acting on a given site, and $V_{ij}$ is the interaction between sites that decays uniformly as a function of Euclidean distance between the sites, i.e.

$$V_{ij} = f\left(\sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}\right). \tag{4.2}$$

Our guiding principle here is to tailor the operator representations towards use with a DMRG-style optimization using a boundary contraction method. Because boundary contractions require truncation of PEPO bonds we choose not to include standard PEPO representations, which can be used to exactly represent long-range interactions [37] or constructed as compressed PEPOs using recently proposed approximate methods [38].

### Sum of MPOs

We begin with the simplest representation, which we call a sum of MPOs and is an extension of one-dimensional long-range operator representations that involve sums of operators [28]. Each term in the Hamiltonian can be represented as an MPO acting on two sites,

$$V_{ij}\left(\mathcal{A}_i\mathcal{B}_j + \mathcal{C}_i\mathcal{D}_j + \mathcal{E}_i\mathcal{F}_j + \cdots\right) = V_{ij}\begin{pmatrix}\mathcal{A}_i & \mathcal{C}_i & \mathcal{E}_i & \cdots\end{pmatrix}\begin{pmatrix}\mathcal{B}_j \\ \mathcal{D}_j \\ \mathcal{F}_j \\ \vdots\end{pmatrix} \tag{4.3}$$

where $\mathcal{W}_a^{[i]} = V_{ij}\begin{pmatrix}\mathcal{A}_i & \mathcal{C}_i & \mathcal{E}_i & \cdots\end{pmatrix}$ and $\mathcal{W}_a^{[j]} = \begin{pmatrix}\mathcal{B}_j & \mathcal{D}_j & \mathcal{F}_j & \cdots\end{pmatrix}^\dagger$ are rank-3 MPO tensors, each containing operators that act only on the physical indices of their respective sites and contain all interacting terms between two sites. Using this approach, the Hamiltonian is represented with $\frac{1}{2}\left(N^2 - N\right)$ MPO pairs and the energy of the state can be computed by contracting the bra and ket PEPS independently with each MPO pair, as illustrated in Figure 4.3, where the MPO tensors are highlighted to distinguish them from the PEPS tensors.

A peculiar aspect of the diagrams in Figure 4.3 is the way that we have rerouted the auxiliary bond of the MPO through the center column. This is done because experience with approximate boundary contraction with operators indicates that truncation of operator bonds can severely impact the local Hamiltonian's accuracy and the optimization's stability and should thus be avoided when possible [37]. Instead of inserting identity tensors to connect MPO pairs within a column, we choose to always have the MPO bonds rerouted through the column being optimized, avoiding MPO bond truncation.

Using this representation of the long-range operators in practice involves contracting representations of each Hamiltonian term in the local basis, i.e. contracting $\mathcal{O}\left(N^2\right)$ separate environments, one for each pairing of interacting sites and one for the norm. To minimize computational cost, we seek to reuse as many intermediates as possible, allowing us to lower the prefactor for the cost of the DMRG algorithm while not changing the formal scaling with system size. Additionally, while not done here, the contraction of these can be easily parallelized such that, with enough resources, the time for contracting all local environments can be equivalent to the time required to contract a single environment. It should also be noted that unlike in infinite PEPS (iPEPS) algorithms, where completed operator environments can be summed, doing likewise here would require summing MPSs and result in further auxiliary bond increases and requisite truncations.

### Rerouted Comb Operator

While the sum of MPOs approach provides a conceptually simple representation that is easily parallelized, the quadratic growth of the number of local Hamiltonian contractions with $N$ is not ideal. We can build upon the recently introduced idea of comb operators [30] to create a similar approach tailored for a DMRG-style optimization. We will first introduce the comb representation of long-range operators, then discuss how they can be modified for a DMRG ground state calculation.

As a precursor to comb operators, an important result from one-dimensional MPO construction is that an exponentially decaying long-range interaction between all sites can have a compact representation as an MPO:

$$\begin{aligned}\mathcal{H}_{exp} &= \sum_{i<j} e^{-\alpha(j-i)}\mathcal{A}_i\mathcal{B}_j \\ &= \sum_{\{a_i\}} \mathcal{W}^{[1]}_{a_1}\mathcal{W}^{[2]}_{a_1,a_2}\cdots\mathcal{W}^{[N-1]}_{a_{N-2},a_{N-1}}\mathcal{W}^{[N]}_{a_{N-1}},\end{aligned} \tag{4.4}$$

where $\mathcal{W}^{[i]}$ are the MPO tensors:

$$\begin{aligned}\mathcal{W}^{[1]}_{a_1} &= \begin{pmatrix} 0 & \mathcal{A}_1 & \mathcal{I}_1 \end{pmatrix} \\ \mathcal{W}^{[i]}_{a_1} &= \begin{pmatrix} \mathcal{I}_1 & 0 & 0 \\ e^{-\alpha}\mathcal{B}_1 & e^{-\alpha}\mathcal{I}_1 & 0 \\ 0 & \mathcal{A}_1 & \mathcal{I}_1 \end{pmatrix} \\ \mathcal{W}^{[N]}_{a_1} &= \begin{pmatrix} \mathcal{I}_1 \\ e^{-\alpha}\mathcal{B}_1 \\ 0 \end{pmatrix}.\end{aligned} \tag{4.5}$$

While this compact representation is not generic for an arbitrary uniform decaying long-range $f(j-i)$, this can be approximated as a sum of exponentials

$$f(j-1) = \sum_{k=1}^{K} \gamma_k e^{-\alpha_k(j-i)} \tag{4.6}$$

and the Hamiltonian can then be represented by an MPO with auxiliary bond dimension $D_{MPO} = K + 2$.

Ideally, an analogous representation could be used for long-range interactions in a two-dimensional lattice. While most physical potentials are dependent on Euclidean distance, extending the exponential MPO representation for use in a PEPO gives potentials as functions of Manhattan distance $f\left(\left|x_i - x_j\right| + \left|y_i - y_j\right|\right)$. Instead of a sum, a product of Gaussian interactions in the horizontal and vertical directions can give a potential dependent on radial distance, i.e.

$$\begin{aligned} f\left(\left|x_i - x_j\right|, \left|y_i - y_j\right|\right) &= f_x\left(\left|x_i - x_j\right|\right) f_y\left(\left|y_i - y_j\right|\right) \\ &= e^{-\alpha\left(x_i - x_j\right)^2} e^{-\alpha\left(y_i - y_j\right)^2} \\ &= e^{-\alpha\left(\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2\right)} \\ &= f\left(r^2\right), \end{aligned} \tag{4.7}$$

where $r$ is the Euclidean distance.

The Gaussian potential, however, does not have a compact exact MPO form like the decaying exponential. Instead, it can be represented approximately using an MPO compression algorithm that can often find compact representations of smooth decaying interactions [29]. Here, the Hamiltonian MPO tensors $\mathcal{W}^{[i]}$ take the form

$$\mathcal{W}^{[i]} = \begin{pmatrix} \mathcal{I} & 0 & 0 \\ v_a^{[i]}\mathcal{A} & X_{a,a'}^{[i]}\mathcal{I} & 0 \\ 0 & w_{a'}^{[i]}\mathcal{B} & 0 \end{pmatrix} \tag{4.8}$$

where $v_a^{[i]}$, $X_{a,a'}^{[i]}$, and $w_{a'}^{[i]}$ are respectively a column vector, a matrix, and a row vector of coefficients, for each site $i$, that are generated via singular value decompositions as detailed in [29]. To illustrate the strength of this compressed representation for the Gaussian potential, Figure 4.4 (a) shows the bond dimension profile of the compressed MPO representation for a sweep of values of $\alpha$, demonstrating that the maximum bond dimension saturates with system size providing a maximum required $D_{MPO}$ of $\mathcal{O}(10)$.

With this, the Gaussian potential can be represented as a product of vertical and horizontal TN operators

$$\begin{aligned} H^{\text{vert.}} &= \sum_{i=1}^{N_y} \sum_{j=i+1}^{N_y} e^{-\alpha(j-i)^2} \left(\mathcal{I}_i\mathcal{B}_j + \mathcal{B}_i\mathcal{I}_j + \mathcal{A}_i\mathcal{B}_j\right) \\ H^{\text{horz.}} &= \sum_{i=1}^{N_x} \sum_{j=i+1}^{N_x} \left(\mathcal{A}_i\mathcal{I}_j + \mathcal{A}_i\mathcal{B}_j\right), \end{aligned} \tag{4.9}$$

where the vertical TN operator is an MPO and the horizontal is a generalized MPO (gMPO), which combine to form a comb operator, explicitly derived in [30] The Gaussian potential is represented by $N_x$ combs and other uniformly decaying

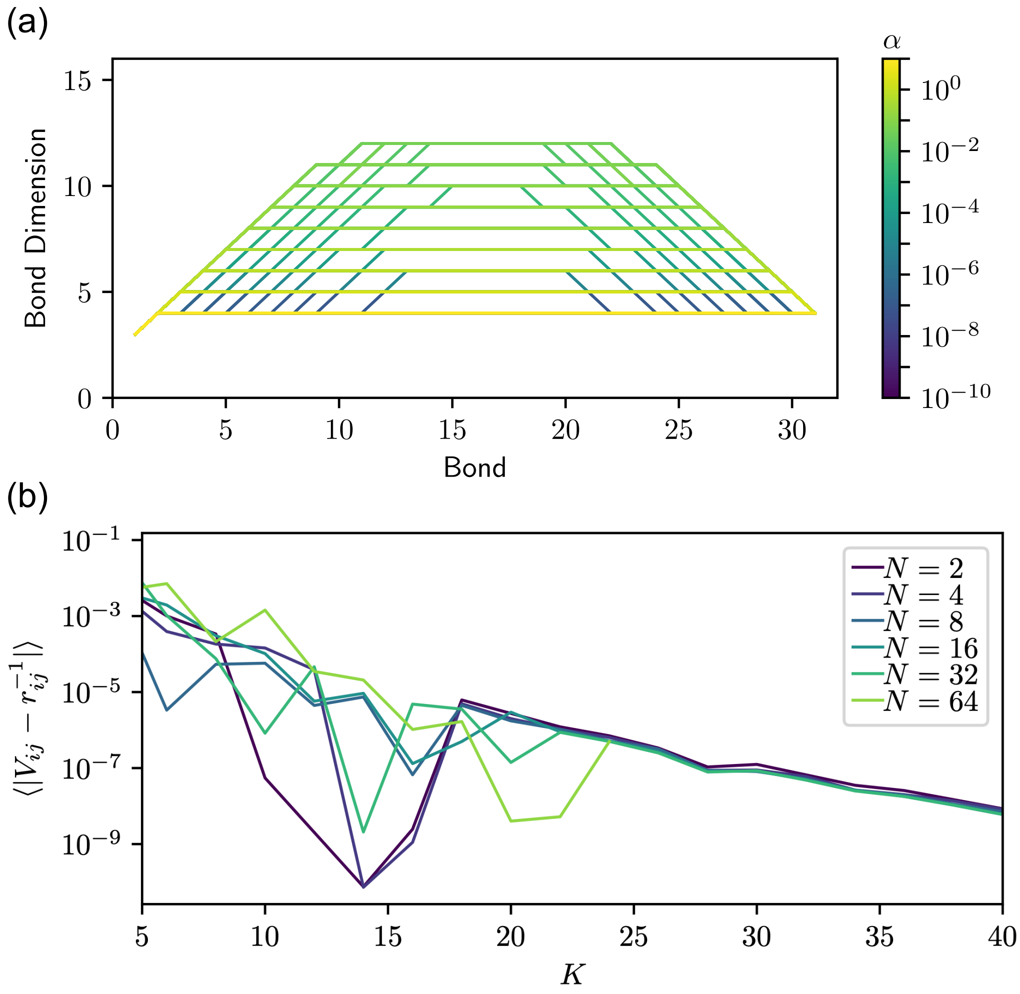

Figure 4.4: (a) The maximum bond dimension, shown as a function of bond position in the MPO, required to represent a long-range Gaussian potential. Each line corresponds to a different $\alpha$, denoted by the line color. The accuracy of the MPO compression was determined by retaining singular values greater than $10^{-10}$. (b) The absolute error in the coulombic interactions $\left\langle \left| V_{ij} - r_{ij}^{-1} \right| \right\rangle$ as a function of the number of summed Gaussians $K$, where the average is taken over all possible $\frac{1}{2}(N-1)N$ interactions. The line color indicates the size of the system with $N_x, N_y \in [2, 64]$.

