# Non-equilibrium strongly-correlated dynamics


**Tomi Harry Johnson**
Worcester College, Oxford

A thesis submitted to the Mathematical and Physical Sciences Division
for the degree of Doctor of Philosophy at the University of Oxford

Hilary Term, 2013

Atomic and Laser Physics,
University of Oxford

# Non-equilibrium strongly-correlated dynamics

Tomi Harry Johnson, Worcester College

A thesis submitted to the Mathematical and Physical Sciences Division
for the degree of Doctor of Philosophy at the University of Oxford

Hilary Term, 2013

**Abstract**

We study non-equilibrium and strongly-correlated dynamics in two contexts. We begin by analysing quantum many-body systems out of equilibrium through the lens of cold atomic impurities in Bose gases. Such highly-imbalanced mixtures provide a controlled arena for the study of interactions, dissipation, decoherence and transport in a many-body quantum environment.

Specifically we investigate the oscillatory dynamics of a trapped and initially highly-localised impurity interacting with a weakly-interacting trapped quasi low-dimensional Bose gas. This relates to and goes beyond a recent experiment by the Inguscio group in Florence. We witness a delicate interplay between the self-trapping of the impurity and the inhomogeneity of the Bose gas, and describe the dissipation of the energy of the impurity through phononic excitations of the Bose gas.

We then study the transport of a driven, periodically-trapped impurity through a quasi one-dimensional Bose gas. We show that placing the weakly-interacting Bose gas in a separate periodic potential leads to a phononic excitation spectrum that closely mimics those in solid state systems. As a result we show that the impurity-Bose gas system exhibits phonon-induced resonances in the impurity current that were predicted to occur in solids decades ago but never clearly observed. Following this, allowing the bosons to interact strongly, we predict the effect of different strongly-correlated phases of the Bose gas on the motion of the impurity. We show that, by observing the impurity, properties of the excitation spectrum of the Bose gas, e.g., gap and bandwidth, may be inferred along with the filling of the bosonic lattice. In other words the impurity acts as a probe of its environment.

To describe the dynamics of such a strongly-correlated system we use the powerful and near-exact time-evolving block decimation (TEBD) method, which we describe in detail. The second part of this thesis then analyses, for the first time, the performance of this method when applied to simulate non-equilibrium classical stochastic processes. We study its efficacy for a well-understood model of transport, the totally-asymmetric exclusion process, and find it to be accurate.

Next, motivated by the inefficiency of sampling-based numerical methods for high variance observables we adapt and apply TEBD to simulate a path-dependent observable whose variance increases exponentially with system size. Specifically we calculate the expected value of the exponential of the work done by a varying magnetic field on a one-dimensional Ising model undergoing Glauber dynamics. We confirm using Jarzynski's equality that the TEBD method remains accurate and efficient. Therefore TEBD and related methods complement and challenge the usual Monte Carlo-based simulators of non-equilibrium stochastic processes.

# Acknowledgements


### Acknowledgements of authorship

This thesis features work that my colleagues and I undertook collaboratively [1, 2, 3, 4, 5]. Therefore I would like to thank and acknowledge all of my coauthors for their role in these studies: Stephen Clark, Dieter Jaksch, Martin Bruderer, Anna Posazhennikova, Wolfgang Belzig, Yongyong Cai, Weizhu Bao and Thomas Elliott.

Of those who deserve particular acknowledgement, I first thank Stephen Clark, who supported me perhaps more than is usual for a second author during the writing of my first publication [1]. This work is presented with modifications in parts of Sec. 3.2 and Ch. 9. Second, I acknowledge Martin Bruderer, the first author of Ref. [2]. This is featured alongside other work in Sec. 5.1. In particular, he first performed the analytical derivation of the atomic current that I present with modifications in Sec. 5.1.4. Third, Yongyong Cai and Weizhu Bao together wrote the code I used to solve the coupled Gross-Pitaevskii equations in Ref. [4], presented here in Sec. 4.2.4. Fourth, while on a summer placement Thomas Elliott ran, with my assistance, some of the simulations used in Ch. 10, which we are now preparing for publication [5]. Last, the fingerprints of Dieter Jaksch can be found throughout the works referenced above as well as this thesis. Their successes are in no small part down to his direction.

### Other acknowledgements

I must also thank my other coauthors, Jacob Biamonte, Pinja Haikka, Sabrina Maniscalco, Richard Walters, Giovanni Cotugno, Mauro Faccin, Sabre Kais, Piotr Migdał, and Ville Bergholm, whose joint work [6, 7, 8, 9, 10] I have unfortunately not had space to include in this thesis. It is not only these coauthors who have influenced my work but other collaborators with whom I have shared interesting discussions that have yet to lead to publications: Vlatko Vedral, John Goold, Pierre-Louis Giscard, Chris Foot, Dominik Hangleiter, Thomas Elliott and Suzanne McEndoo.

My office-mates have kept me company and made the years most enjoyable; it has been a pleasure to work, chat and drink tea with Stephen Clark, Javier Prior, Benjamin Rowlands, Libby Heaney, Thomas Grujic, Manuel Pino, Craig Morgan, Simon Fry, David Mcgonegle, Jaskaran Kahlon, Thomas Elliott, Giovanni Cotugno, Sarah Al-Assam and Dominik Hangleiter (though they were not all in the office at the same time!). A similar thanks must be extended to numerous others in the Clarendon Laboratory, especially those in the groups of Dieter Jaksch, Vlatko Vedral, Chris Foot and Igor Mekhov.

Importantly, I am eternally grateful to my supervisors Dieter Jaksch and Vlatko Vedral for the huge amount of time they have put into supporting, mentoring and teaching me, from my undergraduate degree all the way to my search for a first Post Doc position. Similarly Stephen Clark has had the bad luck to share an office with me for the duration of my DPhil (excellent luck from my point of view), thereby making him the principal victim of my endless questioning. His answers have certainly shaped this thesis.

It would be most unwise to go without the acknowledging the ever-important sources of money that have allowed me to perform the work presented in this thesis: the EPSRC and the Department of Physics for their doctoral training award, and CTQ Singapore for supporting me during two academic visits.

To end on a personal note, I would like thank my family and friends for providing the perfect level of distraction from my work. I am most grateful to my parents for their immeasurable support in innumerable areas of my life. My father, for instance, has been single-handedly building my house around me while I wrote this thesis. Finally, thank you to Jessica for sharing the whole experience with me and making it infinitely more enjoyable; any correct uses of the semicolon in this thesis are entirely down to her influence.

# Contents

# Chapter 1

# Introduction

Non-equilibrium models are fundamental to our understanding of the world around us. They describe the process by which things are measured, computations are performed, and energy and mass are transported[1]. Yet non-equilibrium models are a challenge to analyse and their properties are not as well-understood as those of their equilibrium counterparts.

Similarly, many fascinating and important systems, e.g., magnets, financial markets or traffic, comprise many strongly-correlated parts. Rarely are the effects of one part decoupled from another and so the combined system, if it is to be understood, must be treated as a whole. This again poses a challenge, numerically as well as analytically.

The aim of this thesis is to study non-equilibrium and strongly-correlated phenomena in two settings. First we look at atomic impurities moving in a gas of cold bosonic atoms. We ask how the environment of the impurity affects its motion, the extent to which the impurity probes the environment and the accuracy with which the combined system simulates driven electronic motion in solids. Second we move from quantum to purely classical dynamics. We examine non-equilibrium stochastic processes, the difficulties in their numerical simulation and adapt methods to

[1] They even describe the biological functioning the reader!

overcome these difficulties.

An additional connection between these two strands is provided by the numerical method that we adapt, the time-evolving block decimation (TEBD) method [11, 12]. In the first part of this thesis we use the method, alongside other analytical and numerical techniques, in the quantum mechanical setting for which it is well-known. In the second part of the thesis we use TEBD to simulate classical stochastic processes for the first time and evaluate its performance relative to the current state of the art.

Let us now separately examine the contexts of the two parts of this thesis, starting with impurities immersed in Bose gases. Cold atom experiments featuring impurities have become increasingly popular, and demonstrate growing levels of control and flexibility. To give several relevant and recent examples, the group of Köhl used *in situ* imaging to observe the gravity-driven transport of an impurity wave-packet through a Tonks-Girardeau gas [13]. Inguscio and collaborators observed the oscillatory dissipative dynamics of a trapped and initially highly-localised impurity in a quasi one-dimensional Bose gas [14]. With his colleagues, Bloch observed the dephasing dynamics of superpositions of zero and one fermionic impurities interacting with bosons in each well of an optical lattice. They used this to measure the boson-fermion interaction strength [15]. Similar in vein, Widera *et al.* observed the fluorescence dynamics of a few impurities in a Bose gas, which allowed them to infer the inter-species scattering length [16]. The same group has recently demonstrated the insertion of a single impurity, or a precisely controlled number of impurities, into a bosonic background gas [17].

Inspired partly by these experimental successes theorists have explored the possible uses of dynamical impurities in cold atom gases. It has been found, for example, that the decohering or dissipative dynamics of an impurity measures the phase fluctuations in a Bose gas [18] or, in one dimension, the Luttinger parameter [19]. The

dephasing of an atom in a Bose gas has been proposed as a highly-controllable system in which to study the onset of non-Markovianity [20] and observe Anderson's orthogonality catastrophe in an ideal Fermi gas [21, 22]. Immersed impurities also mimic collisionally-induced electron conduction in solids [23] as well as condensed matter phenomena such as polarons [24].

The work forming the first part of this thesis is closely related to the above developments. We begin with a low-temperature theoretical analysis of the oscillatory dissipative dynamics observed by the Inguscio group (see above) [14]. Then we improve the fidelity with which an impurity-Bose gas systems simulates collisionally-induced electron conduction. We do this by trapping the Bose gas in a periodic potential to manipulate the spectrum of its elementary phononic excitations. Finally we consider the same system but with a deeper bosonic lattice potential, whereby the now strongly-interacting bosons are described by the one-dimensional Bose-Hubbard model [25, 26]. We study the extent to which the impurity may be used as a probe of the strongly-correlated phases of this model.

These strongly-correlated phases are typically more difficult to describe analytically than those of weakly-interacting bosons. Therefore we use the TEBD method [11, 12], a numerical method closely related to density-matrix renormalisation group (DMRG) algorithms [27, 28, 29], that near-exactly simulates strongly-correlated dynamics in one-dimensional quantum lattice systems.

This discussion leads nicely to the historical context of the second part of this thesis. Following the success of DMRG, but prior to the invention of TEBD in 2003, researchers explored the application of DMRG methods to classical stochastic non-equilibrium processes. They began by applying DMRG to find the stationary distributions of archetypal processes, finding that they are in principle well-approximated using this type of method. However their particular method became numerically unstable for large systems [30, 31, 32]. Using related algorithms [33, 34, 35] they

considered transient dynamics in the thermodynamic limit, however, the algorithms broke down before they reached very long times. In the second part of this thesis we follow in the footsteps of these researchers, but using the TEBD method. We simulate the dynamics of classical Markovian processes and also, by considering long times, find their stationary distributions. Our adaption, as we will reveal, shows no sign of the long-time or large-system-size instabilities that plagued DMRG approaches.

The motivation for adapting quantum-inspired algorithms for the field of non-equilibrium stochastic processes is clear from a brief review (which we provide in this thesis) of the numerical methods typically used to simulate these processes. The principal method is dynamical Monte Carlo (DMC) [36, 37, 38, 39, 40], a method that has been studied for half a century by thousands of researchers. It is hugely adaptable and there exist many variants, each making a specific type of simulation more efficient. However, at its heart, DMC is a sampling method. Paths through configuration space are sampled and the values an observable takes for the sampled paths are averaged to estimate the expected value of that observable. As a result, to achieve a given accuracy the required number of sampled paths increases linearly with the variance of the observable of interest. This makes the method less useful for high variances. Much effort [41, 42, 43, 44, 45, 46, 47] has been put in to reducing the computational resources needed to simulate high variance observables, however, there is no systematic approach that works generally. Our aim in introducing TEBD to the stochastic process community is to provide a complementary method with a different source of error, i.e., it is not sampling based. Therefore our main focus in the second part of this thesis is to demonstrate that TEBD efficiently simulates high variance dynamical path-dependent observables in strongly-correlated stochastic processes.

To close this introduction we describe the layout of the thesis. As mentioned

above it is divided into two parts; Part I concerns the dynamics of impurity atoms in Bose gases and Part II the simulation of non-equilibrium classical stochastic processes.

Part I is broken down as follows. We begin with two introductory chapters. Chapter 2 contains our summary of the relevant theoretical descriptions of Bose gases, in harmonic traps or optical lattices. Then in Ch. 3 we review the numerical methods that we will use in conjunction with these descriptions. As part of this the TEBD algorithm is described in detail. The next pair of chapters contain our novel contribution to the field of cold atoms. First in Ch. 4, inspired by the experiment in Ref. [14], we study the so-called breathing of an initially highly-localised impurity as its width oscillates in a quasi one-dimensional harmonic trap at low temperatures. We show how its dynamics are affected by its immersion in a weakly-interacting Bose gas. The two main effects are self-trapping, which is the trapping of the impurity by a self-induced density dip in the surrounding Bose gas, and the dissipation of energy into the phonon modes of the Bose gas. This dissipation process is starkly different to any friction force. Second, in Ch. 5 we study a driven impurity in a periodic potential, where the motion of the impurity is induced by collisions with a quasi one-dimensional Bose gas. If the bosons are weakly-interacting this models the electronic current in a solid when a voltage is applied. We consider the effects of a separate periodic potential on the Bose gas and show that it reproduces the Van Hove singularity found in the density of phonon states at the band edge in solids. This reveals itself through resonances in the current. We then analyse the dynamics of the impurity if the bosons are strongly-interacting and consider how the dynamics changes for the different strongly-correlated phases of the Bose gas. The emphasis is put on how the impurity acts as a probe of its environment, with details of the spectrum of the environment, e.g., gap and bandwidth, imprinted onto the dependence of the current of the impurity on forcing.

A description of the layout of Part II is left until the beginning of that part.

# Part I

# Trapped cold neutral atoms

# Chapter 2

# Bose gases

In this part we study the non-equilibrium dynamics of atomic impurities immersed in a gas of bosonic atoms. We therefore begin by discussing Bose gases in this chapter. We cannot hope to give a complete picture of the decades of work relating to this topic and so details are left to the many excellent reviews and books cited in this chapter. Our aim here is to summarise the properties and common theoretical descriptions of Bose gases that we will make use of in later chapters.

The chapter is laid out as follows. In Sec. 2.1 we discuss the condensation of a Bose gas and how the condensate wavefunction is described by the Gross-Pitaevskii (GP) equations. Further we introduce the Bogoliubov modes, which capture the elementary collective excitations of the bosons beyond the condensate. Then in Sec. 2.2 we look at quasi low-dimensional Bose gases, the effects of phase fluctuations and the formation of quasi condensates. We discuss in what regimes the GP and Bogoliubov approaches still describe the Bose gas. Following this, in Sec. 2.3 we move on to bosons in optical lattices, focusing on their realisation of the one-dimensional Bose-Hubbard model. Finally we summarise in Sec. 2.4.

## 2.1 Three-dimensional Bose gas

### 2.1.1 Ideal Bose gas

#### Homogeneous system

Many properties of Bose gases at low temperatures are closely related to the phenomenon of Bose-Einstein condensation (BEC). This phenomenon is most easily introduced via the ideal (non-interacting but thermally equilibrated) Bose gas in a three-dimensional box of volume $\mathcal{V}$ with periodic boundary conditions. We write the mass of the bosons as $m_b$. In second quantisation the bosonic Hamiltonian is written

$$\hat{H}_b = \int \mathrm{d}\mathbf{r} \hat{\phi}^\dagger(\mathbf{r}) h(\mathbf{r}) \hat{\phi}(\mathbf{r}), \tag{2.1}$$

where the single-particle Hamiltonian is

$$h(\mathbf{r}) = -\frac{\hbar^2 \nabla^2}{2m_b}. \tag{2.2}$$

Here $\hat{\phi}^\dagger(\mathbf{r})$ and $\hat{\phi}(\mathbf{r})$ are the boson field operators. The single-particle energy eigenfunctions are the momentum eigenfunctions $\varphi_{\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}/\sqrt{\mathcal{V}}$ with energies $\varepsilon_{\mathbf{k}} = \hbar^2 k^2 / 2m_b$. Choosing to initially work in the grand canonical ensemble, their average occupation is given by the Bose distribution

$$N_{\mathbf{k}} = \frac{1}{\mathrm{e}^{\beta(\varepsilon_{\mathbf{k}} - \mu_b)} - 1}, \tag{2.3}$$

where $\beta = 1/k_B T$ is the inverse temperature and the chemical potential $\mu_b$ fixes the average number of bosons $N = \sum_{\mathbf{k}} N_{\mathbf{k}}$. Following e.g. Ref. [48], approximating the discrete set of momentum eigenfunctions by a continuum we find that the occupation

of the zero-momentum eigenfunction obeys

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^{3/2}, \tag{2.4}$$

below a critical temperature

$$T_c = \frac{2\pi\hbar^2}{k_B m_b}\left(\frac{n}{g_{3/2}(1)}\right)^{2/3}, \tag{2.5}$$

and is zero above this temperature. Here $n = N/\mathcal{V}$ is the density of bosons and $g_{3/2}(1) \approx 2.612$. We see that below a critical temperature $T_c$ a single-particle eigenfunction is macroscopically occupied. The phenomenon of a macroscopically occupied single-particle modefunction is called BEC[1].