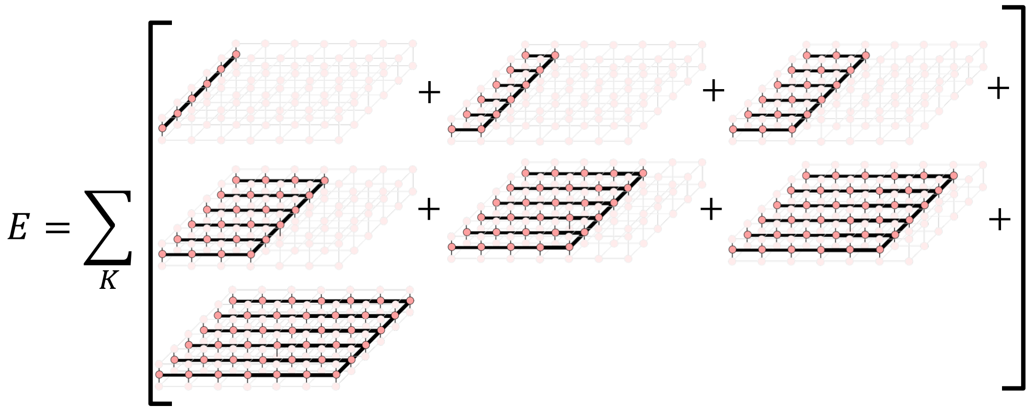

Figure 4.5: The original formulation of the comb operator representation for long-range two-dimensional Hamiltonians, showing how the energy is given by a sum over $K$ Gaussian potentials, each given by a sum of $N_x$ comb operators. The thickness of the operator bonds is proportional to the required bond dimension.

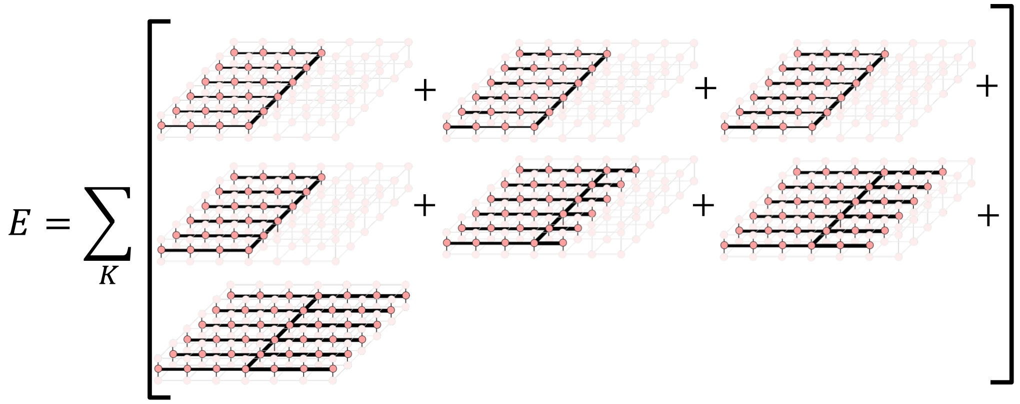

Figure 4.6: The rerouted comb operator representation for long-range two-dimensional Hamiltonians where the $K$ Gaussian potentials are represented by combs whose backbones from Figure 4.5 have been rerouted to the optimization column. The thickness of the operator bonds is proportional to the required bond dimension.

potentials can then be approximated by sums of these

$$V_{ij} \approx \sum_{k=1}^{K} \gamma_k e^{-\alpha_k \left( \left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 \right)}, \tag{4.10}$$

where $\gamma_k$ and $\alpha_k$ may be determined via robust algorithms that allow arbitrary accuracy by varying $K$ [39–41]. In Figure 4.4(b) we show how the accuracy of this approximation for a coulombic potential changes as $K$ increases. For a $64 \times 64$ lattice, the smallest coulombic interaction coefficient is $O\left(10^{-2}\right)$, meaning that $K = 10$ gives, on average, two digits of accuracy for these smallest terms, and increasing to $K = 20$ gives 3-4 digits of accuracy. Thus the energy of a long-ranged Hamiltonian can be computed via the summation of combs shown in Figure 4.5.

While a linearly scaling algorithm for computing the energy of a PEPS with comb operators was introduced alongside the operators themselves, the extension of this to a ground state calculation is not trivial. Though this is theoretically possible via myriad optimization algorithms, such as automatic differentiation [42] or conjugate gradient methods [43], their use in TN algorithms is not as well understood as the DMRG approach. Unfortunately, because of the structure of the comb operators, their direct use in a DMRG-style optimization requires the undesirable truncation of operator bonds and non-ideal computational scaling with system size, $O\left(N_x^2 N_y\right)$.

Focusing on the former difficulty, here we choose to create rerouted comb operators, where the backbone of the comb can be moved from column to column as the optimization proceeds. The structure of this type of operator is shown in Figure 4.6, where the comb backbones are rerouted so the optimization can take place on the center column with no operator bond truncation. In this figure, the thickness of the operator bonds corresponds to the bond dimension; all operator bonds are of size $O\left(D_{MPO}\right)$, where $D_{MPO}$ is the size of the approximate MPO representation of the Gaussian potential. Following the representation of TN operators as signals and rules introduced in [37], Appendix 4.A explicitly gives all of the tensors. With this structure, a DMRG-style calculation can be performed, optionally parallelized over $N_x$ separate combs and $K$ Gaussians and written to reuse boundary environments when possible, without truncating the operator bonds.

### Comb-Like PEPO

While the previous section proposed a representation that overcomes the truncation of operator bonds, it neglected to circumvent the need for multiple combs. Alternatively, we can create a single operator that contains all interactions for the

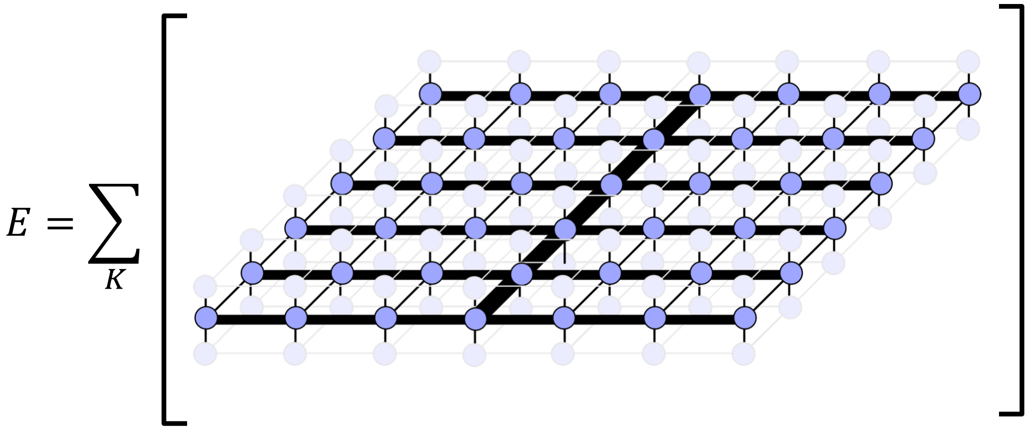

Figure 4.7: A TN diagram giving the comb PEPO representation of the Gaussian potentials used to approximate long-range Hamiltonian terms. The line thickness corresponds to the bond dimension of the PEPO bonds and the backbone (thickest bonds) can be transferred to any column in the lattice to allow for minimal truncation of operator bonds.

Gaussian potential, avoiding the associated incorrect scaling required when using multiple combs, but requiring that operator bonds are truncated. Because boundary contractions scale poorly with the bond dimension of the PEPO, $\mathcal{O}\left(D_{PEPO}^{7}\right)$ and truncating operator bonds leads to inaccuracy and instability, using PEPOs with mostly uniform bonds, constructed via compression algorithms for arbitrary long-range interactions, is currently challenging [38]. As an alternative, we propose using a comb-like PEPO, shown in Figure 4.7, where PEPO tensors are connected to nearest-neighbors like in a standard PEPO, but the vertical bonds are of dimension $D_{PEPO} = 2$ in all columns except where the optimization is occurring. The horizontal bonds have sizes $\mathcal{O}(D_{MPO})$ and the vertical bonds have sizes $\mathcal{O}(D_{MPO}^{2})$, where $D_{MPO}$ is again the size of the approximate MPO representation of the Gaussian potential. Appendix 4.B explicitly gives the tensor structure for these operators.

This long-range operator representation allows us to now implement a DMRG-style ground state calculation with the correct linear scaling with system size. It also minimizes the amount of truncation done to operator tensors, so we would expect that comb-like PEPOs provide better accuracy and stability than a standard PEPO with uniformly distributed bond dimensions.

## 4.4 Theoretical Comparison of Long-Range Operator Representations

Having now introduced three TN representations of operators with uniformly decaying long-range terms, our attention turns toward determining which of these works

Table 4.1: A summary of the theoretical computational cost for a DMRG calculation with each long-range operator representation, including the cost of boundary contractions and compressions, top and bottom environment contractions, and the local eigenproblem update (including the local environment contraction). We also indicate which representation requires truncation of operator bonds.

| | **Sum of MPOs** | **Rerouted Combs** | **Comb-Like PEPO** |
|---|---|---|---|
| Boundary Contractions | $N_x^3, N_y^3, \chi^3, D^5$ | $N_x^2, N_y, \chi^3, D^5, D_{op}^2, K$ | $N_x, N_y, \chi^3, D^5, D_{op}^2, K$ |
| Top/Bottom Environment Contractions | $N_x^3, N_y^3, \chi^3, D^6$ | $N_x^2, N_y, \chi^3, D^6, D_{op}^4, K$ | $N_x, N_y, \chi^3, D^6, D_{op}^6, K$ |
| Local Update | $N_x^3, N_y^3, \chi^3, D^6$ | $N_x^2, N_y, \chi^3, D^6, D_{op}^4, K$ | $N_x, N_y, \chi^3, D^6, D_{op}^6, K$ |
| Requires Operator Truncation | No | No | Yes |

best in practice. While Section 4.5 provides results from numerical experiments probing this, we first provide an overview of the theoretical computational cost of a DMRG calculation with each of these representations and discuss some key differences between the three.

The computational costs are summarized in Table 4.1, where we show how the three main parts of the DMRG-style ground state algorithm scale with the system size ($N_x$, $N_y$), auxiliary PEPS bond dimension ($D$), boundary MPS bond dimension ($\chi$), operator bond dimension ($D_{op}$), and the number of Gaussians used ($K$). We separate the scaling of each of these with commas to emphasize that while each constituent step within these three processes scales differently we only show the most expensive cost with respect to each parameter (i.e. if the local update is done via an implicit iterative eigensolver using a sum of MPOs representation, some contractions require $O\left(N_x^3 N_y^3 \chi^3 D^5\right)$ operations and others requires $O\left(N_x^3 N_y^3 \chi^2 D^6\right)$, giving the scaling costs shown in Table 4.1).

In comparing the costs of the three methods, we first note that for all steps, the three representations have the same cost scaling with $\chi$ and $D$. While we have asserted up to this point that, in practice, the boundary contraction is the most expensive step of a PEPS optimization, Table 4.1 naively would indicate that the other two steps are more expensive. This is not accurate because while the boundary contraction has constituent steps that scale with $O\left(\chi^3 D^5\right)$, both the top/bottom environment contractions and local update have one step that scales as $O\left(\chi^2 D^6\right)$ and another that scales as $O\left(\chi^3 D^5\right)$. Because the general rule of thumb states that $\chi$ scales with $O\left(D^2\right)$, the $O\left(\chi^3 D^5\right)$ step dominates, putting the scaling of all three steps on equal footing. In practice, we continue to find that, for practically sized systems, the boundary contraction is the most expensive step.

The comparison between the three representations then hinges on two things. First, because we expect $D_{op}$ and $K$ to saturate with system size, we would expect the comb-like PEPO to become more efficient than either other methods for sufficiently large systems. Since PEPS calculations are currently limited to $N_x, N_y = O(10)$, it is unclear whether the additional costs incurred by including operator bonds and summing over Gaussians provide a practical advantage over summing over MPOs. Second, a further complication is incurred when we consider the scaling of $\chi$. As mentioned, a commonly used guide indicates that $\chi$ scales quadratically with the PEPS auxiliary bond dimension and is constant with $N$. This rule of thumb, however, comes from experience contracting environments without operator terms. We would expect that $\chi$ scales best for the sum of MPOs, because terms are treated separately, and worst for the comb-like PEPO, because of the operator bond truncation, but a critical question is then how $\chi$ must scale with $D$ and $N$ for an accurate and stable calculation.

## 4.5 Results

Because we are unaware of previous results for DMRG-style ground state calculations on a finite PEPS without canonicalization (some calculations have been done for infinite systems without canonicalization [32] and for finite systems with canonicalization [22]) we begin by benchmarking this algorithm, using the sum of MPOs operator representation, for a nearest-neighbor Heisenberg model

$$\mathcal{H} = \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j \tag{4.11}$$

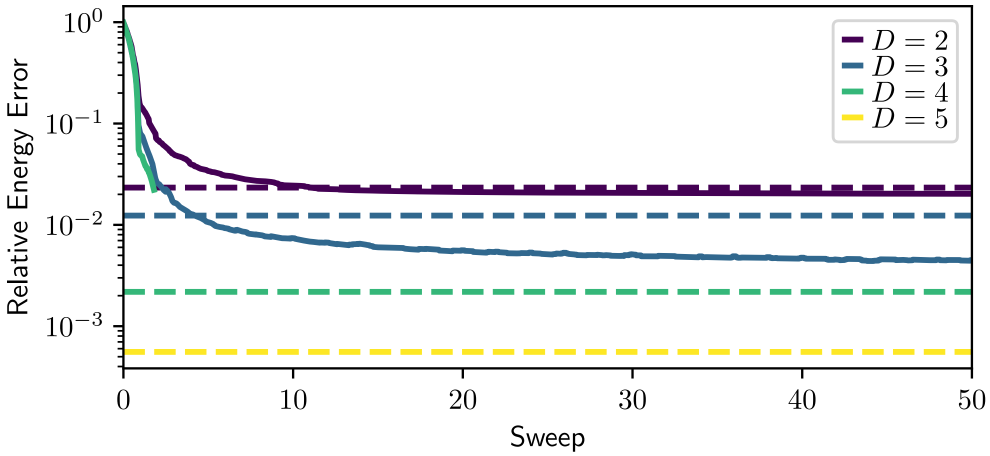

Figure 4.8: Convergence of the energy for a DMRG-style ground state calculation for the nearest-neighbor Heisenberg model on a $10 \times 10$ lattice with $D \in [2, 3]$, where $D$ is indicated by the line color. Solid lines show how the energy converges during the DMRG-style optimizations while dashed lines show converged reference energies computed using the standard full update procedure. The error is computed using the $D = 6$ full update results as a reference.

on a square lattice, where the sum is over all nearest-neighbor lattice sites and where $\vec{S}_i = \begin{pmatrix} \sigma_x & \sigma_y & \sigma_z \end{pmatrix}$, with the spin operators defined as

$$\begin{aligned} \sigma_x &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_y &= \frac{1}{2i}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ \sigma_z &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{4.12}$$

Figure 4.8 shows the convergence of the energy for a system of size $N_x = N_y = 10$ with $\chi = 128$ for $D \in [2, 3]$ (solid lines) compared against converged energies from the standard full update procedure (dashed lines), taken from [34]. The reference values are converged with respect to the imaginary time step size and $\chi$, and the relative error is computed using the $D = 6$ full update energy as a reference. While we observe that the energy converges smoothly from a random initial guess, we surprisingly find that the DMRG-style optimization provides a lower energy than fully converged TEBD results. This has also been previously observed in infinite PEPS calculations and it is postulated that this might be attributable to the way local tensors can affect Hamiltonian terms at distant sites [32].

We can now move to a Hamiltonian with long-range interactions. We choose to study the coulombic Heisenberg model, whose Hamiltonian is given by

$$\mathcal{H} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}} \vec{\mathcal{S}}_i \cdot \vec{\mathcal{S}}_j \tag{4.13}$$

where $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. This model has been previously used as a test-bed for long-range PEPS algorithms [44], with the long-range hopping terms making it a particularly challenging system.

In Figure 4.9(a) we show the convergence of the energy using a DMRG-style optimization for each of the proposed Hamiltonian representations. These calculations are done on a $6 \times 6$ lattice with $D \in [2, 3]$ and $\chi = 181$; for the rerouted combs and comb PEPO calculations, we use $K = 20$ Gaussians. As would be expected, for a given bond dimension, the energies converge mostly smoothly and at roughly the same convergence rate for all three Hamiltonian representations.

To begin a comparison of the computational efficiencies, Figure 4.9(b) shows the same curves as a function of the computational time, with each optimization run on eight processors with parallelization done only at the level of tensor operations. We emphasize that this means these are not parallelized over sums of MPOs, individual combs, or Gaussians, which would provide decreased computational times for all methods. We find that, when the same $\chi$ is used for all approaches, the calculation using the sum of MPOs representation is fastest. This indicates that, for small systems, the cost of summing over Gaussians and including many operator tensors overwhelms the cost associated with summing over all site pairs. This ordering would be expected to change as we study larger systems. Surprisingly, by comparing the ordering of the $D = 2$ and $D = 3$ curves we also find that for modest bond dimensions the increased size of the comb PEPO tensors makes it the least efficient representation in practice.

From these initial results alone, it is clear that the most computationally efficient Hamiltonian representation is a complex function of the system size and bond dimension. An additional layer of complexity, however, emerges when we continue to perform additional optimization sweeps for the $D = 3$ calculations, as shown in Figure 4.9(c). After approximately five sweeps, the comb PEPO calculation becomes unstable. This instability is attributable to the approximate nature of the local

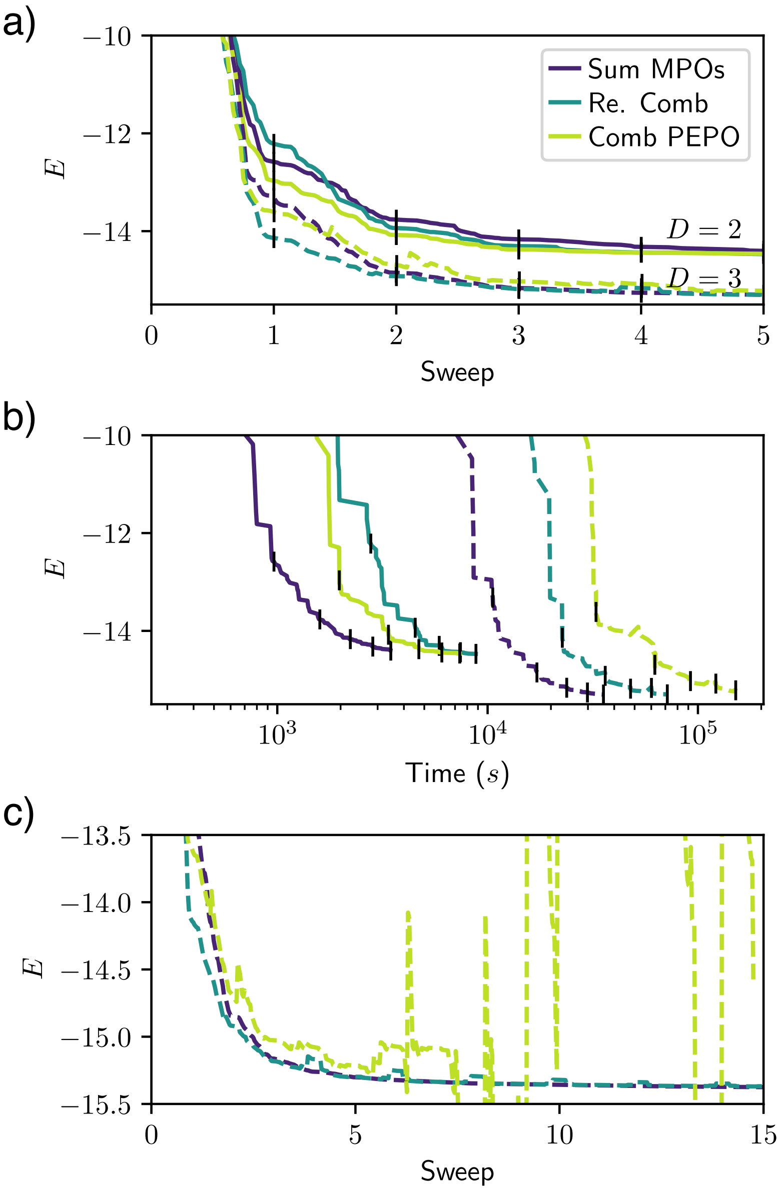

Figure 4.9: An analysis of the convergence of the DMRG-style optimization using each of the three operator representations for a $6 \times 6$ lattice with $\chi = 181$ and $K = 20$. The color corresponds to the operator representation and the line style denotes the bond dimension $D \in [2, 3]$. (a) shows the convergence in energy for each of the operator representations over the first five lattice sweeps while (b) shows this same result as a function of wall time. (c) shows instability in the convergence induced by inaccurate local environments for insufficiently large $\chi$.