While it may not necessarily be the case that the grand canonical ensemble is always the correct ensemble to work within, note that the phenomenon of BEC and the other results presented above may also be derived in the canonical [49] and microcanonical [50] ensembles.

### Spherically symmetric harmonic trap

A similar analysis may be performed for bosons in a harmonic trap, in which case the single-particle Hamiltonian is

$$h(\mathbf{r}) = -\frac{\hbar^2\nabla^2}{2m_b} + v_b(\mathbf{r}), \tag{2.6}$$

with $v_b(\mathbf{r}) = \frac{1}{2}m_b\Omega_b^2 r^2$. The corresponding single-particle eigenfunctions and energies have the well-known harmonic oscillator form (see e.g. Ref. [51]).

[1] Many definitions of BEC demand that the temperature is finite or require that the continuum, or thermodynamic, limit is taken. Here we merely require that a large fraction of a large number of bosons occupy the same single-particle mode, since this is the only criterion relevant to the theoretical treatments that we will use.

As for the homogenous gas the lowest energy eigenfunction, this time a Gaussian, has a macroscopic occupation [48]

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_c}\right)^3, \tag{2.7}$$

below a critical temperature

$$T_c = \frac{\hbar\Omega_b}{k_B}\left(\frac{N}{\zeta_3}\right)^{1/3}, \tag{2.8}$$

where $\zeta_3 \approx 1.202$.

Again the above features may also be obtained by working in the canonical ensemble. A detailed discussion of the predictions of different ensembles is found in Refs. [52, 53].

### 2.1.2 Weakly-interacting Bose gas

A realistic description of Bose gases must include interactions between the bosons. Here we include two-body interactions via a potential $v_{\rm two-body}$, such that the full Hamiltonian is written

$$\hat{H}_b = \int \mathrm{d}\mathbf{r}\hat{\phi}^\dagger(\mathbf{r})h(\mathbf{r})\hat{\phi}(\mathbf{r}) + \frac{1}{2}\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\hat{\phi}^\dagger(\mathbf{r})\hat{\phi}^\dagger(\mathbf{r}')v_{\rm two-body}(\mathbf{r}-\mathbf{r}')\hat{\phi}(\mathbf{r}')\hat{\phi}(\mathbf{r}). \tag{2.9}$$

We denote the $s$-wave scattering length of the potential $v_{\rm two-body}$ by $a_s$, which for the potentials in this thesis will always be positive. A standard approximation in cold atom physics is to replace $v_{\rm two-body}(\mathbf{r})$ by the pseudo-potential $g\delta(\mathbf{r})$ [54, 55, 56, 57], where $g = 4\pi a_s\hbar^2/m_b$ is called the interaction strength. This approximation is valid so long as the temperature $T$ is low such that the thermal wavelength $\lambda_T =$

$\sqrt{2\pi\hbar^2/m_b k_B T}$ is much larger than the scattering length[2] $a_s$. Further the expected distance between particles must be large compared to $a_s$. The first condition

$$\lambda_T \gg a_s, \tag{2.10}$$

defines the temperatures at which this description is valid. For bosonic number densities $n$ that vary slowly on the scale of the typical inter-boson distance the second condition is given by

$$n a_s^3 \ll 1. \tag{2.11}$$

This expresses what is meant by our restriction to weakly-interacting bosons, where the number density of bosons $n$ may be spatially dependent.

A typical experiment with cold bosonic atoms might feature number densities $n \approx 10^{13}$ cm$^{-3}$ thus inter-boson distances $n^{-1/3} \approx 1$ $\mu$m. Meanwhile scattering lengths are on the order $a_s \approx$ nm, so the weakly-interacting condition of Eq. (2.11) is fulfilled [58]. For masses $m_b \approx 50$ amu the low temperature condition of Eq. (2.10) is met provided $T \ll 1$ mK, which is readily achieved. In what follows we work within this regime and so write

$$\hat{H}_b = \int \mathrm{d}\mathbf{r} \hat{\phi}^\dagger(\mathbf{r}) \left( h(\mathbf{r}) + \frac{g}{2} \hat{\phi}^\dagger(\mathbf{r}) \hat{\phi}(\mathbf{r}) \right) \hat{\phi}(\mathbf{r}). \tag{2.12}$$

### GP condensate

Inspired by the $T = 0$ results discussed in the previous section, an approach (see e.g. Ref. [59]) for describing interacting bosons relies on the *a priori* assumption of the macroscopic occupation of a single-particle modefunction $\varphi_0(\mathbf{r})$, by $N \gg 1$ bosons[3].

[2] If the Bose gas is out of thermal equilibrium then the condition is that the kinetic energy of bosons must be much smaller than $\hbar^2/2m_b a_s^2$.

[3] Unlike for the ideal gas, $\varphi_0(\mathbf{r})$ is not necessarily an eigenfunction of the single-particle Hamiltonian.

Following this approach we make the Hartree-Fock ansatz for the ground state and write the state of the condensed bosons as

$$|\,\Psi\rangle = (1/\sqrt{N!})(\hat{a}_0^\dagger)^N\,|\,0\rangle\,, \tag{2.13}$$

where $\hat{a}_0^\dagger = \int \mathrm{d}\mathbf{r}\varphi_0(\mathbf{r})\hat{\phi}^\dagger(\mathbf{r})$ creates a boson in the condensate and $|\,0\rangle$ is the vacuum state with no bosons. To obtain the ground state we minimise $\langle\Psi\,|\,(\hat{H}_b - \mu_b\hat{N})\,|\,\Psi\rangle$, where $\mu_b$ is the Lagrange multiplier associated with fixing the number of bosons in the condensate and $\hat{N} = \int \mathrm{d}\mathbf{r}\hat{\phi}^\dagger(\mathbf{r})\hat{\phi}(\mathbf{r})$ is the corresponding operator. For large $N \gg 1$ this results in the time-independent $T = 0$ GP equation [60, 61, 62]

$$\left(h(\mathbf{r}) - \mu_b + g|\varphi_0(\mathbf{r})|^2\right)\varphi_0(\mathbf{r}) = 0, \tag{2.14}$$

where we have renormalised the condensate wavefunction such that $\int \mathrm{d}\mathbf{r}|\varphi_0(\mathbf{r})|^2 = N$. It also follows from the analysis that $\mu_b$ is the chemical potential.

There are several other ways to arrive at Eq. (2.14) (see e.g. Refs. [63, 64, 48]). Further in Ref. [65] it is shown, in the limit of infinite particle number $N \to \infty$ with fixed $Na_s$, that Eq. (2.13) describes the exact ground state.

The same treatment can be extended to describe the (possibly non-equilibrium) evolution of the condensate wavefunction by solving the time-dependent Schrödinger equation within the same ansatz of Eq. (2.13), allowing the condensate function $\varphi_0(\mathbf{r}, t)$ and trapping potential $v_b(\mathbf{r}, t)$ to be time-dependent. This leads to the time-dependent GP equation

$$\left(h(\mathbf{r}, t) + g|\varphi_0(\mathbf{r}, t)|^2\right)\varphi_0(\mathbf{r}, t) = \mathrm{i}\hbar\frac{\partial}{\partial t}\varphi_0(\mathbf{r}, t). \tag{2.15}$$

The validity of the ansatz in Eq. (2.13) is discussed in Refs. [66, 67, 68], where it is found to be correct to the lowest order in $na_s^3$, i.e., valid for weakly-interacting

bosons.

If the condensate initially takes its equilibrium value $\varphi_0(\mathbf{r})$, satisfying the time-independent GP Eq. (2.14), then it evolves as $\varphi_0(\mathbf{r},t) = \varphi_0(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mu_b t/\hbar}$. Let us now look for solutions to the time-dependent GP Eq. (2.15) if the condensate is initially slightly deviated from its equilibrium value $\varphi_0(\mathbf{r})$, i.e., given by $\varphi_0(\mathbf{r}) + \delta\varphi(\mathbf{r},0)$ the equilibrium value plus some small deformation $\delta\varphi(\mathbf{r},0)$. We make the ansatz $[\varphi_0(\mathbf{r}) + \delta\varphi(\mathbf{r},t)]\,\mathrm{e}^{-\mathrm{i}\mu_b t/\hbar}$ for the condensate with $\delta\varphi(\mathbf{r},t) = \sum_\nu [u_\nu(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega_\nu t} + v_\nu^*(\mathbf{r})\mathrm{e}^{\mathrm{i}\omega_\nu t}]$. Neglecting all terms higher than second order in the deformation $\delta\varphi$, the ansatz provides a valid solution to Eq. (2.15) if $u_\nu$, $v_\nu$ and $\hbar\omega_\nu$ solve the Bogoliubov-de Gennes equations [69, 48]

$$h(\mathbf{r})u_\nu(\mathbf{r}) + g|\varphi_0(\mathbf{r})|^2(2u_\nu(\mathbf{r}) + v_\nu(\mathbf{r})) = (\mu_b + \hbar\omega_\nu)u_\nu(\mathbf{r}), \tag{2.16a}$$

$$h(\mathbf{r})v_\nu(\mathbf{r}) + g|\varphi_0(\mathbf{r})|^2(2v_\nu(\mathbf{r}) + u_\nu(\mathbf{r})) = (\mu_b - \hbar\omega_\nu)v_\nu(\mathbf{r}). \tag{2.16b}$$

The Bogoliubov-de Gennes solutions therefore describe the collective dynamical excitations of a condensate around its $T = 0$ equilibrium value.

### Elementary excitations

Here we introduce the low-energy quantised excitations of the weakly-interacting gas, first found for a uniform system by Bogoliubov [70]. Perhaps the (mathematically) simplest way to arrive at these excitations for a trapped gas (see e.g. Ref. [71]) is to follow Bogoliubov and expand the field operator

$$\hat{\phi}(\mathbf{r}) = \varphi_0(\mathbf{r}) + \delta\hat{\phi}(\mathbf{r}). \tag{2.17}$$

This is the sum of a $c$-number $\varphi_0(\mathbf{r})$, which solves the GP Eq. (2.14), and a deformation $\delta\hat{\phi}(\mathbf{r})$. Note that $\varphi_0(\mathbf{r})$ is now normalised to the average number of particles in

the condensate $N_0 \approx N$, which may be depleted from the total number $N$. The deformation $\delta\hat{\phi}(\mathbf{r})$ then represents excitations on top of the condensate and is assumed to be small.

We first insert the expansion of Eq. (2.17) into the Hamiltonian of Eq. (2.12) and neglect terms higher than second order in the deformation $\delta\hat{\phi}$. Then expressing the deformation $\delta\hat{\phi} = \sum_\nu [u_\nu(\mathbf{r})\hat{d}_\nu + v_\nu^*(\mathbf{r})\hat{d}_\nu^\dagger]$ in terms of bosonic Bogoliubov modes $\hat{d}_\nu$, the Hamiltonian of the Bose gas satisfies, up to the addition of a constant,

$$\hat{H}_b - \mu_b \hat{N} = \sum_\nu \hbar\omega_\nu \hat{d}_\nu^\dagger \hat{d}_\nu \,. \tag{2.18}$$

Here $u_\nu$, $v_\nu$ and $\hbar\omega_\nu$ solve the Bogoliubov-de Gennes Eqs. (2.16). This demonstrates that there is a direct correspondence between the elementary single-phonon Bogoliubov states and the small dynamical oscillations of the condensate considered in the previous subsection.

Extending the above Bogoliubov approach it is found that $\hat{H}_b - \mu_b \hat{N}$ in Eq. (2.18) is also the Hamiltonian governing the dynamics of the system [48].

For a Bose gas in a volume $\mathcal{V}$ with periodic boundary conditions and $v_b(\mathbf{r}) = 0$ the Bogoliubov-de Gennes Eqs. (2.16) may be solved analytically (see e.g. Ref. [48]). In this case the condensate $\varphi_0(\mathbf{r}) = \sqrt{n_0}$ is uniform, where $n_0 = N_0/\mathcal{V}$. The Bogoliubov operator $\hat{d}_\mathbf{k}^\dagger$ then creates a phonon with momentum $\hbar\mathbf{k} \neq 0$. The phonons describe plane waves $u_\mathbf{k}(\mathbf{r}) = \tilde{u}_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}}$, $v_\mathbf{k}(\mathbf{r}) = \tilde{v}_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}}$ with

$$\tilde{u}_\mathbf{k} = \frac{1}{\sqrt{2\mathcal{V}}} \left( \frac{\varepsilon_\mathbf{k} + g n_0}{\hbar\omega_\mathbf{k}} + 1 \right)^{1/2}, \tag{2.19a}$$

$$\tilde{v}_\mathbf{k} = -\frac{1}{\sqrt{2\mathcal{V}}} \left( \frac{\varepsilon_\mathbf{k} + g n_0}{\hbar\omega_\mathbf{k}} - 1 \right)^{1/2}, \tag{2.19b}$$

$$\hbar\omega_\mathbf{k} = \sqrt{\varepsilon_\mathbf{k} \left(\varepsilon_\mathbf{k} + 2 g n_0\right)}, \tag{2.19c}$$

where $\varepsilon_{\mathbf{k}} = \hbar^2 k^2 / 2m_b$ are the single-particle energies.

A similar analysis (dating back to 1959 [72]) may be performed in the canonical ensemble without violating number-conservation [66, 67, 68, 73]. The Hamiltonian is again found to take the similar form

$$\hat{H}_b = \sum_{\nu} \hbar\omega_\nu \hat{d}_\nu^\dagger \hat{d}_\nu \,, \tag{2.20}$$

up to some constant. The phonon spectrum $\hbar\omega_\nu$ is the same as found above, however, the exact form of the Bogoliubov creation $\hat{d}_\nu^\dagger$ and annihilation $\hat{d}_\nu$ operators are adjusted so that the Hamiltonian is number conserving.

Within both types of analysis, number conserving or non-conserving, a more complete discussion reveals that the assumption of small depletion $1 - N_0/N \ll 1$ is valid for small $na_s^3 \ll 1$ (cf. Eq. (2.11)) and low temperatures, making both analyses consistent for weakly-interacting bosons. See Refs. [66, 74] ([75]) for the number (non-)conserving case. Similarly, the differences between the number conserving and non-conserving predictions are small for weakly-interacting bosons $na_s^3 \ll 1$ and a large (average) number of bosons $N$.

## 2.2 Quasi low-dimensional Bose gases

In the previous section we found that when ideal bosons can move in all three spatial dimensions there is a macroscopic occupation of a single-particle eigenstate. We then used the macroscopic occupation of a single-particle mode as an assumption from which to describe the condensate wavefunction and non-condensate excitations for weakly-interacting bosons. For quasi low-dimensional systems, where motion is restricted to some subset of the spatial dimensions, this simple argument is less convincing since condensation is less prevalent, even for ideal bosons. Here we

describe a more careful analysis that justifies the use of the GP equations and Bogoliubov excitations for describing quasi low-dimensional weakly-interacting Bose gases. We follow the discussion presented in Ref. [76].

We begin by assuming a strong harmonic trapping along one (two) dimensions, with trapping frequency $\Omega_\perp$. By strong we mean that only zero-point motion occurs in this (these) dimension(s),

$$\hbar\Omega_\perp \gg k_B T, ng. \tag{2.21}$$

If this is the case then the Bose gas takes on a quasi two- (one-)dimensional character, i.e., the system is described by the Hamiltonian of Eq. (2.12) in a space of reduced spatial dimension with $g$ now an effective interaction strength. For the typical bosonic densities, scattering lengths and masses given above $ng \approx 10^{-32}$ J $\approx k_B \times$ 1 nK. If $k_B T$ is larger than this, say $k_b \times 100$ nK, then temperature is the limiting factor and, for this example temperature, Eq. (2.21) reduces to the condition

$$\Omega_\perp \gg 10^4 \text{ rad}\,\text{s}^{-1}. \tag{2.22}$$

If we further restrict the confinement to be weak enough that

$$l_\perp = \sqrt{\hbar/m_b\Omega_\perp} \gg a_s, \tag{2.23}$$

then the three-dimensional pseudo-potential approximation introduced in Sec. 2.1.2 remains valid. In this case, averaging over the Gaussian(s) describing the bosons in the strongly-confined dimension(s), the interaction strength takes the effective value $g = 2\sqrt{2}\hbar\Omega_\perp l_\perp a_s$ [77] ($g = 2\hbar\Omega_\perp a_s$ [78]) for a quasi two- (one-)dimensional system. For bosonic masses $m_b \approx 50$ amu, Eq. (2.23) is readily satisfied by setting $\Omega_b \ll 10^9$ rad s$^{-1}$ and there is no problem in simultaneously satisfying Eq. (2.22).