Hamiltonian and norm. While this can be circumvented by increasing the boundary bond dimension further, the curves from the three optimizations indicate that the threshold $\chi$ for a stable optimization is dependent on operator representation. Here, for example, with fixed $\chi$, the comb PEPO calculation fails after about five sweeps, but the convergence of the sum of MPOs calculation is monotonically converging while the rerouted combs calculation exhibits small fluctuations. This introduces deeper complexity into predicting the most efficient Hamiltonian representation, requiring us to consider system size, bond dimension, and the required boundary bond dimension for stability.

Because minimizing this instability would enable more efficient calculations, in Figure 4.10 we focus for a moment on understanding its nature and attempting to avoid it, studying how various properties of the local optimization problem behave as the boundary bond dimension increases. To set up these calculations, we performed 100 sweeps of a DMRG-style calculation on a $6 \times 6$ PEPS with $D = 2$, $K = 20$, $\chi = 181$, using rerouted combs to converge to $E = -14.76$ (shown for reference by the grey dashed line in all three plots). We then repeatedly performed boundary contractions to compute the local Hamiltonian and norm at increasing values of $\chi \in [1, 100]$, and, to provide a reference for computing relative errors, performed an exact contraction (i.e. $\chi = \infty$).

In Figure 4.10(a) we show how the locally optimized energy (solid lines), diagonal local Hamiltonian terms (dashed lines), and off-diagonal local Hamiltonian terms (dotted lines) converge as a function of $\chi$. As a rule of thumb, when the energy dips below the gray line, the optimization is considered to be continuing to converge and the corresponding $\chi$ indicates what is needed for a stable calculation. Immediately obvious is the stark difference between the three representations; the sum of MPOs, rerouted combs, and comb PEPOs respectively require $\chi \approx 30$, $\chi \approx 100$, and $\chi \gg 100$. Because the long-range Heisenberg model includes long-range particle hopping (as opposed to long-range potentials that affect nearest-neighbor hopping rates) it is considered especially challenging. To quantify the effects of this, we compare the convergence of the diagonal and off-diagonal elements in the local Hamiltonian, noting that long-range hopping terms involving other sites are included in the diagonal of the local Hamiltonian. This comparison is provided in the difference between the dashed and dotted lines of Figure 4.10(a), where we find that off-diagonal elements converge more slowly. While it remains unclear to what degree, this suggests that the long-range hopping in the Heisenberg model

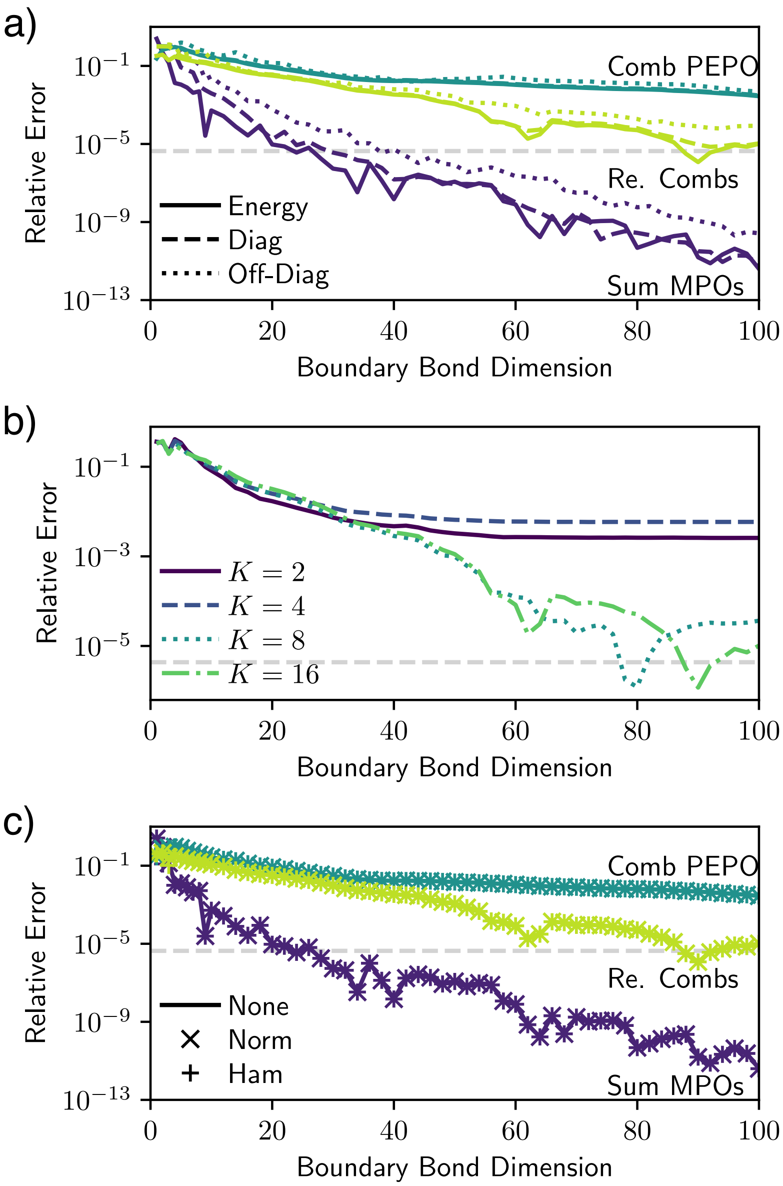

Figure 4.10: A characterization of the root causes of unstable optimizations and the efficacy of proposed solutions. All plots show how the boundary bond dimension $\chi$ affects the accuracy of the local optimization for an already converged $6 \times 6$ calculation with $D = 2$. The gray line on each plot shows the previously converged energy and all relative errors are computed by comparing to an exact contraction ($\chi = \infty$). In (a), we compare the accuracy of the three operator representations with the solid, dashed, and dotted lines respectively corresponding to the relative error in the locally optimized energy, diagonal Hamiltonian terms, and off-diagonal Hamiltonian terms. (b) shows how the accuracy of the local energy depends on the number of Gaussian interactions for the rerouted comb operator representation. (c) shows the effect of doing a local stabilization of the localized norm and Hamiltonian.

exacerbates truncation errors for representations.

As highlighted in Figure 4.4, the accuracy of either of the comb operator representations is severely limited by the number of Gaussians ($K$) used to estimate the coulombic potential. To determine the combined effects of this and finite $\chi$ on the local optimization, we studied the convergence of the local energy with $\chi$ for four different values of $K$ in Figure 4.10(b). From this, while it is clear that having too few Gaussians can severely limit the optimization's accuracy, for sufficiently large $K$, the accuracy will be remain limited by $\chi$.

As a last consideration towards providing increased stability, in standard PEPS optimization algorithms, local tricks are often employed to help stabilize calculations. For example, because the local norm environment is the density matrix, it is physically required to be positive and Hermitian, which can be enforced locally. For a Hermitian Hamiltonian, we may also enforce Hermiticity by locally averaging the Hamiltonian and its transpose. Figure 4.10(c) explores the effects of these local stabilizations on the energy convergence. Surprisingly, unless $\chi$ is very small (i.e. $\chi < 10$), these tricks have no noticeable effect on the accuracy of the local eigenproblem, regardless of the representation used.

With no obvious means towards further stabilizing calculations, we took a practical approach and sought to determine how the required $\chi$ and corresponding computational time scale with system size and bond dimension. To do this, we ran DMRG-style optimizations for the coulombic Heisenberg model with system sizes $N_x = N_y \in [3, 4, 6, 8, 10]$ and $D \in [2, 3]$, systematically increasing $\chi$ until the energy optimizations remained stable for 5 sweeps (with the stability criteria requiring that energy fluctuations are $< 5\%$). To match the stability of standard TEBD optimization methods, $\chi$ should be constant as a function of $N$ and scale quadratically with $D$.

Figure 4.11(a) shows how the required boundary bond dimension scales as a function of system size for optimizations with the three operator representations. The solid lines are for PEPS with $D = 2$ while the dashed lines are for $D = 3$. When using a sum of MPOs, the required $\chi$ remains constant and small with system size. This contrasts sharply against the comb PEPO calculations where we find that $\chi$ must scale exponentially with system size! While the comb PEPOs were constructed to minimize operator bond compression, the difficulty to conduct stable optimizations with these PEPOs emphasizes that when using a boundary contraction method, operator bond truncation should be avoided. The rerouted combs require

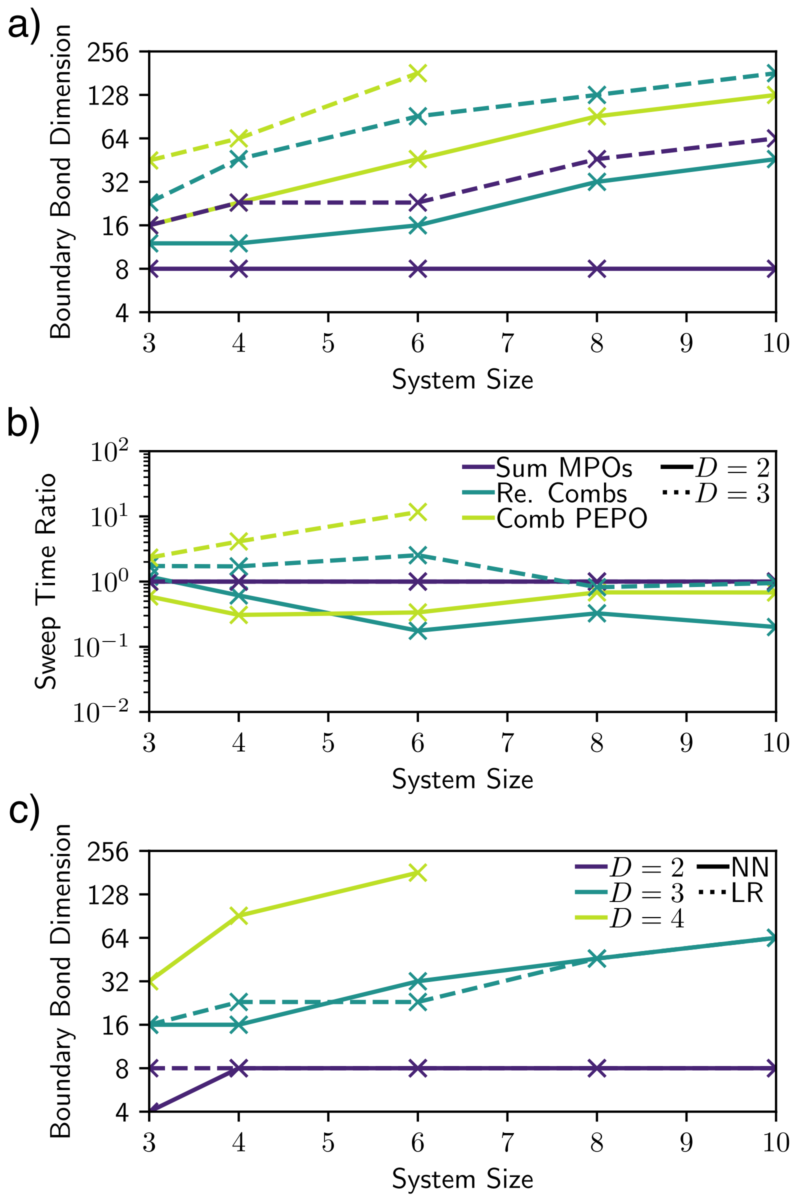

Figure 4.11: A comparison of the computational cost and difficulty for performing a DMRG-style optimization with each of the proposed long-range Hamiltonian representations. In (a) we display the required boundary bond dimension for a stable calculation as a function of the system size $N_x = N_x \in [3, 10]$ for $D \in [2, 3]$. (b) provides compares the computational efficiency of each operator representation by showing the ratio of the time required for a DMRG sweep through all sites against the time required when using the sum of MPOs representation. The legend in (b) is shared for plots (a) and (b). (c) compares the required $\chi$ for a stable calculation as a function of the system size and bond dimension when using a sum of MPOs for Heisenberg Hamiltonians with nearest-neighbor (NN) and long-range (LR) terms.

intermediate boundary bond dimensions that grow, but relatively slowly, with lattice size. Looking at the $D = 3$ results, these observations persist, except that for even the sum of MPOs approach, the required $\chi$ must grow with system size.