In preparation for the next subsections we make the further approximation that

the Bose gas is weakly-interacting in the quasi low-dimensional sense. We write the condition for this in terms of $n$, which from now on stands for the bosonic density averaged over the strongly-confined direction(s)[4]. The criterion for a weakly-interacting gas is that the kinetic energy $\hbar^2 n^{2/\mathtt{D}}/2m_b$ of a boson confined to a box of length equal to the expected inter-particle distance $n^{-1/\mathtt{D}}$, where $\mathtt{D}$ is the effective number of spatial dimensions, is much larger than the interaction energy per particle $ng$ [76]. For three dimensions $\mathtt{D} = 3$ this condition becomes Eq. (2.11), i.e., the weakly-interacting condition we previously assumed. For $\mathtt{D} = 2$ ($\mathtt{D} = 1$) it becomes $m_b g/\hbar^2 \ll 1$ ($m_b g/\hbar^2 n \ll 1$). Note that the importance of interactions actually decreases with increasing $n$ for quasi one-dimensional systems.

### 2.2.1 Uniform quasi two-dimensional Bose gas

We first describe the GP and Bogoliubov approaches for a uniform quasi two-dimensional weakly-interacting Bose gas in a volume $\mathcal{V}$ with periodic boundaries [79, 80]. At $T = 0$ things are similar to the three-dimensional case (cf. Sec. 2.1.2). Assuming a condensate[5] exists with small depletion $1 - N_0/N \ll 1$ then the condensate is described by the uniform solution to the GP equation and interaction-induced excitations above the condensate are represented by the Bogoliubov modes. The occupation of the Bogoliubov modes is small, which is consistent with the *a priori* assumption of a condensate [76].

However at finite temperatures $T > 0$, unlike in the three-dimensional case, assuming a condensate leads to a divergence in the occupation of the Bogoliubov modes — a sign that this assumption is false. The result is consistent with those of Mermin and Wagner [81], and Hohenberg [82], which prohibit the formation of a

[4]To be specific, the quasi two- (one-)dimensional density is obtained from the three-dimensional density via $n_{\mathtt{D}=2} = \pi l_\perp n_{\mathtt{D}=3}(r_\perp = 0)$ ($n_{\mathtt{D}=1} = \pi^2 l_\perp^2 n_{\mathtt{D}=3}(r_\perp = 0)$), where $n_{\mathtt{D}=3}(r_\perp = 0)$ is the three-dimensional density in the centre of the strongly-confining potential(s). In what follows we omit the dimension-specifying subscripts as this will be clear from the context.

[5]Once again we go against some conventions and refer to BEC at $T = 0$.

condensate for a uniform two-dimensional Bose gas at a finite temperature.

The cause of this lack of condensation is revealed by further investigation [83, 77], which we now outline in a number non-conserving approach. Firstly the field operator $\hat{\phi}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\hat{\theta}(\mathbf{r})}\sqrt{\hat{n}(\mathbf{r})}$ is written in terms of Hermitian density and phase operators satisfying $[\hat{n}(\mathbf{r}), \hat{\theta}(\mathbf{r}')] = \mathrm{i}\delta(\mathbf{r} - \mathbf{r}')$ [76]. The density is expanded as $\hat{n}(\mathbf{r}) = n_0(\mathbf{r}) + \delta\hat{n}(\mathbf{r})$. The vectors $\mathbf{r}$ and $\mathbf{r}'$ now denote positions in the space of reduced dimension. Secondly these expansions are inserted into the Hamiltonian of Eq. (2.12) and linearised with respect to $\delta\hat{n}(\mathbf{r})$ and $\nabla\hat{\theta}$, which are assumed to be small. Crucially this approximation is valid only for weakly-interacting bosons as defined in the previous subsection. Making this approximation we find that $n_0(\mathbf{r})$ is the density of the solution to the GP Eq. (2.14) and that the excitations of the Bose gas are described by the Bogoliubov modes, as before. The main difference to the three-dimensional case is that these excitations induce significant long-wavelength fluctuations of the phase. At low temperatures the phase fluctuations result in a power law decay of the single-particle density-matrix $\mathsf{g}(r) = n_0(\lambda_T/r)^{T/T_d}$ at large distances $r$, where $T_d = 2\pi\hbar^2 n_0/k_B m_b$ is the temperature associated with the onset of quantum degeneracy [84].

Note that an alternative definition of BEC is that $\mathsf{g}(r)$ takes on a non-zero value in the limit $r \to \infty$ [85]. This motivates the introduction of a phase coherence length $l_\phi = \lambda_T \mathrm{e}^{T_d/2T}$ defined as the distance at which the single-particle density-matrix decays by a factor $\mathrm{e}^{-1/2}$, i.e., $\mathsf{g}(l_\phi) = \mathsf{g}(0)\mathrm{e}^{-1/2}$. Over distances much shorter than $l_\phi$ the effects of the phase fluctuations are small enough that the decay of the single-particle density-matrix may be ignored and the Bose gas approximated by a condensate. Otherwise the phase fluctuations must be taken into account. Therefore the system may be divided into parts of size $l \ll l_\phi$ that each contain a phase-coherent condensate. If this is possible while also satisfying $l \gg \xi$, where $\xi = \hbar/\sqrt{2m_b ng}$ is the healing (correlation) length of the condensate, then we call

the state of the whole Bose gas a quasi condensate [79].

Substituting in realistic numbers that are consistent with the approximations given above, e.g., $m_b = 50$ amu, $n = 10^{15}$ m$^{-2}$ and $\Omega_\perp = 10^5$ rad s$^{-1}$, we obtain $\xi \approx 0.1$ $\mu$m, $\lambda_T \approx 0.2$ $\mu$m and $T_d \approx 60$ $\mu$K. For a reasonable temperature $T = 100$ nK this results in a coherence length $l_\phi \approx 10^{13}\xi$, which is well within the realm of a quasi condensate.

### 2.2.2 Uniform quasi one-dimensional Bose gas

A similar approach is applicable to the quasi one-dimensional weakly-interacting Bose gas [79, 80, 86]. In this case even at $T = 0$ there is no macroscopic occupation of a single mode, instead the single-particle density-matrix $\mathsf{g}(r)$ decays as a power law at large $r$. Phase fluctuations induced by interactions (not only those induced by temperature) therefore play a significant role in the uniform quasi one-dimensional system [87, 88]. At finite $T > 0$ the addition of thermal phase fluctuations cause the decay of the single-particle density-matrix $\mathsf{g}(r) = \exp(-m_b k_b T r/2n_0\hbar^2)$ to be exponential [83, 84].

As with quasi two-dimensional systems, the solution to the GP Eq. (2.14) still describes the zeroth order bosonic density and the Bogoliubov modes still describe the low energy excitations of the Bose gas. Also it is still possible that the phase correlation length $l_\phi = T_d/2\pi n_0 T$, with degeneracy temperature $T_d = 2\pi\hbar^2 n_0^2/k_B m_b$, is much longer than the healing length $\xi$, i.e., the Bose gas forms a quasi condensate [79, 86].

Again, substituting in realistic numbers demonstrates that quasi condensation is achievable. For $m_b = 50$ amu, $n_0 \approx n = 10^7$ m$^{-1}$ and $\Omega_\perp = 1 \times 10^5$ rad s$^{-1}$ we obtain $\xi \approx 0.5\mu$m and $T_d \approx 6\mu$K. For a low temperature $T = 10$nK this results in a correlation length $l_\phi \approx 21\xi$.

### 2.2.3 Trapped quasi low-dimensional Bose gases

Here we briefly summarise how the above results change if a harmonic trap of frequency $\Omega_b$ is applied along the previously unconfined directions (direction) of the quasi two- (one-)dimensional Bose gas. We write $v_b(\mathbf{r}) = \frac{1}{2}m_b\Omega_b^2 r^2$, where $\mathbf{r}$ denotes a position in the space of reduced dimension.

The trapping induces a finiteness to the size of the Bose gas, limiting its capacity to support the low wavelength excitations associated with phase fluctuations. The result is that, as in three dimensions, ordinary BEC[6] occurs for a large number of ideal non-interacting bosons, where the zero-momentum eigenstate is occupied by a macroscopic number $N_0$ of particles satisfying

$$\frac{N_0}{N} = 1 - \left[\frac{T}{T_c}\right]^2, \qquad \left(\frac{N_0}{N} = 1 - \frac{T}{T_c},\right) \tag{2.24}$$

above a critical temperature [89] ([90])

$$T_c = \frac{\hbar\Omega_b}{k_B}\left[\frac{N}{\zeta_2}\right]^{1/2}, \qquad \left(T_c = \frac{\hbar\Omega_b N}{k_B \ln N},\right) \tag{2.25}$$

where $\zeta_2 = \pi^2/6$. The width of the BEC transition is greater than for three dimensions, with occupation of the lowest energy eigenstate occurring over a temperature range

$$\Delta T \approx T_c\sqrt{\frac{\ln N}{N}}, \qquad \left(\Delta T \approx \frac{T_c}{\ln N}.\right) \tag{2.26}$$

Therefore only at very low temperatures $T \ll T_c$ is there a macroscopic occupation of a single-particle mode.

However, once again there is more to say. Including interactions, the bosons

[6]It is often said that BEC does not occur in a trapped one-dimensional ideal Bose gas, since the fraction of bosons in the single-particle ground state is zero in the continuum limit. However for a large finite number of bosons there is a macroscopic occupation of this mode and we here call this BEC. See Ref. [90] for more details.

may form a quasi condensate even above these temperatures. Let us describe this in the Thomas-Fermi regime [48]. This is the regime in which the trapping potential frequency $\Omega_b$ is so low that the density of the quasi low-dimensional Bose gas varies over length scales much larger than the healing length $\xi$. In this case the kinetic energy contribution to the GP Eq. (2.14) may be neglected, resulting in the condensate density

$$n_0(\mathbf{r}) = \frac{\mu_b - v_b(\mathbf{r})}{g}, \tag{2.27}$$

for $\mu_b > v_b(\mathbf{r})$ and zero otherwise. The size of the Bose gas is then given by the Thomas-Fermi radius $R = \sqrt{2\mu_b/m_b\Omega_b^2}$ satisfying $v_b(R) = \mu_b$ and the regime of validity for the Thomas-Fermi approximation may be expressed as $R \gg \xi$.

In this regime (see Refs. [76, 77, 86]) and for temperatures $T \ll T_d$, where $T_d$ is the degeneracy temperature of the trapped bosons, a quasi condensate exists with a coherence length $l_\phi \gg \xi$. The degeneracy temperature may be written $T_d = \pi\hbar^2\Omega_b^2R^2/k_Bg$ ($T_d = N\hbar\Omega_b/k_B$) for a two- (one-)dimensional Bose gas. Taking $N = 10^5$ ($N = 10^3$) particles and trapping frequency $\Omega_b =$ Hz we find $T_d = 100$ nK ($T_d = 10$ nK).

Note that quasi one-dimensional weakly-interacting Bose gases, quasi condensation and the associated phase fluctuations have all been observed experimentally [91, 92, 93, 94].

## 2.3 Bose gas in an optical lattice

We now return briefly to three dimensions and consider bosons in a periodic potential $v_b$. Such a potential is created using standing waves of laser light, slightly detuned from some resonance at wavelength $\lambda$, that exert a spatially-periodic AC Stark shift on the ground internal state of the bosons [95, 58]. Typically $\lambda \approx 1000$ nm. If a very intense pair of counter-propagating lasers are setup then the potential along

the direction of these lasers forms a sinusoid with very deep wells. If the potential at the minimum of each well corresponds to a harmonic trap with large frequency $\Omega_\perp \gg ng, k_BT$ then each well realises a quasi two-dimensional plane. Similarly if a near-identical[7] pair of lasers are setup in an orthogonal direction then the system divides into quasi one-dimensional tubes. In what follows next and in much of the later chapters we specialise to this case and consider a single quasi one-dimensional tube (which may or may not, e.g., Ref. [94], be created in the way described above). In particular, we here consider the situation in which an additional less intense pair of counter-propagating lasers is set up along the tube and establish a periodic potential $v_b(r) = v_{\text{depth}} \sin^2(2\pi r/\lambda)$ with lattice parameter $a = \lambda/2$, where $r$ is now the position within the quasi one-dimensional system.

### 2.3.1 Realising the Bose-Hubbard model

This system may be tackled using the GP and Bogoliubov treatments of the previous section [48]. However here we present an alternative description, first introduced by Jaksch *et al.* [25], that is applicable over a wider regime. Specifically it is still valid when the interaction energy of the bosons dominates the kinetic energy and large density fluctuations become important, i.e., when the bosons are no longer weakly-interacting[8].

The atoms are usually prepared such that they remain in a subspace described by a small number of the lowest energy Bloch bands (see e.g. Ref. [96] for an introduction to Bloch functions and bands). Here we consider the energy of the bosons, including the thermal $k_BT$ and interaction $ng$ energies, to be small compared to the

[7] We write near here to recognise that the two pairs of lasers must be sightly detuned from each other so that they do not interfere.

[8] Note that a strongly-interacting Bose lattice gas is compatible with the two-body contact potential approximation. It only when the density changes on a length scale much larger than the inter-particle distance that the condition for weak interactions coincides with the validity of the contact approximation. This is not the case for the system we consider in this section.

difference in energy between the lowest and next lowest bands, allowing us to focus entirely on the lowest energy band. We write the Bloch functions of this band as $\varphi_k(r)$, where the wave-vector $k$ takes values inside the first Brillouin zone of the reciprocal lattice $-\pi/a \leq k < \pi/a$.

To simplify the description further we express the state of the atoms in terms of a basis of highly-localised single-particle wavefunctions called Wannier functions [97]. These are given by the Fourier transform of the Bloch functions forming the lowest energy band

$$w_\ell(r) = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} \mathrm{d}k \; e^{-\mathrm{i}ka\ell} \varphi_k(r). \tag{2.28}$$

Here $\ell$ labels the primitive cell or lattice site in which the Wannier function $w_\ell$ is centred. It follows from Eq. (2.28) that the Wannier functions are identical up to translation

$$w_\ell(r) = w_{\ell'}(r + (\ell' - \ell)a). \tag{2.29}$$

These states therefore form a complete basis of orthogonal translationally symmetric functions in which we may expand the field operator

$$\hat{\phi}(r) = \sum_\ell w_\ell(r) \hat{b}_\ell. \tag{2.30}$$

Inserting this into the many-body Hamiltonian of Eq. (2.12) we obtain

$$\hat{H}_b = -\sum_{\ell\ell'} J_{\ell\ell'} \hat{b}_\ell^\dagger \hat{b}_{\ell'} + \frac{1}{2} \sum_{\ell\ell'\ell''\ell'''} U_{\ell\ell'\ell''\ell'''} \hat{b}_\ell^\dagger \hat{b}_{\ell'}^\dagger \hat{b}_{\ell''} \hat{b}_{\ell'''}, \tag{2.31}$$

with hopping and interaction parameters

$$J_{\ell\ell'} = -\int \mathrm{d}r \; w_\ell^*(r) h(r) w_{\ell'}(r), \tag{2.32a}$$

$$U_{\ell\ell'\ell''\ell'''} = g \int \mathrm{d}r \; w_\ell^*(r) w_{\ell'}^*(r) w_{\ell''}(r) w_{\ell'''}(r). \tag{2.32b}$$

Due to Eq. (2.29) these parameters are invariant under a simultaneous translation of the Wannier functions by a direct lattice vector.

The Wannier functions of Eq. (2.28) are not unique since there is total freedom in the choice of the phase of each Bloch function $\varphi_k$. The common approach that we follow here is to choose the phases such that $w_\ell$ are well-localised, and so all $J_{\ell\ell'}$ and $U_{\ell\ell'\ell''\ell'''}$ corresponding to hopping or interactions between distant lattice sites are negligible. This leaves the Bose-Hubbard model

$$\hat{H}_b = -\sum_{\langle \ell\ell' \rangle} J_{\ell\ell'} \hat{b}_\ell^\dagger \hat{b}_{\ell'} + \frac{1}{2} \sum_{\langle \ell\ell'\ell''\ell''' \rangle} U_{\ell\ell'\ell''\ell'''} \hat{b}_\ell^\dagger \hat{b}_{\ell'}^\dagger \hat{b}_{\ell''} \hat{b}_{\ell'''}, \tag{2.33}$$

where the angular brackets indicate that the sum is restricted to local terms, e.g., same-site, nearest-neighbour, or next-nearest-neighbour etc. The range of the terms that need to be kept will depend on how local the Wannier functions $w_\ell$ can be, which in turn is dependent on the depth $v_{\text{depth}}$ of the potential.

We therefore calculate the most localised Wannier functions, i.e., those that minimise the spread $(\Delta r)^2 = \int \mathrm{d}r w_\ell^*(r)(r-\bar{r}_\ell)^2 w_\ell(r)$, where $\bar{r}_\ell = \int \mathrm{d}r w_\ell^*(r) r w_\ell(r)$ is the Wannier centre [98]. Then from these maximally-localised Wannier functions we calculate the Hubbard parameters $J_{\ell\ell'}$ and $U_{\ell\ell'\ell''\ell'''}$. The approach we use is described in Ref. [8]. There we introduce an *ab initio* computational method for calculating the maximally-localised Wannier functions and corresponding Hubbard parameters for arbitrary periodic potentials, bands and dimensions (unfortunately, we do not say more about that work in this thesis).

Using this approach, in Fig. 2.1(a) we plot the maximally-localised Wannier function centred at the origin for a potential depth $v_{\text{depth}} = 3E_R$, where $E_R = h^2/2m_b\lambda^2$ is the recoil energy. The Wannier function decays exponentially, as is clear from the inset. For large $v_{\text{depth}}$ the Wannier functions become increasingly localised. The results of this are shown in Figs. 2.1(b) and (c), where we plot the dominant

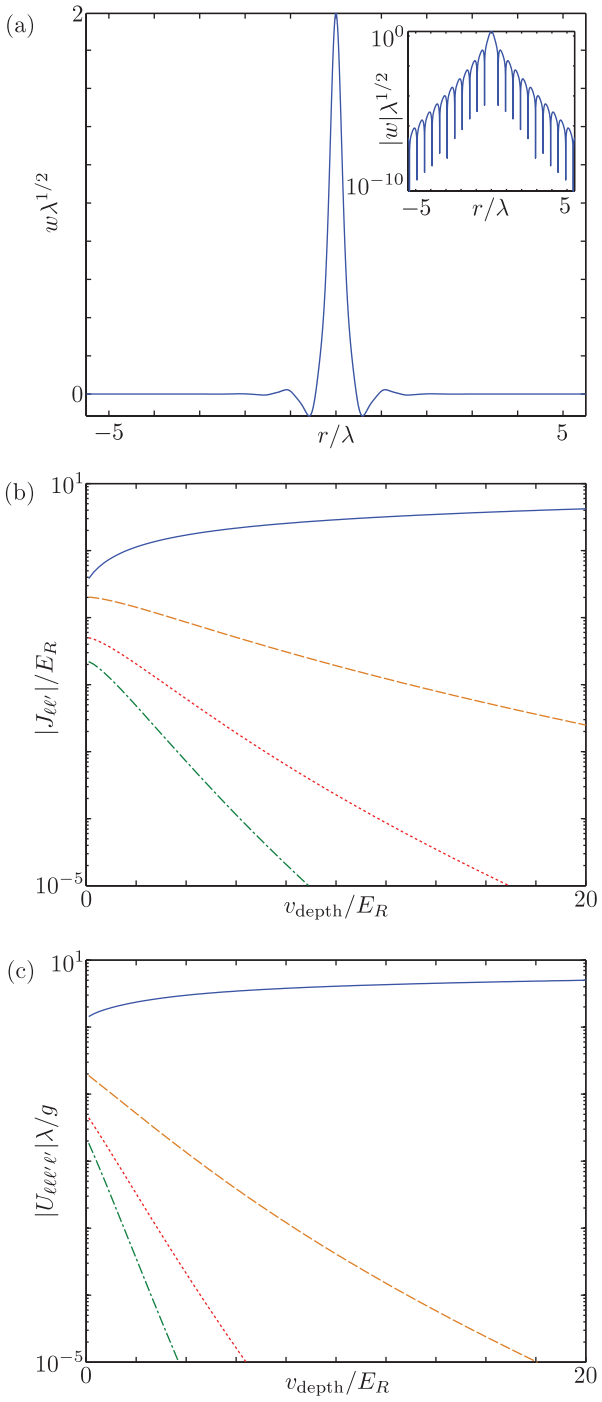


**Figure 2.1**. Hubbard model for bosons in a one-dimensional optical lattice. (a) The maximally-localised Wannier function for lattice depth $v_{\mathrm{depth}} = 3E_R$. The inset shows the absolute value of the same function with a logarithmic $y$-axis. (b) The magnitude of hopping and (c) interaction parameters corresponding to sites separated by $|\ell' - \ell| = 0$ (blue full line), $a$ (orange dashed line), $2a$ (red dotted line) and $3a$ (green dot-dashed line).

hopping and interaction parameters as a function of $v_{\text{depth}}$ (those parameters not shown are negligible). For all but the shallowest potentials $v_{\text{depth}} \lesssim 5E_R$, only the nearest-neighbour hopping and on-site interaction are significant (we may ignore the on-site hopping since it merely corresponds to the addition of a constant to the potential $v_b$). Therefore the Hamiltonian reduces to the single-band nearest-neighbour Bose-Hubbard model

$$\hat{H}_b = -J_b \sum_{\langle \ell\ell' \rangle} \hat{b}_\ell^\dagger \hat{b}_{\ell'} + \frac{U_b}{2} \sum_\ell \hat{b}_\ell^\dagger \hat{b}_\ell^\dagger \hat{b}_\ell \hat{b}_\ell, \tag{2.34}$$

where $J_b, U_b > 0$ and the angular brackets indicate that the sum is restricted to nearest-neighbour sites.

From the plots of $J_b$ in terms of $E_R$ and $U_b$ in terms of the energy $g/\lambda$ we see that to achieve the regime $U_b/J_b \lesssim 1$ we must have $g/\lambda \lesssim E_R$. For a typical wavelength $\lambda = 1$ $\mu$m, scattering length $a_s = 1$ nm and light atoms $m_b = 5$ amu, the regime $g/\lambda \approx E_R$ is obtained by choosing a small radial trapping frequency $\Omega_\perp = 10^2$ rad s$^{-1}$. If the bosons are to retain a quasi one-dimensional nature it is then required that the temperature and maximum three-dimensional particle density are less than 1 nK and $10^{18}$ m$^{-3}$, respectively. Despite these parameters being challenging to obtain the one-dimensional characteristics in the excitation spectrum of Rb in an optical lattice with $U_b/J_b \approx 1$ have been observed experimentally [99]. Alternatively, the necessity of such low temperatures may be avoided by exploiting a Feshbach resonance [100, 101, 102, 103] (which, of course, brings its own experimental challenges) to decrease $g$ without having to make $\Omega_\perp$ small.

Finally, let us comment on the wider significance of the interpretation of cold atoms and optical lattices in terms of Hubbard models. First it allows us, at least in one dimension, to apply sophisticated computational methods (cf. Ch. 3) to simulate the full dynamics of the system for arbitrarily large $v_{\text{depth}}$. Second it provides an

intuitive picture of bosons hopping between spatially-separated sites in a lattice, the occupation of which can be measured through high-resolution imaging [104]. Last, local Hubbard models, which are often not analytically or numerically tractable, are important in understanding solid state phenomena (e.g. the two-dimensional fermionic Hubbard model has been proposed as a model of high-temperature superconductivity [105]). So studying them in a highly controllable and clean environment using cold atoms in optical lattices adds to our understanding of condensed matter systems [95].

### 2.3.2 Properties of the Bose-Hubbard model

We now briefly review some properties of the Bose-Hubbard model [26] that we will use later in Ch. 5. We begin by considering a small $v_{\mathrm{depth}}$ and the weakly-interacting limit $U_b n_b/J_b \ll 1$, where $n_b = N/M$ and $M$ is the number of lattice sites. We assume periodic boundary conditions. The analysis is very similar to what was described earlier for a weakly-interacting gas in a uniform potential. Therefore here we merely state the common number non-conserving approach and note that a number conserving analogue exists as before.

This approach starts by assuming *a priori* a macroscopic occupation of the lowest single-particle eigenvector at $U_b/J_b = 0$. The single-particle eigenvectors and eigenenergies are just those of the tight binding model [96]. The eigenvectors have elements $(\varphi_k)_\ell = \mathrm{e}^{\mathrm{i}ka\ell}/\sqrt{M}$, where $\hbar k$ is the crystal momentum, and the eigenspectrum is $\varepsilon_k = 4J_b \sin^2(ka/2)$. Therefore it is the vector with zero crystal momentum that is macroscopically occupied. We then replace the creation $\hat{a}_0^\dagger$ and annihilation $\hat{a}_0$ operators for this eigenvector with the $c$-number $\sqrt{N_0}$, where $N_0 \approx N$, and expand the Hamiltonian to second order in the creation and annihilation operators corresponding to vectors with non-zero crystal momenta. Applying a Bogoliubov

transformation results in [106, 107, 108]

$$\hat{H}_b - \mu_b \hat{N} = \sum_k \hbar\omega_k \hat{d}_k^\dagger \hat{d}_k \,, \tag{2.35}$$

up to the addition of a constant. Here the Bogoliubov phonon creation $\hat{d}_k^\dagger$ and annihilation $\hat{d}_k$ operators are superpositions of the creation and annihilation operators for the single-particle eigenvectors with wave-vector $k$ and $-k$. The phonon spectrum is given by $\hbar\omega_k = \sqrt{\varepsilon_k(\varepsilon_k + 2U_b n_{b,0})}$, where $n_{b,0} = N_0/M \approx n_b$.

For weak interactions $U_b n_b \ll J_b$ the *a priori* assumption of the macroscopic occupation of a single-particle mode made above may be justified as for the continuous quasi one-dimensional Bose gas. Phase fluctuations prohibit BEC but the length scale over which the phase fluctuates can be made large compared to the correlation length of the system. So long as we are concerned with length scales for which the phase is roughly constant the approach outlined above is valid.

For strong interactions, i.e., large $U_b n_b \gtrsim J_b$, the Bogoliubov approach breaks down. This is particularly apparent in the limit $U_b/J_b \to \infty$ in which the bosons map to the same number of non-interacting identical fermions with energy spectrum $\varepsilon_k$ [109].

For the commensurately filled case of integer $n_b$ an increasing interaction strength $U_b/J_b$ takes the gas through a continuous superfluid to Mott insulator transition, which for $n_b = 1$ occurs at $U_b/J_b \approx 3.37$ [110, 111] in the thermodynamic limit. The ground state deep in this Mott insulator regime ($U_b/J_b \gg 1$) is the state in which each lattice site contains exactly one boson.

## 2.4 Chapter summary

In this chapter we have introduced several common descriptions of cold bosonic atoms in anisotropic traps for which the system takes on a low-dimensional character. We have discussed their regimes of validity and given examples of experimental parameters for which these models may be realised.

Following this chapter, whenever we discuss one or two-dimensional Bose gases we are of course discussing quasi two- (one-)dimensional systems. Similarly, whenever we describe a condensate we include the possibility of a quasi condensate whose phase fluctuations may be neglected over the distances in which we are interested.

# Chapter 3

# Numerical techniques

Having introduced Bose gases and several models that describe them, we now present some numerical techniques for analysing these models.

The methods we present first, in Sec. 3.1, are for solving (possibly coupled) time-independent and time-dependent non-linear Schrödinger equations, such as the Gross-Pitaevskii (GP) Eqs. (2.14) and (2.15). Then in Sec. 3.2 we introduce methods for describing the ground state and dynamics of one-dimensional quantum lattice systems, such as the Bose-Hubbard model of Eq. (2.35).

Each of these methods will be naturally extended in the next two chapters (Chs. 4 and 5) to simulate an impurity interacting with a Bose gas. The lattice-based methods will also be extremely important for the second part of this thesis, where we use them to describe the dynamics of classical stochastic lattice systems.

## 3.1 GP solvers

Here we introduce the methods that we will use in Ch. 4 to solve coupled time-independent and time-dependent GP equations. We explain the methods by applying them to the single-species time-independent and time-dependent GP Eqs. (2.14) and (2.15) in a single dimension. The approach is readily extended to higher

dimensions and coupled equations. However, we do not go into the details of how they are extended in this thesis and instead point the reader to Refs. [112, 113], which discuss more complicated uses of the same methods.

### 3.1.1 Continuous normalised gradient flow

We begin by describing how to find the ground state solution to the one-dimensional time-independent GP Eq. (2.14), which we restate here for clarity as

$$\mu_b\varphi(r) = \left(-\frac{\hbar^2}{2m_b}\frac{\partial^2}{\partial r^2} + v_b(r) + g|\varphi(r)|^2\right)\varphi(r), \tag{3.1}$$

where $\varphi(r)$ is normalised to the number of particles in the condensate

$$||\varphi||_2 = \int \mathrm{d}r|\varphi(r)|^2 = N_0. \tag{3.2}$$

A function $\varphi(r)$ solves Eq. (3.1) iff it extremises the energy functional [48]

$$E[\varphi] = \int \mathrm{d}r\varphi^*(r)\left(-\frac{\hbar^2}{2m_b}\frac{\partial^2}{\partial r^2} + v_b(r) + \frac{g}{2}|\varphi(r)|^2\right)\varphi(r). \tag{3.3}$$

The corresponding value of $\mu_b$ is given by

$$N_0\mu_b = E[\varphi] + \frac{g}{2}\int \mathrm{d}r|\varphi(r)|^4. \tag{3.4}$$

In particular, the ground state minimises both the energy functional in Eq. (3.4) and $\mu_b$.

The above leads to a simple iterative steepest-descent method for calculating the ground state wavefunction $\varphi(r)$ [114, 115], termed normalised gradient flow, which

we now outline. First, take some wavefunction $\varphi(r)$ and calculate

$$\frac{\delta E[\varphi]}{\delta \varphi(r)} = \left( -\frac{\hbar^2}{2m_b}\frac{\partial^2}{\partial r^2} + v_b(r) + \frac{g}{2}|\varphi(r)|^2 \right) \varphi(r). \tag{3.5}$$

Then make the replacement $\varphi(r) \mapsto \varphi(r) - \frac{\delta E[\varphi]}{\delta \varphi(r)}$. Last, renormalise $\varphi(r)$ according to Eq. (3.2). When repeating these steps iteratively the true ground state wavefunction is a valid convergence point. However there is no guarantee that an iteration step does not increase $E[\varphi]$ and the ground state may not be the only point of convergence.

This problem of non-monotonically decreasing energy is due to the handling of the normalisation of the wavefunction and may be overcome by continuously renormalising the gradient [116, 117, 118]. This is called continuous normalised gradient flow and is equivalent to evolving $\varphi(r,t)$, now a function of time $t$, according to the differential equation

$$\frac{\partial \varphi(r,t)}{\partial t} = -\frac{1}{2}\frac{\delta E[\varphi]}{\delta \varphi(r,t)} + \frac{E[\varphi]}{||\varphi(t)||_2}\varphi(r,t). \tag{3.6}$$

In this case $E[\varphi]$ decreases monotonically in time $t$. Equation (3.6) can be integrated numerically using normalisation preserving splitting methods, such as those we present in the next subsection.

### 3.1.2 Time-splitting

We now turn to solving the time-dependent GP Eq. (2.15) in one dimension

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varphi(r,t) = \left( -\frac{\hbar^2}{2m_b}\frac{\partial^2}{\partial r^2} + v_b(r) + g|\varphi(r,t)|^2 \right) \varphi(r,t). \tag{3.7}$$

The approach we take utilises splitting, which we now describe. Consider the differential equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varphi(r,t) = (\mathcal{D}_1 + \mathcal{D}_2)\,\varphi(r,t), \tag{3.8}$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ are some differential (with respect to $r$) operators. Splitting methods break down the solving of this possibly complicated differential Eq. (3.8) into the solving of the differential equations

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varphi_1(r,t) = \mathcal{D}_1\varphi_1(r,t), \tag{3.9a}$$

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\varphi_2(r,t) = \mathcal{D}_2\varphi_2(r,t), \tag{3.9b}$$

which may be easier.

Let us be more explicit. Consider dividing the evolution according to Eq. (3.8) into $n_{\mathrm{steps}}$ time-steps of duration $\delta t = t_f/n_{\mathrm{steps}}$. Here $t_f$ is the duration of the simulation, which starts at time $t = 0$. Given $\varphi(r,t_j)$, the solution to Eq. (3.8) at the time $t_j = j\delta t$, the task is to construct $\varphi(r,t_{j+1})$.

The most simple splitting method, named after Marchuk and Yanenko [119, 120], is to first find $\varphi_1(r,t_{j+1})$, the solution to Eq. (3.9a) at the time $t_{j+1}$ given the initial condition $\varphi_1(r,t_j) = \varphi(r,t_j)$. Then find $\varphi_2(r,t_{j+1})$, the solution to Eq. (3.9b) at the time $t_{j+1}$ given the initial condition $\varphi_2(r,t_j) = \varphi_1(r,t_{j+1})$. The final solution $\varphi_2(r,t_{j+1})$ is an approximation of the desired solution $\varphi(r,t_{j+1})$ with the error being $\mathcal{O}(\delta t^2)$. We may express this in the following way

$$\begin{aligned}\varphi(t_{j+1}) &= \mathrm{e}^{-\mathrm{i}(\mathcal{D}_1+\mathcal{D}_2)\delta t/\hbar}\varphi(t_j) \\ &= \mathrm{e}^{-\mathrm{i}\mathcal{D}_1\delta t/\hbar}\mathrm{e}^{-\mathrm{i}\mathcal{D}_2\delta t/\hbar}\varphi(t_j) + \mathcal{O}(\delta t^2).\end{aligned} \tag{3.10}$$

The scaling of this error is improved using so-called Strang splitting (named after Ref. [121]), which is a second order method and the one used in this thesis. The

Strang splitting method is summarised by the equation

$$\varphi(t_{j+1}) = \mathrm{e}^{-\mathrm{i}\mathcal{D}_1\delta t/2\hbar}\mathrm{e}^{-\mathrm{i}\mathcal{D}_2\delta t/\hbar}\mathrm{e}^{-\mathrm{i}\mathcal{D}_1\delta t/2\hbar}\varphi(t_j) + \mathcal{O}(\delta t^3). \tag{3.11}$$

Strang splitting is unconditionally stable if the discrete counterparts of $\mathcal{D}_1$ and $\mathcal{D}_2$ are positive definite matrices, which will be the case in what follows.