Because constructing the local representation of a nearest-neighbor Hamiltonian using a sum of MPOs is nearly identical to doing this for a long-range Hamiltonian, we guessed that the required bond dimension for a stable calculation should scale similarly for both of these calculations. Figure 4.11(c) shows results exploring this and indeed finds that the long-range and nearest-neighbor calculations require the same boundary bond dimension for stable calculations. Because we know that accurate ground states can be computed with limited $\chi$ for nearest-neighbor models, this indicates that the instability is due to the local optimization using the approximate Hamiltonian. It is thus likely that in future work, by slowing down the convergence of the DMRG procedure, we can better stabilize the optimization.

Finally, we conclude which of the operator representations is most efficient for long-range DMRG-style optimizations using Figure 4.11(d). Here, we show the required time to do a stable sweep through the lattice for each operator representation, as a ratio against the same for the sum of MPOs representation. While for the $D = 2$ results the comb PEPO and rerouted combs provide an advantage, it is clear from the $D = 3$ results that the sum of MPOs is more efficient for small systems. We also immediately see again that using a comb PEPO is the least efficient approach because of the operator truncation. For larger systems, the rerouted combs representation becomes slightly faster than the sum of MPOs, meaning that either is appropriate for use in practical calculations.

## 4.6 Conclusions

In this paper, we have introduced three approaches to representing long-range Hamiltonians for two-dimensional DMRG-style calculations, with each of these tailored towards use with a DMRG-style optimization. We found that while all three representations could provide accurate ground state energies, if the local Hamiltonian was not sufficiently accurate, as controlled by the boundary bond dimension, then the DMRG optimization was unstable. While none of the naive approaches towards stabilizing the optimization provided substantive change, we hypothesized that slowing down the convergence rate will dampen instabilities and allow for accurate calculations with moderate boundary bond dimensions.

Regarding which of the operator representations is most efficient for use with a

DMRG-style optimization, we find this to be a complex function of system size, PEPS bond dimension, and the Hamiltonian. In general, we find that using a PEPO, even one where operator bond truncation is minimized, is inadvisable, requiring extremely large boundary bond dimensions and becoming prohibitively expensive quickly. For small systems with moderate bond dimensions, we find that the sum of MPOs approach is most efficient. As the system becomes larger, the rerouted combs approach becomes competitive with the sum of MPOs approach and it appears that both are viable candidates for practical calculations. We expect that this work will serve as a starting point towards carrying out calculations for more complex physical systems where interacting particles are of central importance, including those in the continuum limit.

## 4.A Rerouted Comb Operator Construction

This appendix provides explicitly the structure of the rerouted comb operators introduced in Section 4.3. We follow the style of [37] and introduce the operators as rules that govern signals between tensors. Note that when considering rerouted comb operators, there are two backbones associated with each comb operator, one for the backbone of the natural comb, as seen in Figure4.5, and another for the backbone of the rerouted comb, as seen in Figure4.6. In the rules given here, we refer to the former as the backbone column and the latter as the optimization column. Rules correspond to blocks of a tensor instead of single entries as would be conventional. The coefficients for the blocks ($v^{[i]}$, $X^{[i]}$, $w^{[i]}$) come from MPO tensors of the form shown in Equation 4.8 and determined for Gaussian long-range interactions via the compression algorithm detailed in [29]. The same coefficients may be used for both the vertical and horizontal interaction terms; for clarity, we use index $i$ and $j$ to respectively refer to the coefficients associated with the $x$ and $y$ coordinate.

Tables 4.2 through 4.10 give the rules for the MPO tensors to the left of the optimization column, with Tables 4.2 to 4.4, Tables 4.5 to 4.7, and Tables 4.8 to 4.10 respectively giving the rules for the bottom, top, and center row MPOs. In each of these series, the first (Tables 4.2, 4.5, and 4.8) are for when the backbone column is in the same column as the operator tensor, while the second (Tables 4.3, 4.6, and 4.9) and third (Tables 4.4, 4.7, and 4.10) are respectively for when the backbone column is to the right and left of the operator tensor.

The MPOs on the right side of the optimization columns are given in Tables 4.11 and 4.12.

The backbone PEPO tensors are given in Tables 4.13 and 4.21. Tables 4.13 to 4.15, Tables 4.16 to 4.18 and Tables 4.19 to 4.21 respectively giving the rules for the bottom, top, and center row combs. In each of these series, the first (Tables 4.13, 4.16, and 4.19) are for when the backbone column is in the same column as the operator tensor, while the second (Tables 4.14, 4.17, and 4.20) and third (Tables 4.15, 4.18, and 4.21) are respectively for when the backbone column is to the left and right of the operator tensor.

Table 4.2: Rules for the teeth tensors in the bottom row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is in the same column as this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,e) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1) | $v^{[i]}\mathcal{I}$ |
| 4 | (0,1) | $\mathcal{A}$ |
| 5 | (0,2) | $\mathcal{B}$ |

Table 4.3: Rules for the teeth tensors in the bottom row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the right of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (0,1) | $w^{[i]}\mathcal{A}$ |
| 3 | (1,1) | $X^{[i]}\mathcal{I}$ |

Table 4.4: Rules for the teeth tensors in the bottom row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the left of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,1) | $\mathcal{I}$ |
| 3 | (2,2) | $\mathcal{I}$ |
| 4 | (e,e) | $\mathcal{I}$ |

Table 4.5: Rules for the teeth tensors in the top row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is in the same column as this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,e) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1) | $v^{[i]}\mathcal{I}$ |
| 4 | (0,2) | $\mathcal{B}$ |
| 5 | (1,3) | $v^{[i]}\mathcal{I}$ |

Table 4.6: Rules for the teeth tensors in the top row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the right of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (0,1) | $w^{[i]}\mathcal{A}$ |
| 3 | (1,1) | $X^{[i]}\mathcal{I}$ |

Table 4.7: Rules for the teeth tensors in the top row of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the left of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,1) | $\mathcal{I}$ |
| 3 | (2,2) | $\mathcal{I}$ |
| 4 | (3,3) | $\mathcal{I}$ |
| 5 | (e,e) | $\mathcal{I}$ |

Table 4.8: Rules for the teeth tensors in the center rows of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is in the same column as this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,e) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1) | $v^{[i]}\mathcal{I}$ |
| 4 | (0,1) | $\mathcal{A}$ |
| 5 | (0,2) | $\mathcal{B}$ |
| 6 | (1,3) | $v^{[i]}\mathcal{I}$ |

Table 4.9: Rules for the teeth tensors in the center rows of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the right of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (0,1) | $w^{[i]}\mathcal{A}$ |
| 3 | (1,1) | $X^{[i]}\mathcal{I}$ |

Table 4.10: Rules for the teeth tensors in the center rows of the rerouted comb operators when the optimization column is to the right of this tensor and the backbone column is to the left of this tensor. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,1) | $\mathcal{I}$ |
| 3 | (2,2) | $\mathcal{I}$ |
| 4 | (3,3) | $\mathcal{I}$ |
| 5 | (e,e) | $\mathcal{I}$ |

Table 4.11: Rules for the teeth tensors in the rows of the rerouted comb operators when the optimization column is to the left of this tensor and the backbone column is to the right of this tensor. Only rules 1-8 are included for operators in the top and bottom rows.

| **Rule Number** | **Signals** (left, right) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,1) | $X^{[i]}\mathcal{I}$ |
| 3 | (e,1) | $w^{[i]}\mathcal{A}$ |
| 4 | (2,2) | $X^{[i]}\mathcal{I}$ |
| 5 | (3,2) | $w^{[i]}\mathcal{A}$ |
| 6 | (3,3) | $\mathcal{I}$ |
| 7 | (4,4) | $\mathcal{I}$ |
| 8 | (e,e) | $\mathcal{I}$ |
| 9 | (5,2) | $w^{[i]}\mathcal{A}$ |
| 10 | (5,5) | $\mathcal{I}$ |

Table 4.12: Rules for the teeth tensors in the top and bottom row of the rerouted comb operators when the optimization column is to the left of this tensor and the backbone column is in the same column as this tensor. Rule 4 is excluded for tensors in the top row and rule 6 is not included in the top or bottom rows.

| **Rule Number** | **Signals** (left) | **Output** |
|---|---|---|
| 1 | (0) | $\mathcal{I}$ |
| 2 | (1) | $v^{[i]}\mathcal{B}$ |
| 3 | (2) | $v^{[i]}\mathcal{I}$ |
| 4 | (3) | $\mathcal{A}$ |
| 5 | (4) | $\mathcal{B}$ |
| 6 | (5) | $\mathcal{A}$ |

Table 4.13: Rules for the backbone tensors in the bottom row of the rerouted comb operators when the optimization column is the same as the backbone column. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals (left, up)** (left, up) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (1,e) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1) | $v^{[i]}w^{[j]}\mathcal{I}$ |
| 4 | (0,1) | $w^{[j]}\mathcal{A}$ |
| 5 | (0,2) | $w^{[j]}\mathcal{B}$ |

Table 4.14: Rules for the backbone tensors in the bottom row of the rerouted comb operators when the optimization column is to the right of the backbone column.

| **Rule Number** | **Signals (left, up)** (left, up) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 1 | (1,1) | $w^{[j]}\mathcal{I}$ |
| 2 | (2,2) | $w^{[j]}\mathcal{I}$ |
| 3 | (e,e) | $\mathcal{I}$ |

Table 4.15: Rules for the backbone tensors in the bottom row of the rerouted comb operators when the optimization column is to the left of the backbone column.

| **Rule Number** | **Signals** (left, right, up) | **Output** |
|---|---|---|
| 1 | (0,0,0) | $\mathcal{I}$ |
| 2 | (1,1,e) | $X^{[i]}\mathcal{I}$ |
| 3 | (0,e,e) | $\mathcal{I}$ |
| 4 | (0,1,e) | $w^{[i]}\mathcal{A}$ |
| 5 | (1,2,1) | $X^{[i]}w^{[j]}\mathcal{I}$ |
| 6 | (0,2,1) | $w^{[i]}w^{[j]}\mathcal{A}$ |
| 7 | (0,3,1) | $w^{[j]}\mathcal{I}$ |
| 8 | (0,4,2) | $w^{[j]}\mathcal{I}$ |

Table 4.16: Rules for the backbone tensors in the top row of the rerouted comb operators when the optimization column is the same as the backbone column. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals (left, down)** (left, down) | **Output** |
|---|---|---|
| 1 | (1,0) | $v^{[i]}\mathcal{B}$ |
| 2 | (0,e) | $\mathcal{I}$ |
| 3 | (0,1) | $v^{[j]}\mathcal{B}$ |
| 4 | (1,2) | $v^{[i]}v^{[j]}\mathcal{I}$ |

Table 4.17: Rules for the backbone tensors in the top row of the rerouted comb operators when the optimization column is to the right of the backbone column.

| **Rule Number** | **Signals** (left, down) | **Output** |
|---|---|---|
| 1 | (0,e) | $\mathcal{I}$ |
| 2 | (e,0) | $\mathcal{I}$ |
| 3 | (2,1) | $v^{[j]}\mathcal{I}$ |
| 4 | (3,2) | $v^{[j]}\mathcal{I}$ |

Table 4.18: Rules for the backbone tensors in the top row of the rerouted comb operators when the optimization column is to the left of the backbone column. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right, down) | **Output** |
|---|---|---|
| 1 | (0,0,e) | $\mathcal{I}$ |
| 2 | (1,1,0) | $X^{[i]}\mathcal{I}$ |
| 3 | (0,1,0) | $w^{[i]}\mathcal{A}$ |
| 4 | (0,e,0) | $\mathcal{I}$ |
| 5 | (0,4,1) | $v^{[j]}\mathcal{I}$ |
| 6 | (1,2,2) | $X^{[i]}v^{[j]}\mathcal{I}$ |
| 7 | (0,3,2) | $v^{[j]}\mathcal{I}$ |
| 8 | (0,2,2) | $v^{[j]}w^{[i]}\mathcal{A}$ |

Table 4.19: Rules for the backbone tensors in the center rows of the rerouted comb operators when the optimization column is the same as the backbone column. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, down, up) | **Output** |
|---|---|---|
| 1 | (0,0,0) | $\mathcal{I}$ |
| 2 | (1,0,e) | $v^{[i]}\mathcal{B}$ |
| 3 | (0,1,e) | $v^{[j]}\mathcal{B}$ |
| 4 | (0,1,1) | $X^{[j]}\mathcal{I}$ |
| 5 | (1,0,1) | $v^{[i]}w^{[j]}\mathcal{I}$ |
| 6 | (0,0,1) | $w^{[j]}\mathcal{A}$ |
| 7 | (1,2,e) | $v^{[i]}v^{[j]}\mathcal{I}$ |
| 8 | (0,2,2) | $X^{[j]}\mathcal{I}$ |
| 9 | (0,0,2) | $w^{[j]}\mathcal{B}$ |
| 10 | (0,e,e) | $\mathcal{I}$ |

Table 4.20: Rules for the backbone tensors in the center rows of the rerouted comb operators when the optimization column is to the right of the backbone column.

| **Rule Number** | **Signals** (left, down, up) | **Output** |
|---|---|---|
| 1 | (0,0,0) | $\mathcal{I}$ |
| 2 | (0,e,e) | $\mathcal{I}$ |
| 3 | (e,0,e) | $\mathcal{I}$ |
| 4 | (0,1,1) | $X^{[j]}\mathcal{I}$ |
| 5 | (2,1,e) | $v^{[j]}\mathcal{I}$ |
| 6 | (1,0,1) | $w^{[j]}\mathcal{I}$ |
| 7 | (3,2,e) | $v^{[j]}\mathcal{B}$ |
| 8 | (0,2,2) | $X^{[j]}\mathcal{I}$ |
| 9 | (2,0,2) | $w^{[j]}\mathcal{I}$ |

Table 4.21: Rules for the backbone tensors in the center rows of the rerouted comb operators when the optimization column is to the left of the backbone column. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right, down, up) | **Output** |
|---|---|---|
| 1 | (0,0,0,0) | $\mathcal{I}$ |
| 2 | (0,0,e,e) | $\mathcal{I}$ |
| 3 | (1,1,0,e) | $X^{[i]}\mathcal{I}$ |
| 4 | (0,1,0,e) | $w^{[i]}\mathcal{A}$ |
| 5 | (0,4,1,e) | $v^{[j]}\mathcal{I}$ |
| 6 | (0,0,1,1) | $X^{[j]}\mathcal{I}$ |
| 7 | (0,e,0,e) | $\mathcal{I}$ |
| 8 | (1,2,0,1) | $X^{[i]}w^{[j]}\mathcal{I}$ |
| 9 | (0,2,0,1) | $w^{[i]}w^{[j]}\mathcal{A}$ |
| 10 | (0,5,0,1) | $w^{[j]}\mathcal{I}$ |
| 11 | (1,2,2,e) | $X^{[i]}v^{[j]}\mathcal{I}$ |
| 12 | (0,2,2,e) | $w^{[i]}v^{[j]}\mathcal{A}$ |
| 13 | (0,3,2,e) | $v^{[j]}\mathcal{I}$ |
| 14 | (0,4,0,2) | $w^{[j]}\mathcal{I}$ |
| 15 | (0,0,2,2) | $X^{[j]}\mathcal{I}$ |

## 4.B Comb-Like PEPO Construction

In this appendix, we present explicitly the structure of the comb-like PEPO operators introduced in Section 4.3 again following the representation of TN operators as signals as done in [37]. The notes regarding the construction of the TN operators in the previous appendix are also valid here. Additionally, rerouting all entries through a single backbone requires reshaped outer products, denoted ⊗.

All operators forms are given in Tables 4.22 through **??**. Tensors left to the optimization column, where the thick backbone of the PEPO is located, are given in Tables 4.22-4.23, while Tables 4.24-4.25 give the tensors to the right of the optimization column. Last, Tables 4.26-**??** give the backbone tensors.

Table 4.22: Rules for PEPO tensors to the left of the optimization column in the top and bottom rows of the comb-like PEPO. If at the left boundary, then only rules where the left signal is 0 are kept. Rule 4 is not included in the top row.

| **Rule Number** | **Signals** (left, right, up/down) | **Output** |
|---|---|---|
| 1 | (0,0,n) | $\mathcal{I}$ |
| 2 | (0,1,n) | $w^{[i]}\mathcal{A}$ |
| 3 | (1,1,n) | $X^{[i]}\mathcal{I}$ |
| 4 | (0,2,y) | $\mathcal{A}$ |
| 5 | (1,2,y) | $v^{[i]}\mathcal{I}$ |
| 6 | (1,e,n) | $v^{[i]}\mathcal{B}$ |
| 7 | (0,3,y) | $\mathcal{B}$ |
| 8 | (2,2,n) | $\mathcal{I}$ |
| 9 | (3,3,n) | $\mathcal{I}$ |
| 10 | (e,e,n) | $\mathcal{I}$ |

Table 4.23: Rules for PEPO tensors to the left of the optimization column in the center rows of the comb-like PEPO. If at the left boundary, then only rules where the left signal is 0 are kept.

| **Rule Number** | **Signals** (left, right, down, up) | **Output** |
|---|---|---|
| 1 | (0,0,n,n) | $\mathcal{I}$ |
| 2 | (0,1,n,n) | $w^{[i]}\mathcal{A}$ |
| 3 | (1,1,n,n) | $X^{[i]}\mathcal{I}$ |
| 4 | (0,2,n,y) | $\mathcal{A}$ |
| 5 | (1,2,n,y) | $v^{[i]}\mathcal{I}$ |
| 6 | (1,2,y,n) | $v^{[i]}\mathcal{I}$ |
| 7 | (1,e,n,n) | $v^{[i]}\mathcal{B}$ |
| 8 | (0,3,y,n) | $\mathcal{B}$ |
| 9 | (0,3,n,y) | $\mathcal{B}$ |
| 10 | (2,2,n,n) | $\mathcal{I}$ |
| 11 | (3,3,n,n) | $\mathcal{I}$ |
| 12 | (e,e,n,n) | $\mathcal{I}$ |
| 13 | (0,0,y,y) | $\mathcal{I}$ |

Table 4.24: Rules for PEPO tensors to the right of the optimization column in the top and bottom rows of the comb-like PEPO. If at the right boundary, then only rules where the right signal is 0 are kept. Rule 10 is only included in the top row.

| **Rule Number** | **Signals** (left, right, up/down) | **Output** |
|---|---|---|
| 1 | (0,0,n) | $\mathcal{I}$ |
| 2 | (1,0,n) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1,n) | $X^{[i]}\mathcal{I}$ |
| 4 | (2,1,y) | $w^{[i]}\mathcal{I}$ |
| 5 | (e,1,n) | $w^{[i]}\mathcal{A}$ |
| 6 | (3,0,y) | $\mathcal{A}$ |
| 7 | (2,2,n) | $\mathcal{I}$ |
| 8 | (3,3,n) | $\mathcal{I}$ |
| 9 | (e,e,n) | $\mathcal{I}$ |
| 10 | (2,0,y) | $\mathcal{B}$ |

Table 4.25: Rules for PEPO tensors to the right of the optimization column in the center rows of the comb-like PEPO. If at the right boundary, then only rules where the right signal is 0 are kept.

| **Rule Number** | **Signals** (left, right, down, up) | **Output** |
|---|---|---|
| 1 | (0,0,n,n) | $\mathcal{I}$ |
| 2 | (1,0,n,n) | $v^{[i]}\mathcal{B}$ |
| 3 | (1,1,n,n) | $X^{[i]}\mathcal{I}$ |
| 4 | (2,0,y,n) | $\mathcal{B}$ |
| 5 | (2,1,y,n) | $w^{[i]}\mathcal{I}$ |
| 6 | (2,1,n,y) | $w^{[i]}\mathcal{I}$ |
| 7 | (e,1,n,n) | $w^{[i]}\mathcal{A}$ |
| 8 | (3,0,y,n) | $\mathcal{A}$ |
| 9 | (3,0,n,y) | $\mathcal{A}$ |
| 10 | (2,2,n,n) | $\mathcal{I}$ |
| 11 | (3,3,n,n) | $\mathcal{I}$ |
| 12 | (e,e,n,n) | $\mathcal{I}$ |
| 13 | (0,0,y,y) | $\mathcal{I}$ |

Table 4.26: Rules for PEPO tensors in the bottom row of the optimization column, if the optimization column is the left-most column.

| **Rule Number** | **Signals** (right, up) | **Output** |
|---|---|---|
| 1 | (0,e) | $\mathcal{I}$ |
| 2 | (3,1) | $w^{[j]}\mathcal{I}$ |
| 3 | (e,0) | $\mathcal{I}$ |
| 4 | (2,2) | $w^{[j]}\mathcal{I}$ |
| 5 | (1,0) | $w^{[i]}\mathcal{A}$ |
| 6 | (0,1) | $w^{[j]}\mathcal{A}$ |
| 7 | (1,2) | $w^{[i]}w^{[j]}\mathcal{I}$ |

Table 4.27: Rules for PEPO tensors in the top row of the optimization column, if the optimization column is the left-most column.

| **Rule Number** | **Signals (right, down)** (right, down) | **Output** |
|---|---|---|
| 1 | (2,1) | $v^{[j]}\mathcal{I}$ |
| 2 | (0,0) | $\mathcal{I}$ |
| 3 | (e,e) | $\mathcal{I}$ |
| 4 | (3,2) | $v^{[j]}\mathcal{I}$ |
| 5 | (1,e) | $w^{[i]}\mathcal{A}$ |
| 6 | (0,1) | $v^{[j]}\mathcal{B}$ |
| 7 | (1,1) | $w^{[i]}v^{[j]}\mathcal{I}$ |
| 8 | (0,2) | $v^{[j]}\mathcal{A}$ |