Returning now to the concrete example of the time-dependent GP Eq. (2.15) and following Refs. [122, 117] we make the division

$$\mathcal{D}_1\varphi_1(r,t) = \left[v_b(r) + g|\varphi_1(r,t)|^2\right]\varphi_1(r,t), \tag{3.12a}$$

$$\mathcal{D}_2 = -\frac{\hbar^2}{2m_b}\frac{\partial^2}{\partial r^2}. \tag{3.12b}$$

The first differential operator $\mathcal{D}_1$ is diagonal in the position basis and so its only effect is to vary the phases at different positions, i.e., it does not affect the norm. Therefore a solution to Eq. (3.9a) exactly satisfies

$$\varphi_1(r,t_{j+1/2}) = \exp\left[-\frac{\mathrm{i}\delta t}{2\hbar}\left(v_b(r) + g|\varphi_1(r,t_j)|^2\right)\right]\varphi_1(r,t_j), \tag{3.13a}$$

$$\varphi_1(r,t_{j+1}) = \exp\left[-\frac{\mathrm{i}\delta t}{2\hbar}\left(v_b(r) + g|\varphi_1(r,t_{j+1/2})|^2\right)\right]\varphi_1(r,t_{j+1/2}). \tag{3.13b}$$

The second differential operator $\mathcal{D}_2$ is diagonal in the momentum basis. To solve Eq. (3.9b) we therefore rotate into the momentum basis, perform the evolution, which in this basis once again results only in the variation of phases for different momenta, and then rotate back into the position basis.

The exact process depends on our system, our boundary conditions and how we represent the solutions. In this thesis we assume a finite system $0 \leq r < (M-1)a$ with box boundary conditions, i.e., $\varphi(0,t) = \varphi((M-1)a,t) = 0$ (valid for a sufficiently trapped Bose gas). We represent the condensate wavefunction by the

values $\varphi(r_\ell, t)$ it takes at the points $r_\ell = (\ell - 1)a$, which form a linear grid of $M$ points with spacing $a$. The appropriate transformation into the momentum basis is therefore the sine transformation

$$\varphi_2(k,t) = \sqrt{\frac{2}{M}} \sum_{\ell=1}^{M} \varphi_2(r_\ell, t) \sin(k r_\ell). \tag{3.14}$$

In this basis the differential operator $\mathcal{D}_2$ evolves momentum basis vectors as

$$\varphi_2(k, t_{j+1}) = \exp\left(-\frac{\mathrm{i}\hbar k^2 \delta t}{2m_b}\right) \varphi_2(k, t_j), \tag{3.15}$$

Bringing Eqs. (3.14), its inverse and (3.15) together, Eq. (3.9b) has the sine-spectral solution

$$\varphi_2(r, t_{j+1}) \approx \frac{2}{M} \sum_{\ell=1}^{M} \sum_{\ell'=1}^{M} \exp\left(-\frac{\mathrm{i}\hbar k^2 \delta t}{2m_b}\right) \sin(k r_\ell) \sin\left(\frac{2\pi(\ell'-1)r}{M^2 a}\right) \varphi_2(r_\ell, t_j). \tag{3.16}$$

The error of this does not depend on the time-step $\delta t$ and is accurate to spectral (infinite) order in the spatial discretisation $a$ [122].

Finally, combining all the steps of Eqs. (3.13) and (3.16) into the Strang splitting approach of Eq. (3.11) results in the desired single time-step evolution. Repeating this for each time-step of a larger evolution is called the second order time-splitting sine-spectral method. It follows from the details above that the method is unconditionally stable, normalisation preserving, and is accurate to spectral order in the spatial resolution and second order in the time-step [123].

## 3.2 Matrix product state methods

Next we introduce a rather different type of numerical method[1]. Instead of continuous wavefunctions, the aim here is to simulate the unitary evolution of a discrete quantum state, specifically the state of a quantum system composed of many small subsystems, which we refer to as lattice sites. This for example includes systems described by Hubbard models (see Sec. 2.3.1), which will be our focus in Ch. 5.

The problem faced in such simulations is that the memory required to store the state of the full system scales exponentially with the number of lattice sites, as does the processing time needed to evolve this state. Here we focus on a well-established approach for overcoming these problems, the time-evolving block decimation (TEBD) algorithm [11, 12], based on matrix product states (MPS). TEBD has been used to compute the dynamics of spin chains [12, 124, 125] and Hubbard models [126, 127] (see Sec. 2.3.1). The algorithm provides access to various dynamical properties like atomic many-body currents [128], non-linear responses to driving fields [129] and the formation and transport of quasi-particles [24, 130].

Our goal of simulating the unitary evolution of a system may be divided into three parts. We wish to first represent the state of the system, then evolve it according to some Hamiltonian and finally extract observable expected values from this evolved state. In this section we will show how each part may be performed efficiently[2]. First in Sec. 3.2.1 we describe how an MPS is used to approximately and efficiently represent a state vector. Second, when this state vector is evolved the MPS representation must be updated. The strength of the MPS structure is that for local Hamiltonians the MPS can be efficiently updated to approximately reflect the evolution of the state vector it represents, as we will show in Sec. 3.2.2;

[1] This section is largely based on a review we first presented in Ref. [1].

[2] Here we take efficient to mean the computational resources needed scale as a low order polynomial in the number of lattice sites.

this is the TEBD algorithm. Last, in Sec. 3.2.3 we demonstrate how, without any further approximation, expected values of local operators may be efficiently calculated from the MPS representing the state vector. We leave some of the more technical derivations to the appendices.

### 3.2.1 Matrix product states

To illustrate MPS we consider a quantum system composed of $M$ sites, each with a local $d$-dimensional Hilbert space. An arbitrary quantum state of this system $|\,\psi\rangle$, normalised according to the $L_2$-norm as $\| \,|\,\psi\rangle\,\|_2 \;=\; 1$, can be expressed as

$$|\,\psi\rangle = \sum_{\mathbf{i}} \psi_{\mathbf{i}}\,|\,\mathbf{i}\rangle\,. \tag{3.17}$$

Here $\mathbf{i} = (i_1, i_2, \cdots, i_M)$ is an $M$-tuple of local indices $i_\ell = 0, \cdots, d-1$ specifying a complete configuration and $|\,\mathbf{i}\rangle = |\,i_1\rangle\,|\,i_2\rangle \cdots |\,i_M\rangle$ are orthonormal configuration states. Since $|\,\psi\rangle$ inhabits the tensor product vector space $(\mathbb{C}^d)^{\otimes M}$, whose dimension grows exponentially with $M$ as $d^M$, an exact description of its amplitudes $\psi_{\mathbf{i}}$ rapidly becomes intractable with increasing $M$. To attempt to overcome this scaling, we expand $\psi_{\mathbf{i}}$ as

$$\psi_{\mathbf{i}} = A^{[1]i_1} A^{[2]i_2} \cdots A^{[M]i_M}, \tag{3.18}$$

where for each $i_\ell$ the factor $A^{[\ell]i_\ell}$ is a complex matrix of size $\chi_{\ell-1} \times \chi_\ell$, aside from the boundary terms $A^{[1]i_1}$ and $A^{[M]i_M}$, which are instead $1 \times \chi_1$ and $\chi_{M-1} \times 1$ complex vectors, respectively[3]. The expansion in Eq. (3.18) is simple to visualise; each site $\ell$, depending on its local configuration $i_\ell$, contributes a matrix $A^{[\ell]i_\ell}$ to a site-ordered multiplication of matrices that results in the amplitude $\psi_{\mathbf{i}}$. The boundary vectors

[3]It follows from Eq. (3.18) that an MPS is invariant under the transformation $A^{[\ell]i_\ell} \mapsto A^{[\ell]i_\ell} Y$, $A^{[\ell+1]i_{\ell+1}} \mapsto Y^{-1} A^{[\ell+1]i_{\ell+1}}$, where $Y$ is a non-singular matrix. The formalism of Sec. 3.2.1 and App. 3.A automatically exploits this freedom so that the tensors $A^{[\ell]}$ satisfy orthonormality conditions.

ensure that a scalar is recovered. Such an MPS can be thought of as a tensor network, as depicted in Fig. 3.1.

A product state, whose MPS requires dimensions $\chi_\ell = 1$ for all $\ell$, has no correlations between different sites while the MPS describing a vector with arbitrarily strong correlations may need matrix dimensions $\chi_\ell$ up to $X_\ell = \min\left(d^\ell, d^{M-\ell}\right)$ and so scale exponentially with the system size $M$. Thus in describing arbitrary vectors an MPS description offers no advantage.

The utility of an MPS formulation, however, is based precisely on the fact that physical states have a structure which allows for a near-exact MPS description while restricting $\chi_\ell \leq \chi$ with $\chi$ much smaller than the largest $X_\ell$. For quantum systems, correlations in a many-body state $|\,\psi\rangle$ are quantified by the entanglement entropy and this in turn has a rigorous connection to the required MPS dimension $\chi$ [11, 131]. This is shown by first splitting the system into two contiguous blocks $\mathtt{L}^{[\ell]} = \{1, \cdots, \ell\}$ and $\mathtt{R}^{[\ell]} = \{\ell+1, \cdots, M\}$. The state is then expanded as $|\,\psi\rangle = \sum_{\mathbf{ij}} \psi_{\mathbf{ij}}\, |\,\mathbf{i}\rangle_{\mathtt{L}}\, |\,\mathbf{j}\rangle_{\mathtt{R}}$, where $\mathbf{i}$ and $\mathbf{j}$ label configurations of the two subsystems and $\psi$ is now the reshaped $d^\ell \times d^{M-\ell}$ matrix of amplitudes. To expose the correlations of this state a singular value decomposition (SVD) [132] is performed that factorises $\psi = UDV^\dagger$, where $U$ and $V$ are unitary matrices, and $D$ is a diagonal matrix with real non-negative elements. This establishes a so-called Schmidt decomposition [133] of $|\,\psi\rangle$ as

$$|\,\psi\rangle = \sum_{\mu=1}^{X_\ell} \lambda_\mu^{[\ell]} |\,\mathtt{L}_\mu^{[\ell]}\rangle |\,\mathtt{R}_\mu^{[\ell]}\rangle. \tag{3.19}$$

Here $|\,\mathtt{L}_\mu^{[\ell]}\rangle$ and $|\,\mathtt{R}_\mu^{[\ell]}\rangle$ are Schmidt states derived from the columns of $U$ and $V$, while $\lambda_\mu^{[\ell]}$ are the singular values (Schmidt coefficients in this context) specified by the diagonal elements of $D$, arranged in non-increasing order with $\mu$. The normalisation

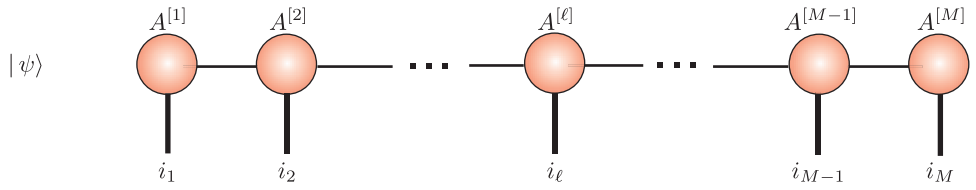


**Figure 3.1.** An MPS for a vector $|\psi\rangle$ as a comb-like tensor network. Each site $\ell$ of the system has a tensor $A^{[\ell]}$ associated with it, illustrated by the shaded circles. The thick vertical leg of a tensor $A^{[\ell]}$ represents the physical index $i_\ell$, while the thinner leg(s) portray the internal index (indices). The boundary tensors have a single internal leg and the others have two. The joining of tensor legs represents the contraction (or matrix multiplication) over the internal indices.

of $|\psi\rangle$, and unitarity of $U$ and $V$ result in

$$\sum_\mu (\lambda_\mu^{[\ell]})^2 = 1, \tag{3.20a}$$

$$\langle \mathtt{L}_\nu^{[\ell]} | \mathtt{L}_\mu^{[\ell]} \rangle = \langle \mathtt{R}_\nu^{[\ell]} | \mathtt{R}_\mu^{[\ell]} \rangle = \delta_{\nu\mu}. \tag{3.20b}$$

The mutual information between the subsystems $\mathtt{L}^{[\ell]}$ and $\mathtt{R}^{[\ell]}$ is equal to twice the entanglement entropy [134], given by

$$S_{\mathtt{LR}} = -\sum_{\mu=1}^{X_\ell} (\lambda_\mu^{[\ell]})^2 \log\left[(\lambda_\mu^{[\ell]})^2\right]. \tag{3.21}$$

The size of the entanglement entropy between a subsystem and the rest of the system for ground states has been extensively studied in the context of area laws (see Ref. [135] for a review). These have shown that $S_{\mathtt{LR}}$ should scale with the boundary connecting the two parts up to logarithmic corrections at criticality, and area laws therefore severely limit the entanglement entropy physical systems can possess. For quantum spin chains with short-range interactions the area law is manifested in an extremely rapid decay of the Schmidt coefficients $\lambda_\mu^{[\ell]}$ [12]. An immediate consequence of this is that an accurate (un-normalised) approximation

$|\,\tilde{\psi}^{[\ell]}\rangle$ of $|\,\psi\rangle$ can be formed by truncating the sum in Eq. (3.19) to the first $\chi$ terms [12] as

$$|\,\tilde{\psi}^{[\ell]}\rangle = \sum_{\mu=1}^{\chi} \lambda_{\mu}^{[\ell]} |\,\mathtt{L}_{\mu}^{[\ell]}\rangle |\,\mathtt{R}_{\mu}^{[\ell]}\rangle. \tag{3.22}$$

The error made in computing the expected value of any quantum mechanical observable with $|\,\tilde{\psi}^{[\ell]}\rangle$ rather than $|\,\psi\rangle$ is bounded by the $L_2$-norm of the residual [133], given by $||\,|\,\psi\rangle - |\,\tilde{\psi}^{[\ell]}\rangle||_2 = [\sum_{\mu=\chi+1}^{X_\ell} (\lambda_{\mu}^{[\ell]})^2]^{1/2}$, and is therefore small if the sum of the squares of the truncated Schmidt coefficients is small [12]. These observations can be directly related to an MPS, as described in App. 3.A, by applying this decomposition and truncation in an iterative sequence to all contiguous bipartitions to form an approximation $|\,\tilde{\psi}\rangle$ [11]. This approximation is an MPS of the form of Eq. (3.18) but with a maximum matrix dimension $\chi$. The errors in this case are bounded by the singular values [131] as

$$||\,|\,\psi\rangle - |\,\tilde{\psi}\rangle||_2 \leq \left[ 2 \sum_{\ell=1}^{M-1} \sum_{\mu=\chi+1}^{X_\ell} (\lambda_{\mu}^{[\ell]})^2 \right]^{1/2}. \tag{3.23}$$

A rapid decay of $\lambda_{\mu}^{[\ell]}$ with $\mu$ for each bipartition thus enables $\chi$ to be kept small while retaining a high accuracy, and results in an MPS description parameterised by $\mathcal{O}(Md\chi^2)$ complex numbers [12]. This near lossless compression of the information in $|\,\psi\rangle$ for physically relevant vectors is responsible for the enormous success of the MPS approach used within density-matrix renormalisation group [27, 28] and TEBD methods [11, 12].

In addition to the physical properties above, the SVD also has a pleasing mathematical feature that makes it especially useful when describing quantum systems: given $|\,\psi\rangle = \sum_{\mathbf{ij}} \psi_{\mathbf{ij}} |\,\mathbf{i}\rangle_{\mathtt{L}} |\,\mathbf{j}\rangle_{\mathtt{R}}$ and $|\,\phi\rangle = \sum_{\mathbf{ij}} \phi_{\mathbf{ij}} |\,\mathbf{i}\rangle_{\mathtt{L}} |\,\mathbf{j}\rangle_{\mathtt{R}}$, a well-known result first proposed by Eckart and Young [136] is that $|\,\tilde{\psi}^{[\ell]}\rangle$ (cf. Eq. (3.22)) is the solution to

the optimisation problem [132]

$$\min_{\substack{\phi\in\mathbb{C}^{d^\ell\times d^{M-\ell}},\\ \mathrm{rank}(\phi)\leq\chi}} || \,|\,\psi\rangle - |\,\phi\rangle\,||_2 = || \,|\,\psi\rangle - |\,\tilde{\psi}^{[\ell]}\rangle||_2, \tag{3.24}$$

for any $1 \leq \chi \leq X_\ell$. Thus the $L_2$-norm error of a truncated SVD will always be less than that of any other factorisation of the same rank. This ensures that if there exists an exact MPS of some dimension $\chi$ then there also exists an exact SVD-generated MPS of dimension $\chi$ or less.

### 3.2.2 Time-evolving block decimation

We have seen that MPS can provide an accurate and efficient description of some state vectors. Our goal in this subsection is to review the TEBD algorithm, devised to update a quantum MPS under unitary time-evolution [11, 12, 127]. The unitary evolution of a state between times $t = 0$ and $t = t_f$ generated by Hamiltonian $\hat{H}$ results in $|\,\psi(t_f)\rangle = \mathcal{T}\exp[-(\mathrm{i}/\hbar)\int_0^{t_f}\mathrm{d}t\hat{H}(t)]\,|\,\psi(0)\rangle$, where $\mathcal{T}$ is the time-ordering operator.

Another evolution we will be interested in is imaginary time-evolution for a time-independent Hamiltonian $\hat{H}$, which gives $|\,\psi(t_f)\rangle = \exp(-\hat{H}t_f/\hbar)\,|\,\psi(0)\rangle$. For long times $t_f \to \infty$ the final state $|\,\psi(t_f)\rangle$ becomes the ground state of $\hat{H}$ after renormalisation. In this section we will present TEBD as used for unitary evolution governed by a (potentially) time-dependent Hamiltonian, remembering that the algorithm for imaginary time-evolution is obtained by removing the time-dependence of the Hamiltonian and replacing $t$ with $-\mathrm{i}t$.

### Time-steps

For anything but the smallest systems neither the initial state $|\,\psi(0)\rangle$ nor the full unitary evolution operator $\mathcal{T}\exp[-(\mathrm{i}/\hbar)\int_0^{t_f}\mathrm{d}t\hat{H}(t)]$ can, in general, be represented exactly. Having applied an MPS to tackle the description of $|\,\psi(0)\rangle$ we now follow the standard approach [12], namely breaking the evolution up into $n_{\mathrm{steps}}$ small time-steps $\delta t = t_f/n_{\mathrm{steps}}$ as

$$\mathcal{T}\exp\left(-\frac{\mathrm{i}}{\hbar}\int_0^{t_f}\mathrm{d}t\hat{H}(t)\right) = \prod_{j=1}^{n_{\mathrm{steps}}}\exp\left(-\frac{\mathrm{i}\hat{H}(t_j)\delta t}{\hbar}\right) + \mathcal{O}\left(\delta t^2\right), \tag{3.25}$$

where $t_j = j\delta t$.

### Trotter expansion

Next we split the evolution operator $\exp(-\mathrm{i}\hat{H}(t_j)\delta t/\hbar)$ of the $j$-th time-step into a product of two-site operators (from here-on, for clarity, we omit the labelling of $j$ in the time $t_j$ of the time-step). To split the evolution operator we first restrict ourselves to considering Hamiltonians $\hat{H}$ composed of a sum of nearest-neighbour two-site terms $\hat{h}_{\ell,\ell+1}$ as

$$\hat{H} = \sum_{\ell=1}^{M-1}\hat{h}_{\ell,\ell+1}, \tag{3.26}$$

where we have incorporated any single-site terms into $\hat{h}_{\ell,\ell+1}$. We can then proceed to split up $\exp(-\mathrm{i}\hat{H}\delta t/\hbar)$ via a second order Suzuki-Trotter expansion [137] as

$$\exp\left(-\frac{\mathrm{i}\hat{H}\delta t}{\hbar}\right) = \left(\prod_{\ell=1}^{M-1}\mathrm{e}^{-\mathrm{i}\hat{h}_{\ell,\ell+1}\delta t/2\hbar}\right)\left(\prod_{\ell=M-1}^{1}\mathrm{e}^{-\mathrm{i}\hat{h}_{\ell,\ell+1}\delta t/2\hbar}\right) + \mathcal{O}\left(\delta t^3\right), \tag{3.27}$$

which represents an approximation because overlapping terms $\hat{h}_{\ell,\ell+1}$ and $\hat{h}_{\ell+1,\ell+2}$ do not generally commute[4]. This expansion therefore applies a sequence of two-

[4]This is reminiscent of the Strang time-splitting method discussed in Sec. 3.1.2.

site nearest-neighbour unitary operators $\hat{S}_\ell = \exp(-\mathrm{i}\hat{h}_{\ell,\ell+1}\delta t/2\hbar)$ sweeping across pairs of sites, as depicted in Fig. 3.2(a). A consequence of this expansion is that the approximation of the evolution operator remains unitary and so preserves the $L_2$-norm of a state vector $|\,\psi\rangle$. For time-dependent Hamiltonians it is clear from Eq. (3.25) that the effect of a finite $\delta t$ is an error that scales as $\mathcal{E}_{\delta t} \sim t_f \delta t$. However if the Hamiltonian is time-independent then the only source of error is the approximate expansion of the evolution operator for each time-step into two-site gates according to Eq. (3.27). In App. 3.B we show that the corresponding $L_2$-norm error incurred by using this expansion, which we call the Trotter error, scales at worst as $\mathcal{E}_{\mathrm{ST}} \sim t_f \delta t^2$.

### Applying a two-site unitary operator

The core of the TEBD algorithm is a procedure for applying a two-site nearest-neighbour operator to an MPS and performing a systematic truncation of the result back into an MPS of dimension $\chi$ [11, 12]. Using this procedure then allows the sequence of operators $\hat{S}_\ell$ making up the Trotterised approximation to $\exp(-\mathrm{i}\hat{H}\delta t)$ to be approximately applied to $|\,\psi\rangle$. Given an MPS of form $|\,\psi\rangle = \sum_{\mathbf{i}} A^{[1]i_1} A^{[2]i_2} \cdots A^{[M]i_M} |\,\mathbf{i}\rangle$ the operator $\hat{S}_\ell$ can be applied exactly to $|\,\psi\rangle$ and only affects the $A^{[\ell]}$ and $A^{[\ell+1]}$ tensors, as shown in step (i) of Fig. 3.2(b). This forms a new order-4 tensor associated to these sites with the contraction written out explicitly as

$$\Theta^{i_\ell i_{\ell+1}}_{\nu\ \mu} = \sum_{i'_\ell, i'_{\ell+1}=0}^{d-1} \langle i_\ell \,|\, \langle i_{\ell+1} \,|\, \hat{S}_\ell \,|\, i'_\ell \rangle \,\big|\, i'_{\ell+1} \rangle \sum_{\upsilon=1}^{\chi_\ell} A^{[\ell]i'_\ell}_{\nu\upsilon} A^{[\ell+1]i'_{\ell+1}}_{\upsilon\mu}. \tag{3.28}$$

The resulting state $\hat{S}_\ell\,|\,\psi\rangle$, depicted in step (ii) of Fig. 3.2(b), no longer has a proper MPS form due to the presence of a structureless two-site $\Theta$ tensor. To establish an MPS form we first reshape the $\Theta$ tensor into a conventional matrix by combining the indices $(i_\ell, \nu)$ and $(i_{\ell+1}, \mu)$ into so-called fat indices that represent the row and

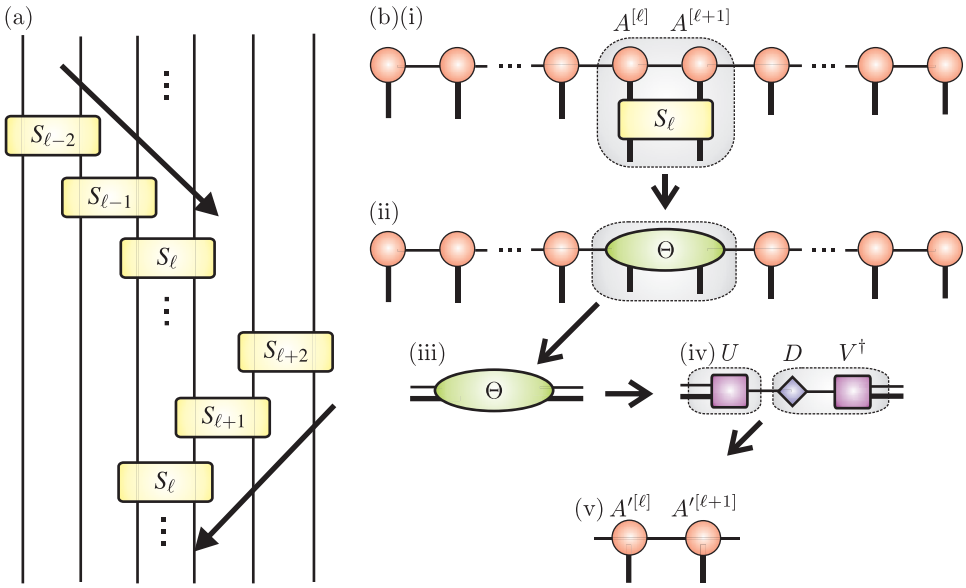


**Figure 3.2**. (a) A circuit diagram showing the second order Suzuki-Trotter expansion used throughout the calculations presented in this thesis. One time-step $\exp(-\mathrm{i}\hat{H}\delta t/\hbar)$ is approximated by applying two-site unitary operators $\hat{S}_\ell$ between each nearest-neighbour pair of sites, sweeping first from left to right and then back again. (b) A schematic tensor diagram of the core procedure in the TEBD algorithm [11, 12], described in the text, for applying a two-site nearest-neighbour operator. Firstly in (i) the physical legs of the tensor representing $\hat{S}_\ell$ are contracted with the relevant legs of the tensors $A^{[\ell]}$ in the MPS. This multiplies the state by $\hat{S}_\ell$ and involves only the tensors shaded. The result of this is a two-site order-4 tensor $\Theta$ shown in (ii). Next in (iii) the $\Theta$ tensor is reshaped as a matrix, resulting in two so-called fat legs. The most computationally expensive step is (iv) where an approximate matrix factorisation of $\Theta$, of the form $\tilde{\Theta} = UDV^\dagger$, is computed. Finally in (v) the reshaping of $U$ and $DV^\dagger$ is equivalent to unravelling their legs to form the newly updated $A'^{[\ell]}$ and $A'^{[\ell+1]}$ tensors. The diagonal matrix $D$ is absorbed into the $V^\dagger$ during the left to right sweep, as shown by the shading in (iv), and into the $U$ for the opposite direction. If the SVD is used then the $A'^{[\ell]}$ tensor, constructed from an orthogonal matrix only, automatically obeys one of the orthonormality constraints detailed in App. 3.A. Which way the matrix $D$ is absorbed then follows from requiring that the $\Theta$ tensor for the next two-site operation has indices that correspond to orthonormal bases, as discussed in Sec. 3.2.2.

columns respectively, as depicted in step (iii) of Fig. 3.2(b). The resulting $\Theta$ matrix is then subject to a matrix factorisation that approximately decomposes it into the form $\tilde{\Theta} = UDV^\dagger$, where $D$ is a $\mathbb{R}^{\chi\times\chi}$ diagonal matrix, $U \in \mathbb{C}^{d\chi\times\chi}, V \in \mathbb{C}^{d\chi\times\chi}$ and $\chi$ is the desired MPS dimension. This is step (iv) in Fig. 3.2(b) and is the most computationally expensive part of the procedure. After absorbing $D$ into either the $U$ or $V^\dagger$ matrix according to the direction of the Trotter sweep the matrices are reshaped into order-3 tensors $A'^{[\ell]}$ and $A'^{[\ell+1]}$. In this way, as shown in step (v) of Fig. 3.2(b), the factorisation is used to impose internal structure on the original $\Theta$ tensor and by replacing $\Theta$ with the new tensors establishes a proper MPS for $\hat{S}_\ell \,|\, \psi\rangle$. Crucially, however, the factorisation also provides the means of truncating the MPS dimension to some desired maximum $\chi$, controlling the growth of the description at the expense of becoming an approximation. We call the $L_2$-norm error due to these approximate factorisations the truncation error $\mathcal{E}_{\text{tr}}$ and we discuss the calculation of a quantity that bounds this error in App. 3.B.

In this part of the thesis the matrix factorisation we choose to apply to $\Theta$ is an SVD. We do this partly because of the analytical properties of the SVD discussed in the previous section. However it is also partly due to the numerical properties of the SVD, which we now discuss. For any $\Theta$ there exists an SVD of the form $\Theta = UDV^T$, where $U$ and $V$ are $\mathbb{C}^{d\chi\times d\chi}$ and $\mathbb{C}^{d\chi\times d\chi}$ unitary matrices respectively, whose columns are the singular vectors, and the elements of the diagonal matrix $D$ are the non-negative singular values $\lambda_\mu$ arranged in non-increasing order with $\mu$ [132]. The SVD is a unique decomposition up to the phases of its left and right singular vectors. These properties, as well as the Eckart-Young theorem in Eq. (3.24), not only make the SVD very mathematically appealing but have also aided in the construction of efficient, accurate and numerically robust algorithms for their computation [132]. Specifically the $L_2$-norm minimising rank-$\chi$ approximation $\tilde{\Theta}$, formed by truncating the inner dimension in the SVD from $d\chi$ to $\chi$, requires only a number of computa-

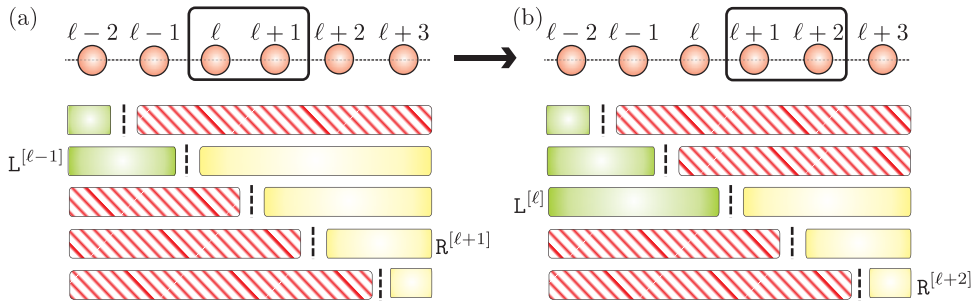


**Figure 3.3**. Following Ref. [138] we illustrate the orthonormality structure of an MPS. Here the boxes correspond to the sets of basis states for the left and right subsystems for different bipartitions (see App. 3.A for details). If these states form an orthonormal basis then the box is shaded and correspondingly it is hashed if they do not. (a) This is the situation obtained by following the procedure described in Fig. 3.2 up until $\hat{S}_{\ell-1}$ has just been applied and $\hat{S}_\ell$ is next to be applied. All splittings to the left of the one between sites $(\ell-1,\ell)$ have left states which are orthonormal while all those to the right have their right states orthonormal. The orthogonality centre is said to be located after site $\ell-1$. If the two-site unitary operator $\hat{S}_\ell$ is applied to sites $(\ell,\ell+1)$ then we can be assured that all the indices of the $\Theta$ tensor are orthonormal since subsystems $\mathtt{L}^{[\ell-1]}$ and $\mathtt{R}^{[\ell+1]}$ have orthonormal bases. (b) This is the situation after the application of $\hat{S}_\ell$. While the action of this non-unitary operator destroys the orthonormality of $\mathtt{R}^{[\ell-1]}$, the unitarity of the SVD decomposition establishes it in $\mathtt{L}^{[\ell]}$. This results in the orthogonality centre moving one site to the right. If we applied another two-site operator $\hat{S}_{\ell+1}$ to sites $(\ell+1,\ell+2)$ then we are again assured of the orthonormality of the tensor indices since subsystems $\mathtt{L}^{[\ell]}$ and $\mathtt{R}^{[\ell+2]}$ have orthonormal bases in (b). Finally, analogous changes in the orthonormality structure occur for two-site operations when moving to the left.

tions $\mathcal{O}(d^3\chi^3)$.

### Optimality of SVD within TEBD

For the calculation of arbitrary expected values, we would like to minimise the $L_2$-norm of the difference between $|\,\psi\rangle$ and its TEBD approximation $|\,\phi\rangle$. Within TEBD it is the local $\Theta$ tensor that is factorised and discussions in the previous section explain how the SVD gives the optimal approximation $\tilde{\Theta}$ to the $\Theta$ tensor with a given rank. Here we examine the optimality of this factorisation for the distance between $|\,\psi\rangle$ and $|\,\phi\rangle$.

A unique property of the $L_2$-norm is that it is unitarily invariant and so we are free to evaluate it in any orthonormal basis. For this reason the SVD truncation $\tilde{\Theta}$ of $\Theta$, which minimises the corresponding local error $\| \, |\,\Theta\rangle - |\,\tilde{\Theta}\rangle \|_2$, also minimises the global error $\| \, |\,\psi\rangle - |\,\phi\rangle \, \|_2$ between the corresponding vectors so long as the indices of $\tilde{\Theta}$ and $\Theta$ correspond to an orthonormal basis. In App. 3.A we discuss the orthonormality constraints that the tensors $A^{[\ell]}$ need to obey for this to be the case. In the context of imaginary time-evolution, where inherently non-unitary local unitary operators $\hat{S}_\ell = \exp(-\hat{h}_{\ell,\ell+1}\delta t/2\hbar)$ are applied, maintaining such orthonormal properties of the tensors $A^{[\ell]}$ is by no means guaranteed[5]. Yet it turns out that by applying the two-site operators in a sweeping sequence across the system the relevant bases for each pair are always orthonormal and the SVD truncation is optimal for each operation [138]. This is described in detail in Fig. 3.3. The second order Suzuki-Trotter expansion introduced earlier is based precisely on such a sweep and is therefore, combined with the SVD, an ideal choice for simulating generic evolution using an MPS if the $L_2$-distance is the cost function to be minimised.