Table 4.28: Rules for PEPO tensors in the center row of the optimization column, if the optimization column is the left-most column.

| **Rule Number** | **Signals** (right, down, up) | **Output** |
|---|---|---|
| 1 | (3,e,1) | $w^{[j]}\mathcal{I}$ |
| 2 | (0,1,1) | $X^{[j]}\mathcal{I}$ |
| 3 | (2,1,0) | $v^{[j]}\mathcal{I}$ |
| 4 | (0,e,e) | $\mathcal{I}$ |
| 5 | (e,e,0) | $\mathcal{I}$ |
| 6 | (0,0,0) | $\mathcal{I}$ |
| 7 | (2,e,2) | $w^{[j]}\mathcal{I}$ |
| 8 | (3,2,0) | $v^{[j]}\mathcal{I}$ |
| 9 | (0,2,2) | $X^{[j]}\mathcal{I}$ |
| 10 | (1,e,0) | $w^{[i]}\mathcal{A}$ |
| 11 | (0,e,1) | $w^{[j]}\mathcal{A}$ |
| 12 | (0,1,1) | $X^{[j]}\mathcal{I}$ |
| 14 | (0,1,0) | $v^{[j]}\mathcal{B}$ |
| 15 | (1,1,0) | $w^{[i]}v^{[j]}\mathcal{I}$ |
| 16 | (0,2,0) | $v^{[j]}\mathcal{A}$ |
| 17 | (0,2,2) | $X^{[j]}\mathcal{I}$ |
| 18 | (1,e,2) | $w^{[i]}w^{[j]}\mathcal{I}$ |

Table 4.29: Rules for PEPO tensors in the bottom row of the optimization column, if the optimization column is the right-most column.

| **Rule Number** | **Signals** (left, up) | **Output** |
|---|---|---|
| 1 | (0,0) | $\mathcal{I}$ |
| 2 | (e,e) | $\mathcal{I}$ |
| 3 | (1,e) | $v^{[i]}\mathcal{B}$ |
| 4 | (2,1) | $w^{[j]}\mathcal{I}$ |
| 5 | (3,2) | $w^{[j]}\mathcal{I}$ |
| 6 | (0,1) | $w^{[j]}\mathcal{A}$ |
| 7 | (1,1) | $v^{[j]}w^{[i]}\mathcal{I}$ |
| 8 | (0,2) | $w^{[j]}\mathcal{B}$ |

Table 4.30: Rules for PEPO tensors in the top row of the optimization column, if the optimization column is the right-most column.

| **Rule Number** | **Signals** (left, down) | **Output** |
|---|---|---|
| 1 | (0,e) | $\mathcal{I}$ |
| 2 | (e,1) | $v^{[j]}\mathcal{I}$ |
| 3 | (e,0) | $\mathcal{I}$ |
| 4 | (2,2) | $v^{[j]}\mathcal{I}$ |
| 5 | (1,0) | $v^{[i]}\mathcal{B}$ |
| 6 | (0,1) | $v^{[j]}\mathcal{B}$ |
| 7 | (1,2) | $v^{[i]}v^{[j]}\mathcal{I}$ |

Table 4.31: Rules for PEPO tensors in the center rows of the optimization column, if the optimization column is the right-most column.

| **Rule Number** | **Signals** (left, down, up) | **Output** |
|---|---|---|
| 1 | (3,1,e) | $v^{[j]}\mathcal{I}$ |
| 2 | (0,1,1) | $X^{[j]}\mathcal{I}$ |
| 3 | (2,0,1) | $w^{[j]}\mathcal{I}$ |
| 4 | (0,e,e) | $\mathcal{I}$ |
| 5 | (e,0,e) | $\mathcal{I}$ |
| 6 | (0,0,0) | $\mathcal{I}$ |
| 7 | (2,2,e) | $v^{[j]}\mathcal{I}$ |
| 8 | (3,0,2) | $w^{[j]}\mathcal{I}$ |
| 9 | (0,2,2) | $X^{[j]}\mathcal{I}$ |
| 10 | (1,0,e) | $v^{[j]}\mathcal{B}$ |
| 11 | (0,1,e) | $v^{[j]}\mathcal{B}$ |
| 12 | (0,0,1) | $w^{[j]}\mathcal{A}$ |
| 13 | (1,0,1) | $v^{[j]}w^{[i]}\mathcal{I}$ |
| 14 | (0,0,2) | $w^{[j]}\mathcal{B}$ |
| 15 | (1,2,e) | $v^{[i]}v^{[j]}\mathcal{I}$ |

Table 4.32: Rules for PEPO tensors in the bottom row of the optimization column, if the optimization column is any of the center columns.

| **Rule Number** | **Signals** (left, right, up) | **Output** |
|---|---|---|
| 1 | (2,0,1) | $w^{[j]}\mathcal{I}$ |
| 2 | (0,0,0) | $\mathcal{I}$ |
| 3 | (3,0,2) | $w^{[j]}\mathcal{I}$ |
| 4 | (e,0,e) | $\mathcal{I}$ |
| 5 | (0,0,2) | $w^{[j]}\mathcal{B}$ |
| 6 | (1,0,e) | $v^{[i]}\mathcal{B}$ |
| 7 | (1,0,1) | $v^{[i]}w^{[j]}\mathcal{I}$ |
| 8 | (0,0,1) | $w^{[j]}\mathcal{A}$ |
| 9 | (0,2,2) | $w^{[j]}\mathcal{I}$ |
| 10 | (0,3,3) | $w^{[j]}\mathcal{I}$ |
| 11 | (0,e,e) | $\mathcal{I}$ |
| 12 | (0,0,3) | $w^{[j]}\mathcal{A}$ |
| 13 | (0,1,e) | $w^{[j]}\mathcal{A}$ |
| 14 | (0,1,4) | $w^{[j]}w^{[i]}\mathcal{I}$ |
| 15 | (1,1,e) | $X^{[i]}\mathcal{I}$ |
| 16 | (1,1,5) | $w^{[j]} \otimes X^{[i]}\mathcal{I}$ |
| 17 | (1,1,6) | $X^{[i]} \otimes w^{[j]}\mathcal{I}$ |

Table 4.33: Rules for PEPO tensors in the top row of the optimization column, if the optimization column is any of the center columns.

| **Rule Number** | **Signals** (left, right, down) | **Output** |
|---|---|---|
| 1 | (0,0,e) | $\mathcal{I}$ |
| 2 | (3,0,1) | $v^{[j]}\mathcal{I}$ |
| 3 | (2,0,2) | $v^{[j]}\mathcal{I}$ |
| 4 | (e,0,0) | $\mathcal{I}$ |
| 5 | (0,0,1) | $v^{[j]}\mathcal{B}$ |
| 6 | (1,0,0) | $v^{[i]}\mathcal{B}$ |
| 7 | (1,0,2) | $v^{[i]}v^{[j]}\mathcal{I}$ |
| 8 | (0,3,2) | $v^{[j]}\mathcal{I}$ |
| 9 | (0,2,3) | $v^{[j]}\mathcal{I}$ |
| 10 | (0,1,3) | $v^{[j]}\mathcal{I}$ |
| 11 | (0,e,0) | $\mathcal{I}$ |
| 12 | (0,1,0) | $w^{[i]}\mathcal{A}$ |
| 13 | (0,1,4) | $v^{[j]}\mathcal{A}$ |
| 14 | (1,0,5) | $v^{[j]} \otimes \nVdash \mathcal{I}$ |
| 15 | (0,1,6) | $\nVdash \otimes v^{[j]}\mathcal{I}$ |

Table 4.34: Rules for PEPO tensors in the center rows of the optimization column, if the optimization column is any of the center columns.

| **Rule Number** | **Signals**<br>(left, right, down, up) | **Output** |
|---|---|---|
| 1 | (3,0,1,e) | $v^{[j]}\mathcal{I}$ |
| 2 | (2,0,0,1) | $w^{[j]}\mathcal{I}$ |
| 3 | (0,0,1,1) | $X^{[j]}\mathcal{I}$ |
| 4 | (2,0,2,e) | $v^{[j]}\mathcal{I}$ |
| 5 | (0,0,2,2) | $X^{[j]}\mathcal{I}$ |
| 6 | (3,0,0,2) | $w^{[j]}\mathcal{I}$ |
| 7 | (0,0,e,e) | $\mathcal{I}$ |
| 8 | (e,0,0,e) | $\mathcal{I}$ |
| 9 | (0,0,0,0) | $\mathcal{I}$ |
| 10 | (0,0,0,2) | $w^{[j]}\mathcal{B}$ |
| 11 | (1,0,0,e) | $v^{[i]}\mathcal{B}$ |
| 12 | (0,0,1,e) | $v^{[j]}\mathcal{B}$ |
| 13 | (1,0,0,1) | $v^{[i]}w^{[j]}\mathcal{I}$ |
| 14 | (0,0,0,1) | $w^{[j]}\mathcal{A}$ |
| 15 | (1,0,2,e) | $v^{[i]}v^{[j]}\mathcal{I}$ |
| 16 | (0,3,2,e) | $v^{[j]}\mathcal{I}$ |
| 17 | (0,2,0,2) | $w^{[j]}\mathcal{I}$ |
| 18 | (0,0,2,2) | $X^{[j]}\mathcal{I}$ |
| 19 | (0,2,3,e) | $v^{[j]}\mathcal{I}$ |
| 20 | (0,0,3,3) | $X^{[j]}\mathcal{I}$ |
| 21 | (0,3,0,3) | $w^{[j]}\mathcal{I}$ |
| 22 | (0,0,e,e) | $\mathcal{I}$ |
| 23 | (0,e,0,e) | $\mathcal{I}$ |
| 24 | (0,0,0,3) | $w^{[j]}\mathcal{A}$ |
| 25 | (0,1,0,e) | $w^{[i]}\mathcal{A}$ |
| 26 | (0,1,0,4) | $w^{[i]}w^{[j]}\mathcal{I}$ |
| 27 | (0,0,4,4) | $X^{[j]}\mathcal{I}$ |
| 28 | (0,0,4,e) | $v^{[j]}\mathcal{A}$ |
| 29 | (0,1,3,e) | $v^{[j]}w^{[i]}\mathcal{I}$ |
| 30 | (0,0,e,e) | $\mathcal{I}$ |
| 31 | (1,1,0,e) | $X^{[i]}\mathcal{I}$ |
| 32 | (0,0,5,5) | $X^{[j]} \otimes \mathbb{I}\mathcal{I}$ |
| 33 | (0,1,0,5) | $w^{[j]} \otimes X^{[i]}\mathcal{I}$ |
| 34 | (1,0,5,e) | $\mathbb{I} \otimes v^{[i]}\mathcal{I}$ |
| 35 | (0,1,6,e) | $v^{[j]} \otimes \mathbb{I}\mathcal{I}$ |
| 36 | (0,0,6,6) | $\mathbb{I} \otimes X^{[j]}\mathcal{I}$ |
| 37 | (1,0,0,6) | $X^{[i]} \otimes w^{[j]}\mathcal{I}$ |

# BIBLIOGRAPHY

[1] U. Schollwöck, *Annals of Physics* **2011**, *326*, 96–192.