### 3.2.3 Calculating expected values

Let us now briefly describe how the expected value $\langle\psi|\hat{O}|\psi\rangle$ of some operator $\hat{O}$ may be efficiently calculated from the MPS representation of $|\,\psi\rangle$. Specifically in this part we will be interested in either a single-site operator $\hat{O}_\ell = \hat{\mathbb{1}}_d^{\otimes(\ell-1)} \otimes \hat{o} \otimes \hat{\mathbb{1}}_d^{\otimes(M-\ell)}$ that acts on the Hilbert space of the $\ell$-th site with the local operator $\hat{o}$, or a two-site operator $\hat{O}_{\ell\ell'} = \hat{\mathbb{1}}_d^{\otimes(\ell-1)} \otimes \hat{o}_1 \otimes \hat{\mathbb{1}}_d^{\otimes(\ell'-\ell-1)} \otimes \hat{o}_2 \otimes \hat{\mathbb{1}}_d^{\otimes(M-\ell')}$ that acts on the $\ell$-th and $\ell'$-th site with $\hat{o}_1$ and $\hat{o}_2$, respectively. Here $\hat{\mathbb{1}}_d$ is the $d$-dimensional identity operator.

The calculation of $\langle\psi|\hat{O}_\ell|\psi\rangle$ is illustrated in Fig. 3.4(a). First, the matrix $o$ with elements $o_{ii'} = \langle i'|\hat{o}|i\rangle$ is built, requiring $\mathcal{O}(d^2)$ computations. Then for each site

[5] A similar situation occurs for classical stochastic evolution, as we will see in Part II, and the evolution of mixed states according to a quantum master equation [138].

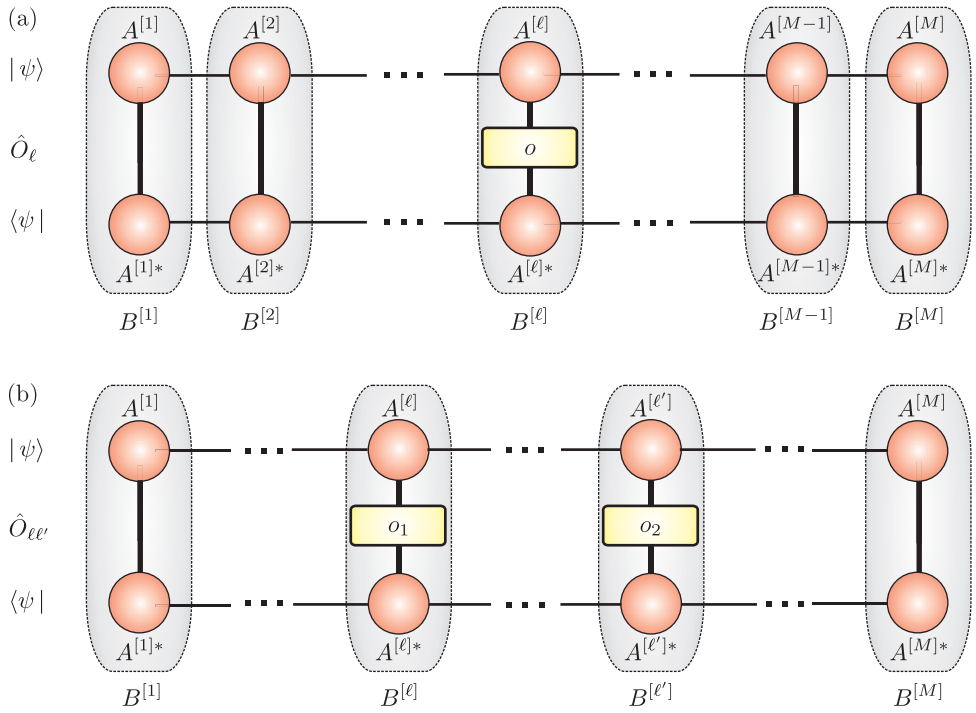


**Figure 3.4**. The calculation of the expected value of a (a) single-site $\langle\psi|\hat{O}_\ell|\psi\rangle$ and (b) two-site $\langle\psi|\hat{O}_{\ell\ell'}|\psi\rangle$ operator. The tensors within the $\ell$-th silver shaded area are contracted to form a tensor $B^{[\ell]}$. When each is reshaped into a matrix, multiplying them together results in the desired expected value.

$\ell' \neq \ell$ we compute the tensor $B^{[\ell']}$ with elements $B^{[\ell']}_{\nu\mu\upsilon\varsigma} = \sum_i A^{[\ell']i*}_{\nu\mu} A^{[\ell']i}_{\upsilon\varsigma}$, requiring $\mathcal{O}(Md\chi^4)$ computations. For site $\ell$ we build a tensor $B^{[\ell]}$ with elements $B^{[\ell]}_{\nu\mu\upsilon\varsigma} = \sum_{ii'} A^{[\ell]i'*}_{\nu\mu} o_{ii'} A^{[\ell]i}_{\upsilon\varsigma}$, requiring $\mathcal{O}(d^2\chi^4)$ computations. Next, the indices of each $B^{[\ell]}$ are combined into two indices $(\nu\mu)$ and $(\upsilon\varsigma)$, thus reshaping each $B^{[\ell']}$ into a matrix of size $\chi^2 \times \chi^2$. Last, the ordered product of these matrices gives the desired expected value $B^{[1]}B^{[2]}\cdots B^{[M]} = \langle\psi|\hat{O}_\ell|\psi\rangle$, requiring $\mathcal{O}(M\chi^4)$ computations. The procedure for $\hat{O}_{\ell\ell'}$ is similar and is illustrated in Fig. 3.4(b).

### 3.2.4 Discussion

In this section we have introduced the TEBD method used to obtain many of the results presented in this thesis. Here we shall briefly mention the advantages, drawbacks and applications of this method (discussions relating to simulating classical stochastic systems are left to Part II).

TEBD has some very favourable qualities. For instance the two sources of error — the discretisation of time, and the small $\chi$ compression — are accessible, controllable and well-understood. As an example of accessibility, we can calculate a bound for the error due to truncation for any expected value calculated during a simulation (see App. 3.B). If the errors are largely caused by the finite time-step then it is clear how to control the error since it scales linearly with $\delta t$ for time-dependent Hamiltonians and, for the particular Trotter expansion used here, as $\delta t^2$ for time-independent Hamiltonians (again, see App. 3.B). The well-understood physical cause of errors, entanglement, is also useful when planning simulations. For example, simulations with periodic boundary conditions are known to require more resources than those with open boundary conditions since the entanglement between the two sides of the system is increased. Even without these useful properties, in practice we can often reassure ourselves that our calculations are correct by varying $\delta t$ and $\chi$, and

checking that the expected values of interest converge satisfactorily.

The main drawback of MPS methods are that they are restricted to problems in one dimension. It is the linear geometry of the system that ensures the area between two subsystems forming a contiguous bipartition does not increase with the size of the system, which then limits their entanglement entropy [135]. If trying to describe, for instance, a two-dimensional system then the entanglement entropy between any two bipartitions, and thus the MPS dimension $\chi$ required to describe the system accurately, will grow quickly with the system size. This makes MPS methods unsuitable for high-dimensional geometries.

However, tensor network states (the generalisation of MPS) have been shown to be effective for geometries beyond one dimension. For example, the TEBD algorithm has been extended to systems with a tree geometry or a geometry with a bounded tree-width [139, 140]. For two dimensions (with potential for higher dimensions) a class of tensor network states called projected entangled pair states (PEPS) have been used to simulate ground states and dynamics [141, 142]. The scaling of the resources with bond dimension $\chi$ for such algorithms is very high, making the simulations expensive and requiring highly optimised code and state-of-the-art hardware for large systems. Further, contracting the PEPS, e.g. in order to calculate expectation values, exactly is inefficient and so approximate methods are needed. Another approach for dimensions greater than one, with similar drawbacks, is multiscale entanglement renormalisation [143]. Finally, though not relevant to the types of dynamics we consider in this thesis, for completeness we mention that MPS and TEBD have been extended for infinite homogeneous systems [144] and quantum fields [145].

To close the chapter we remark on an interesting connection between simulating quantum evolution on a classical computer and performing other types of classical computational tasks. Tensor network-based algorithms applied to weakly-entangled

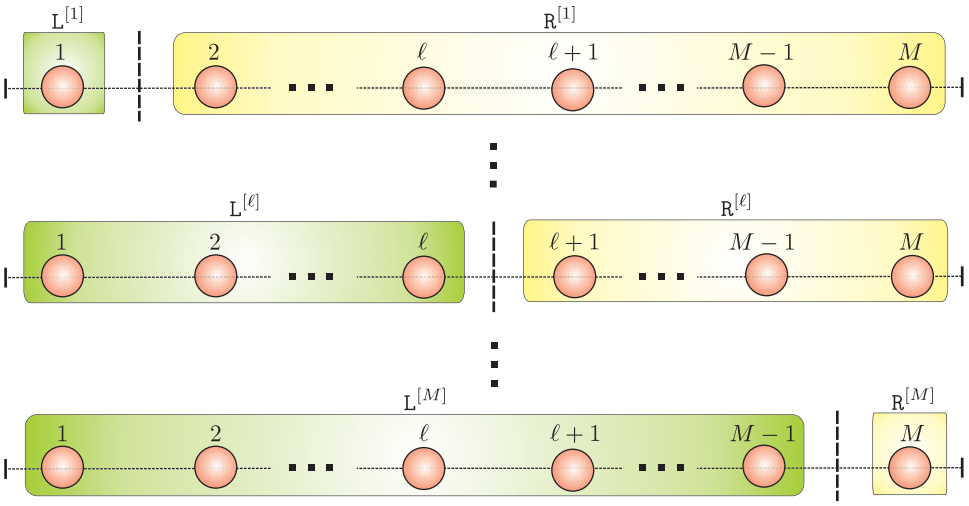


**Figure 3.5**. The sequence of $M-1$ contiguous partitions (according to the labelling imposed) of the system in which the Schmidt decompositions are computed.

systems are among a class of simulators called efficient strong simulators. This means that calculating the probability of a measurement outcome to $n_{\text{dec}}$ decimal places requires resources that grow at most polynomially in $n_{\text{dec}}$ and the system size $M$. In work not presented in this thesis we have shown that any efficient strong simulator can be used to efficiently solve seemingly unrelated search problems [6]. For the specific case of tensor network methods this means that search problems whose solutions are checked by a weakly-entangling circuit are efficiently solvable.

## 3.A Constructing an MPS via repeated SVD

In this appendix we follow Vidal [11] and show how, through the repeated use of the SVD (Schmidt decomposition), an MPS of any vector $|\psi\rangle$ can, in principle, be found. To do this we first compute the Schmidt decomposition of $|\psi\rangle$ for every contiguous bipartition into blocks $\mathtt{L}^{[\ell]}$ and $\mathtt{R}^{[\ell]}$, as depicted in Fig. 3.5. Starting from the left boundary, each left Schmidt state $|\mathtt{L}^{[\ell]}_{\mu_\ell}\rangle$ is iteratively expanded in terms of

the local basis of the rightmost site of the block $|\, i_\ell \rangle$ and the left Schmidt states $|\, \mathtt{L}^{[\ell-1]}_{\mu_{\ell-1}} \rangle$ of the neighbouring splitting one site further to the left. This gives the following set of expansions

$$
\begin{aligned}
|\, \mathtt{L}^{[1]}_{\mu_1} \rangle &= \sum_{i_1=0}^{d-1} A^{[1]i_1}_{\mu_1} \,|\, i_1 \rangle \,, && \text{(3.29a)} \\
|\, \mathtt{L}^{[2]}_{\mu_2} \rangle &= \sum_{i_2=0}^{d-1} \sum_{\mu_1=1}^{X_1} A^{[2]i_2}_{\mu_1\mu_2} |\, \mathtt{L}^{[1]}_{\mu_1} \rangle \,|\, i_2 \rangle \,, && \text{(3.29b)} \\
&\vdots \\
|\, \mathtt{L}^{[\ell]}_{\mu_\ell} \rangle &= \sum_{i_\ell=0}^{d-1} \sum_{\mu_{\ell-1}=1}^{X_{\ell-1}} A^{[\ell]i_\ell}_{\mu_{\ell-1}\mu_\ell} |\, \mathtt{L}^{[\ell-1]}_{\mu_{\ell-1}} \rangle \,|\, i_\ell \rangle \,, && \text{(3.29c)} \\
&\vdots \\
|\mathtt{L}^{[M]}_{\mu_M=1} \rangle &= |\, \psi \rangle = \sum_{i_M=0}^{d-1} \sum_{\mu_{M-1}=1}^{X_{M-1}} A^{[M]i_M}_{\mu_{M-1}} |\, \mathtt{L}^{[M-1]}_{\mu_{M-1}} \rangle \,|\, i_M \rangle \,. && \text{(3.29d)}
\end{aligned}
$$

The expansion coefficients can be seen to define the tensors $A^{[\ell]}$ of an MPS representation of $|\, \psi \rangle$ by taking the last expansion of $|\, \psi \rangle$ in Eq. (3.29d) and inserting all the others into it until no Schmidt states remain

$$
|\, \psi \rangle = \sum_{\mathbf{i}} \left\{ \sum_{\mu_1=1}^{X_1} \sum_{\mu_2=1}^{X_2} \cdots \sum_{\mu_{M-1}=1}^{X_{M-1}} A^{[1]i_1}_{\mu_1} A^{[2]i_2}_{\mu_1\mu_2} \cdots A^{[M]i_M}_{\mu_{M-1}} \right\} |\, \mathbf{i} \rangle \,. \tag{3.30}
$$

The term in the parentheses is identical to the matrix product expansion of the amplitude $\psi_{\mathbf{i}}$ given in Eq. (3.18) but with all the summations written out explicitly. Truncating each summation (equivalent to truncating the Schmidt decomposition or SVD) to $\chi$ terms results in $|\, \tilde{\psi} \rangle$, an approximate MPS of dimension $\chi$.

While the full product of matrices $A^{[\ell]i_\ell}$ generates $|\, \psi \rangle$, a partial product up to a site $\ell$ instead forms an MPS of a left Schmidt state $|\, \mathtt{L}^{[\ell]}_{\mu_\ell} \rangle$ as

$$
|\, \mathtt{L}^{[\ell]}_{\mu_\ell} \rangle = \sum_{\mathbf{i}} (A^{[1]i_1} A^{[2]i_2} \cdots A^{[\ell]i_\ell})_{\mu_\ell} \,|\, \mathbf{i} \rangle_{\mathtt{L}} \,. \tag{3.31}
$$

The orthonormality of the left Schmidt states is equivalent to all tensors $A^{[\ell]}$ derived from this construction obeying a left orthonormality property [127]

$$\sum_{i_\ell=0}^{d-1} (A^{[\ell]i_\ell})^\dagger A^{[\ell]i_\ell} = \mathbb{1}_{X_\ell}, \tag{3.32}$$

where $\mathbb{1}_{X_\ell}$ is the $X_\ell \times X_\ell$ identity matrix. Notice that we could have equally performed this construction starting from the right boundary and using the right Schmidt states $|\,\mathtt{R}^{[\ell]}_{\mu_\ell}\rangle$. This would result in a different set of tensors $A^{[\ell]}$ obeying a right orthonormality property

$$\sum_{i_\ell=0}^{d-1} A^{[\ell]i_\ell} (A^{[\ell]i_\ell})^\dagger = \mathbb{1}_{X_{\ell-1}}. \tag{3.33}$$

If we were to apply the left procedure up to and including site $\ell$ and the right procedure up to an including site $\ell+1$ then the resulting MPS would be of the form

$$|\,\psi\rangle = \sum_{\mathbf{i}} A^{[1]i_1} A^{[2]i_2} \cdots A^{[\ell]i_\ell} D^{[\ell]}\, A^{[\ell+1]i_{\ell+1}} \cdots A^{[M]i_M}\,|\,\mathbf{i}\rangle\,, \tag{3.34}$$

where $D^{[\ell]}$ is the diagonal matrix of singular values $\lambda^{[\ell]}_{\mu_\ell}$. This MPS has an orthogonality centre, or twist in its handedness, located precisely at the bipartition after site $\ell$, since the constraints Eq. (3.32) and Eq. (3.33) are obeyed by the tensors to the left and right of $D^{[\ell]}$, respectively. This is crucial since we can readily extract from the MPS in Eq. (3.34) an expansion of $|\,\psi\rangle$ in an orthonormal basis

$$|\,\psi\rangle = \sum_{\mu_{\ell-1},\mu_{\ell+1}} \sum_{i_\ell, i_{\ell+1}} \psi^{i_\ell i_{\ell+1}}_{\mu_{\ell-1}\mu_{\ell+1}} |\,\mathtt{L}^{[\ell-1]}_{\mu_{\ell-1}}\rangle\,|\,i_\ell\rangle\,|\,i_{\ell+1}\rangle\,|\,\mathtt{R}^{[\ell+1]}_{\mu_{\ell+1}}\rangle, \tag{3.35}$$

by forming the two-site tensor about the orthogonality centre

$$\psi^{i_\ell i_{\ell+1}}_{\mu_{\ell-1}\mu_{\ell+1}} = \sum_{\mu_\ell} A^{[\ell]i_\ell}_{\mu_{\ell-1}\mu_\ell} \lambda^{[\ell]}_{\mu_\ell} A^{[\ell+1]i_{\ell+1}}_{\mu_\ell\mu_{\ell+1}}. \tag{3.36}$$

It is only for this pair of sites, or $(\ell-1,\ell)$ and $(\ell+1,\ell+2)$ adjacent to the orthogonality centre, that a tensor of amplitudes for $|\,\psi\rangle$ can be formed from components of the MPS in Eq. (3.34) with all indices corresponding to orthonormal states. As discussed in Sec. 3.2.2, moving the orthogonality centre so as to ensure that it is always located adjacent to any pair of sites on which a two-site operator is applied is essential for the optimality of the subsequent SVD decimation.