[2] F. Verstraete, V. Murg, J. I. Cirac, *Advances in Physics* **2008**, *57*, 143–224.

[3] R. Orús, *Annals of Physics* **2014**, *349*, 117–158.

[4] G. Evenbly, G. Vidal, *Physical Review Letters* **2015**, *115*, 180405.

[5] P. Helms, G. K.-L. Chan, *Physical Review Letters* **2020**, *125*, 140601,

[6] P. Helms, U. Ray, G. K.-L. Chan, *Physical Review E* **2019**, *100*, 022101,

[7] T. H. Johnson, T. Elliott, S. R. Clark, D. Jaksch, *Physical Review Letters* **2015**, *114*, 090602.

[8] S. Efthymiou, J. Hidary, S. Leichenauer, *arXiv:1906.06329* **2019**.

[9] C. Roberts, A. Milsted, M. Ganahl, A. Zalcman, B. Fontaine, Y. Zou, J. Hidary, G. Vidal, S. Leichenauer, *arXiv:1905.01330* **2019**.

[10] A. Kardashin, A. Uvarov, J. Biamonte, *arXiv:1804.02398* **2018**.

[11] E. M. Stoudenmire, D. J. Schwab, *arXiv:1605.05775* **2016**.

[12] S. Cheng, L. Wang, P. Zhang, *Physical Review B* **2021**, *103*, 125117.

[13] Y. Zhou, E. M. Stoudenmire, X. Waintal, *Physical Review X* **2020**, *10*, 041038.

[14] W. Huggins, P. Patil, B. Mitchell, K. B. Whaley, E. M. Stoudenmire, *Quantum Science and technology* **2019**, *4*, 024001.

[15] Y. Pang, T. Hao, A. Dugad, Y. Zhou, E. Solomonik, *arXiv:2006.15234* **2020**.

[16] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, F. Verstraete, *Physical Review B* **2016**, *94*, 165116.

[17] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, C. Hubig, *Annals of Physics* **2019**, *411*, 167998.

[18] F. Verstraete, J. J. Garcia-Ripoll, J. I. Cirac, *Physical Review Letters* **2004**, *93*, 207204.

[19] E. Stoudenmire, S. R. White, *New Journal of Physics* **2010**, *12*, 055026.

[20] F. Verstraete, J. I. Cirac, *arXiv:cond-mat/0407066* **2004**.

[21] R. Haghshenas, M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2019**, *100*, 054404.

[22] K. Hyatt, E. M. Stoudenmire, *arXiv:1908.08833* **2019**.

[23] M. P. Zaletel, F. Pollmann, *Physical Review Letters* **2020**, *124*, 037201.

[24] F. Verstraete, J. I. Cirac, *Physical Review Letters* **2010**, *104*, 190405.

[25] R. Haghshenas, Z.-H. Cui, G. K.-L. Chan, *Physical Review Research* **2021**, *3*, 023057.

[26] A. Tilloy, J. I. Cirac, *Physical Review X* **2019**, *9*, 021040.

[27] M. Dolfi, B. Bauer, M. Troyer, Z. Ristivojevic, *Physical Review Letters* **2012**, *109*, 020604.

[28] G. K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, S. R. White, *The Journal of Chemical Physics* **2016**, *145*, 014102.

[29] E. M. Stoudenmire, S. R. White, *Physical Review Letters* **2017**, *119*, 046401.

[30] M. J. O'Rourke, G. K.-L. Chan, *Physical Review B* **2020**, *101*, 205142.

[31] M. J. O'Rourke, Z. Li, G. K.-L. Chan, *Physical Review B* **2018**, *98*, 205127.

[32] P. Corboz, *Physical Review B* **2016**, *94*, 035133.

[33] S. R. White, *Physical Review Letters* **1992**, *69*, 2863.

[34] M. Lubasch, J. I. Cirac, M.-C. Banuls, *New Journal of Physics* **2014**, *16*, 033014.

[35] M. Lubasch, J. I. Cirac, M.-C. Banuls, *Physical Review B* **2014**, *90*, 064425.

[36] J. Jordan, R. Orús, G. Vidal, F. Verstraete, J. I. Cirac, *Physical Review Letters* **2008**, *101*, 250602.

[37] F. Fröwis, V. Nebendahl, W. Dür, *Physical Review A* **2010**, *81*, 062337.

[38] L. Lin, Y. Tong, *SIAM Journal on Scientific Computing* **2021**, *43*, A164–A192.

[39] D. Braess, W. Hackbusch, *IMA journal of numerical analysis* **2005**, *25*, 685–697.

[40] G. Beylkin, L. Monzón, *Applied and Computational Harmonic Analysis* **2005**, *19*, 17–48.

[41] G. Beylkin, L. Monzón, *Applied and Computational Harmonic Analysis* **2010**, *28*, 131–149.

[42] H.-J. Liao, J.-G. Liu, L. Wang, T. Xiang, *Physical Review X* **2019**, *9*, 031041.

[43] L. Vanderstraeten, J. Haegeman, P. Corboz, F. Verstraete, *Physical Review B* **2016**, *94*, 155123.

[44] M. J. O'Rourke, Z. Li, G. K.-L. Chan, *Physical Review B* **2018**, *98*, 205127.

## *Chapter 5*

# CONCLUSIONS AND PROSPECTS

In this final chapter we provide an overview of the work presented in this thesis and discuss promising future directions of work at the intersection of tensor networks and classical nonequilibrium statistical mechanics.

## 5.1 Conclusions

The work presented in this thesis represents a continuous evolution in the application of tensor networks to studying classical nonequilibrium statistical mechanics (NESM) from simple one-dimensional models towards fully two-dimensional systems with long-range interactions. In this section, we will begin by summarizing the phenomenological insight we uncovered regarding the simple exclusion process then summarize our findings on the utility of tensor networks (TNs) in studying these systems. In the following section, we close by discussing potential areas of future theoretical and algorithmic work at the intersection of TNs and NESM.

Regarding the physical behavior of the asymmetric simple exclusion process (ASEP), the work presented in Chapters 2 and 3 used large deviation theory (LDT) to provide the first studies of both the fluctuations of the current and the dynamical phase behavior of the multi-lane and two-dimensional ASEP models. In one dimension, dynamical phase transitions have proven to be of particular physical importance as they are linked to dynamical heterogeneity. From the perspective of a single particle in the lattice, dynamical heterogeneity is manifest by stochastic transitions between multiple types of dynamical behaviors and the probability of these transitions is dictated by the distance between the dynamical phase transition and typicality. In the phase space we explored, the dynamical phase transition observed is between jammed and flowing phases, also called the high-density, low-density coexistence phase and the hyperuniform or maximal current phase respectively. In the one-dimensional system, the existence of this dynamical phase transition at typicality results in the shock phase where a jam forms in the lattice as particles wait to be removed and the starting point of this jam moves stochastically through the lattice.

In the multi-lane system, we discovered that the existence of the phase transition depends strongly on the latitudinal hopping parameters and particularly the insertion and removal rates at the top and bottom boundaries. When these boundaries allow particles to exit the lattice relatively quickly, i.e. open boundary conditions, the particles will exit the lattice before becoming jammed, circumventing the dynamical phase transition. When these boundaries are instead closed, the multi-lane behavior mimics what is seen in one dimension. Moving to the fully two-dimensional system, we discovered that the jammed phase persists given appropriate physical parameters. In this case, the dynamical phase transition does not occur at typicality, but we expect that parameters can be tuned to cause the two-dimensional shock phase.

In this thesis, the utility of the TN algorithms, introduced for computing large deviation functions in the classical nonequilibrium setting, was of central importance. In Chapter 2, we demonstrated that, like in the quantum setting, a one-dimensional TN, the matrix product state (MPS), and an associated optimization algorithm, the density matrix renormalization group (DMRG) algorithm, provide accurate and compact state representations. Specifically, we found that we were able to compute large deviation functions with 3-4 digits of accuracy using an MPS with a bond dimension of $O(100)$, while state-of-the-art calculations allow bond dimensions that are two orders of magnitude larger. The primary algorithmic difference between the classical and quantum setting is the non-hermiticity of the Hamiltonian, here the tilted generator, which can be overcome by using a non-hermitian local eigensolver. We were also able to use the MPS to study multi-lane systems, though the width of the lattice was limited, as is common for quantum systems.

Moving to two dimensions, we employed time evolution via block decimation (TEBD) algorithms to compute formally two-dimensional TN states, called projected entangled pair states (PEPS), and associated large deviation functions. There were no algorithmic differences between these classical nonequilibrium calculations and the standard quantum implementations and we found that we were again able to compute large deviation functions and characterize system behavior with relatively low computational effort. Here, we required a bond dimension of 2 to 3 while state-of-the-art calculations use bond dimensions of $< 20$.

While the results presented in Chapters 2 and 3 strongly emphasize the utility of TNs in this context, TN algorithms are young and there remain many opportunities for algorithmic improvements to extend their sucess to more complex settings. Importantly, many novel numerical routines should be extensible to both quantum and classical systems. As an example of this, Chapter 4 presented work towards studying two-dimensional systems with long-range interactions. The complexity of this problem was highlighted by introducing three seemingly equivalent approaches and demonstrating the counterintuitive comparative computational cost scaling with system size. While we hope to soon use these methods for realistic systems, the progress reported there showed two viable approaches and is promising for future work in this area.

## 5.2 Prospective Work

The work presented in this thesis has provided strong evidence for the effectiveness of using TNs to study fluctuations in nonequilibrium classical systems and we expect that these results will make this a commonly used tool in this field. To extend the applicability of TNs to the complex chemical or physical systems practitioners study provides many opportunities for algorithmic developments. For example, while we have introduced algorithms to study particles with long-range interactions, such as the Lennard-Jones potential commonly used in statistical mechanics, these particles are normally not confined to a lattice, thus requiring algorithms that allow continuous spatial degrees of freedom. Additionally, our results have been limited to one- and two-dimensional TNs and systems; to continue this trajectory we would require the development of TN algorithms for fully three-dimensional systems. These results have also all assumed relatively regular system geometries, here square lattices; to extend the success of TNs to study complex stochastic systems, such as gene expression networks, we would also require methods to treat non-standard geometries. Finally, a common theme in quantum TN algorithm development has been the creation and optimization of formally infinite TNs. While infinite algorithms would allow us to circumvent finite-size scaling used with Monte Carlo results, one must consider how to incorporate boundary effects that are commonly present in the nonequilibrium setting.

In addition to these numerical developments that would make TNs a powerful tool in classical NESM, the use of TNs may also provide theoretical insights that would be difficult to recognize using other methods. For example, in the quantum setting, the size of the tensors is directly related to the entanglement between parts of the lattice and TNs give direct access to the system's entanglement entropy. While this quantity is still computed when using TNs for classical systems, the physical interpretation is unclear and could provide deeper insight into the physical behavior. In quantum systems, this easy access to the entanglement entropy provided a route towards finding area laws and detecting topological phase transitions. While we are unaware of classical analogs for either of these, the formalism of TNs will allow for their straightforward discovery if they do indeed exist.

In summary, TNs are powerful numerical tools for studying fluctuations in the nonequilibrium setting, capable of providing meaningful physical insight and numerical accuracy. Promising areas for future developments in TN algorithms, applications, and theoretical interpretations make them well-poised to play an important role in furthering our understanding of nonequilibrium systems and behaviors.