## 3.B TEBD errors

Let us start by deriving the scaling of the error introduced by the finite time-step $\delta t$. As said previously, for a time-dependent Hamiltonian the dominant source of error is due to the approximation of the true time-dependence of the Hamiltonian by some piece-wise constant Hamiltonian that ignores the time-dependence within each time-step $\delta t$. For a fixed simulation time $t_f$ the resulting error is linear in the time-step $\mathcal{E}_{\delta t} \sim t_f \delta t$. However, for a time-independent Hamiltonian this is not a source of error and instead the only error is that due to the Suzuki-Trotter expansion (see Eq. (3.27)). Let us now consider the error due to this approximation for the time-independent case. In approximating the full unitary evolution $\exp(-\mathrm{i}\hat{H}t_f/\hbar)$ by a product of $n_{\text{steps}}$ time-steps, each an identical unitary evolution

$\hat{S} = (\prod_{\ell=1}^{M-1} \hat{S}_\ell)(\prod_{\ell=M-1}^{1} \hat{S}_\ell)$, we quantify the error by the $L_2$-norm of the residual

$$\begin{aligned}
\mathcal{E}_{\mathrm{ST}} &= \| \left\{ \mathrm{e}^{-\mathrm{i}\hat{H}t_f/\hbar} - \left[ \prod_{j=1}^{n_{\mathrm{steps}}} \hat{S} \right] \right\} | \, \psi(0) \rangle \, \|_2 \\
&\leq \| (\mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S}) \left[ \prod_{j=1}^{n_{\mathrm{steps}}-1} \hat{S} \right] | \, \psi(0) \rangle \, \|_2 \\
&+ \| \mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} (\mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S}) \left[ \prod_{j=1}^{n_{\mathrm{steps}}-2} \hat{S} \right] | \, \psi(0) \rangle \, \|_2 + \\
&\cdots + \| \left[ \prod_{j=3}^{n_{\mathrm{steps}}} \mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} \right] (\mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S}) \hat{S} \, | \, \psi(0) \rangle \, \|_2 \\
&+ \| \left[ \prod_{j=2}^{n_{\mathrm{steps}}} \mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} \right] (\mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S}) \, | \, \psi(0) \rangle \, \|_2 \\
&\leq n_{\mathrm{steps}} \| \mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S} \|_2 .
\end{aligned} \tag{3.37}$$

Here we have used that the $L_2$-norm is both convex and invariant under the action of a unitary operator [133]. Since $\| \mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar} - \hat{S} \|_2 \sim \delta t^3$ and $n_{\mathrm{steps}} = t_f/\delta t$ the error scales at worst as $\mathcal{E}_{\mathrm{ST}} \sim t_f \delta t^2$.

Next let us return to the general time-dependent case and consider the error due to truncating the MPS description after applying each two-site gate. For convenience we relabel the total set of two-site unitary evolution operators over the whole simulation as $\hat{S}_{l_{\max}} \hat{S}_{l_{\max}-1} \cdots \hat{S}_l \cdots \hat{S}_2 \hat{S}_1 \, | \, \psi(0) \rangle$. Here $l_{\max} = 2 n_{\mathrm{steps}} M$ and the index $l$ denotes the position in the sequence.

Consider applying the two-site operator $\hat{S}_l$ to the state $| \, \psi_{l-1} \rangle$ after $l-1$ gates have been applied. Rather than being applied exactly, i.e., $\hat{S}_l \, | \, \psi_{l-1} \rangle$, the subsequent factorisation and truncation back to an MPS with a maximum inner-dimension of $\chi$ produces a state $| \, \psi_l \rangle$. This can be thought of as being equivalent to implementing exactly some other, generally non-unitary, transformation $\hat{T}_l$ that depends on $l$, $| \, \psi_{l-1} \rangle$ and $\chi$ such that $| \, \psi_l \rangle = \hat{T}_l \, | \, \psi_{l-1} \rangle$. The error in doing this is then quantified

by the $L_2$-norm $\epsilon_l = ||(\hat{S}_l - \hat{T}_l)\,|\,\psi_{l-1}\rangle\,||_2$. Once the sequence has been performed up to $l_{\max}$ the accumulated approximation is $\hat{T}_{l_{\max}}\hat{T}_{l_{\max}-1}\cdots\hat{T}_1\,|\,\psi(0)\rangle$. For each step $l$ it is possible to extract from the TEBD algorithm the $L_2$-norm error $\epsilon_l = ||(\hat{S}_l - \hat{T}_l)\hat{T}_{l-1}\cdots\hat{T}_1\,|\,\psi(0)\rangle\,||_2$ caused by truncating after applying $\hat{S}_l$ to the accumulative approximate state up to point $l-1$. This information then provides an upper-bound to the total accumulative $L_2$-norm error due to truncation $\mathcal{E}_{\mathrm{tr},||\cdot||_2}$ after $l_{\max}$ two-site operations since

$$\begin{aligned}
\mathcal{E}_{\mathrm{tr},||\cdot||_2} &= \|\left\{\prod_{l=1}^{l_{\max}}\hat{S}_l - \prod_{l=1}^{l_{\max}}\hat{T}_l\right\}|\,\psi(0)\rangle\,\|_2 \\
&\leq \|(\hat{S}_{l_{\max}} - \hat{T}_{l_{\max}})\prod_{l=1}^{l_{\max}-1}\hat{T}_l\,|\,\psi(0)\rangle\,\|_2 \\
&+ \|\hat{S}_{l_{\max}}(\hat{S}_{l_{\max}-1} - \hat{T}_{l_{\max}-1})\prod_{l=1}^{l_{\max}-2}\hat{T}_l\,|\,\psi(0)\rangle\,\|_2 + \\
&\cdots + \|\prod_{l=3}^{l_{\max}}\hat{S}_l(\hat{S}_2 - \hat{T}_2)\hat{T}_2\,|\,\psi(0)\rangle\,\|_2 + \|\prod_{l=2}^{l_{\max}}\hat{S}_l(\hat{S}_1 - \hat{T}_1)\,|\,\psi(0)\rangle\,\|_2 \\
&\leq \|(\hat{S}_{l_{\max}} - \hat{T}_{l_{\max}})\prod_{l=1}^{l_{\max}-1}\hat{T}_l\,|\,\psi(0)\rangle\,\|_2 + \|(\hat{S}_{l_{\max}-1} - \hat{T}_{l_{\max}-1})\prod_{l=1}^{l_{\max}-2}\hat{T}_l\,|\,\psi(0)\rangle\,\|_2 \\
&\cdots + \|(\hat{S}_2 - \hat{T}_2)\hat{T}_1\,|\,\psi(0)\rangle\,\|_2 + \|(\hat{S}_1 - \hat{T}_1)\,|\,\psi(0)\rangle\,\|_2 \\
&= \sum_{l=1}^{l_{\max}}\epsilon_l = \mathcal{E}^{+}_{\mathrm{tr},||\cdot||_2}.
\end{aligned} \tag{3.38}$$

Again we have used the convexity of the $L_2$-norm and its unitary invariance [133]. Thus the truncation error $\mathcal{E}_{\mathrm{tr},||\cdot||_2}$ grows at worst additively during the Trotter sequence and can be monitored within the algorithm by computing the sum of the upper-bounds for each two-site operation. This gives us a (typically conservative) upper-bound $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_2} \geq \mathcal{E}_{\mathrm{tr},||\cdot||_2}$.

# Chapter 4

# Breathing oscillations of a trapped impurity in a Bose gas

Having introduced Bose gases and their description in Ch. 2, and numerical methods for their simulation in Ch. 3, we are now in a position to present our novel physical contribution to the field. Specifically, for the rest of this part we will describe and simulate an atomic impurity in a Bose gas. In Ch. 5 we will focus on the transport of an externally driven impurity through a Bose gas, while in this chapter we focus on the dissipative dynamics of a trapped impurity initialised in a highly-localised non-equilibrium state within a Bose gas[1].

Before doing this, let us briefly discuss the more general topic of cold atomic mixtures, i.e., combinations of different elements or atoms of the same element but with different internal states, that this work fits into. Such mixtures realise a rich variety of quantum matter with highly exotic phases that mimic and more often go beyond what is seen in traditional condensed matter systems [95], e.g., spinor condensates [146], spin Hamiltonians [147, 148] and quantum spin liquid crystals [149, 150, 151]. Moreover, mixing a Fermi gas with another species provides a means of thermalising the Fermi gas at low temperatures (since $s$-wave scattering between

[1]The main results of this chapter were first presented by us in Ref. [4].

identical fermions is prohibited), allowing, for instance, the application of evaporative cooling [152]. Of particular relevance to our work, highly-imbalanced mixtures make it possible to investigate the transport, interactions and decoherence of single atom impurities immersed in a background atomic gas [153, 13, 21, 20, 22, 17]. As a prominent example, signatures of polaron effects, caused by impurity-induced density fluctuations of the background gas, have been investigated theoretically [24, 154] and observed in experiments [217, 156]. Impurities also offer a means of probing the environment in which they are immersed [19, 157, 16, 15, 158, 21, 159], for example, extracting work functions [160, 161] or temperature [18, 162].

Catani *et al.* recently created a harmonically trapped impurity suspended in a separately trapped Bose gas [14]. They studied the dynamics of the system following a sudden lowering of the trap frequency of the impurity. Primarily, they observed breathing oscillations of the width of the impurity density distribution for various impurity-Bose gas interaction strengths. Several features of the experiment were amenable to interpretation in terms of a quantum Langevin equation in conjunction with a polaronic mass shift, however, we show here that at even lower temperatures a different model is required to fully describe the dynamics of the system. Developing such a model is the goal of this chapter.

Specifically, our model provides a versatile intuitive physical picture of the breathing oscillations of an impurity in a Bose gas at zero temperature. At first our analysis is based on a variational approach in the Gross-Pitaevskii (GP) regime of the Bose gas (cf. Sec. 2.1.2). We show that the impurity density distribution, which we describe by a Gaussian, has a width $\sigma$ obeying a Newtonian equation of motion for a fictitious particle with position $\sigma$. The potential governing this motion accounts for the quantum pressure of the impurity, the inhomogeneity of the trapped Bose gas and the localised distortion of this background gas induced by the impurity. The latter leads to a strong confinement of the impurity, known as

self-trapping [163, 164, 165, 166]. We subsequently extend our model by including excitations of the Bose gas in the form of Bogoliubov phonons (cf. Sec. 2.1.2). A variational ansatz describing the phonons as a product of coherent states allows us to track the evolution of the system, including the exchange of energy between the impurity and the phonon bath. The dynamics turns out to be non-Markovian because of the back-action of phonons created by the impurity and we show that the timescale of memory effects can be varied by adjusting the trapping parameters, allowing for a comprehensive study of the transition between Markovian and non-Markovian dynamics.

Our results reproduce the main features observed in the experiment of Catani *et al.* despite neglecting finite temperature effects. They also demonstrate that cooling the system further than in the experiment would introduce novel effects. We obtain the experimentally observed amplitude reduction of the breathing oscillations for large impurity-Bose gas interactions. However, unlike in the experiment, this reduction takes the form of a sudden quench in the breathing amplitude once the impurity-Bose gas coupling substantially exceeds the boson-boson interaction strength. Also we find that the damping of the breathing oscillations, due to the net dissipation of energy into the Bose gas, is highly non-Markovian, contrary to their theoretical treatment of the system [14]. Moreover, our model explains the different behaviours of attractive and repulsive impurities as a consequence of the inhomogeneity of the trapped Bose gas while attractive impurities were not considered in the experiment.

Before we present our results, we give the relevant details of the experiment by Catani *et al.* in Sec. 4.1. Section 4.2 then contains our theoretical description of this system. We close in Sec. 4.3 by comparing our work with more recent theoretical work by others on the same system.

## 4.1 Experimental setup and procedure

### 4.1.1 Setup

We begin by briefly describing the experimental setup and procedure used by Catani *et al.* [14]. They took a mixture of $^{87}$Rb and $^{41}$K featuring a magnetically-tunable inter-species scattering length. They cooled and trapped this mixture with two orthogonal pairs of intense counter-propagating laser beams, creating an array of quasi one-dimensional non-interacting tubes along the remaining direction (cf. Secs. 2.2 and 2.3), which we refer to as the axial direction. The residual potential along this axial direction, due to the transverse profile of the laser beams, was harmonic with frequency $\Omega_b = 2\pi \times 62$ Hz ($\Omega_2 = 2\pi \times 87$ Hz) for the Rb (K) atoms.

The experiment featured roughly 180 Rb atoms and 1.4 K atoms per tube. The effective one-dimensional inter-species interaction strength was varied between zero and 30 times the intra-species Rb value $g = 2.36 \times 10^{-37}$ Jm$^{-1}$. At their disposal they also had a species-specific dipole potential (SSDP), which when turned on resulted in an additional tight harmonic trap for the K atoms of frequency $2\pi \times 1$ kHz.

At any point in the experiment the density of the K atoms could be measured by adding a tight optical lattice in the axial direction to freeze the motion of these atoms and then taking an *in situ* absorption image.

### 4.1.2 Procedure

In the experiment the SSDP was initially turned on and the mixture brought into thermal equilibrium. Then the SSDP was suddenly extinguished and the mixture left to evolve for some time $t$. Following this, the density of the K atoms was imaged. The width $\sigma$ of the imaged density distribution was found to oscillate as a function of $t$.

### 4.1.3 Differences

In the next section we model a single tube in a slight variation on the Catani *et al.* experiment. The result for many tubes can be obtained by averaging over the results predicted for each tube.

The main difference between the model we consider in the next section and the Catani *et al.* experiment is that we will consider the system to be at zero temperature. We also utilise the GP and Bogoliubov descriptions of the Bose gas, which for a quasi one-dimensional system are only strictly valid in the regime $m_b g/\hbar^2 n \ll 1$, where $g$ ($n$) is the effective one-dimensional interaction strength (density) of the Bose gas (see Sec. 2.2). In the Catani *et al.* experiment the one-dimensional density took a maximum value $n = 7\ \mu\text{m}^{-1}$ that together with the interaction strength $g$ given in the previous subsection results in $m_b g/\hbar^2 n \approx 5$. So the experiment took place outside the regime in which our theoretical approaches are strictly valid. Further, we ignore any interactions between K atoms by assuming there to be only a single atom and generalise our model to allow the quasi dimensionality of the system to be increased beyond one.

## 4.2 Theoretical description

### 4.2.1 Model

The specific model we consider consists of a single impurity of mass $m_a$ in a harmonic trap with frequency $\Omega_a$. Together with this are many identical bosons of mass $m_b$ that weakly interact with strength $g$ and are trapped by a separate external potential $v_b(\mathbf{r})$. The impurity interacts with the bosons with interaction strength $\eta g$, where the dimensionless parameter $\eta$ controls the relative strengths of interactions. As such the many-body Hamiltonian for the combined system is given by $\hat{H} = \hat{H}_a + \hat{H}_b + \hat{H}_{ab}$,

where

$$\hat{H}_a = \int \mathrm{d}\mathbf{r}\hat{\varkappa}^\dagger \left(-\frac{\hbar^2\nabla^2}{2m_a} + \frac{m_a}{2}\Omega_a^2 r^2\right)\hat{\varkappa}, \tag{4.1a}$$

$$\hat{H}_b = \int \mathrm{d}\mathbf{r}\hat{\phi}^\dagger \left(-\frac{\hbar^2\nabla^2}{2m_b} + v_b + \frac{g}{2}\hat{\phi}^\dagger\hat{\phi}\right)\hat{\phi}, \tag{4.1b}$$

$$\hat{H}_{ab} = \eta g \int \mathrm{d}\mathbf{r}\hat{\varkappa}^\dagger\hat{\varkappa}\hat{\phi}^\dagger\hat{\phi}, \tag{4.1c}$$

describe the impurity, bosons and impurity-Bose gas coupling, respectively. Here $\hat{\varkappa}(\mathbf{r})$ and $\hat{\phi}(\mathbf{r})$ are the impurity and boson field operators, respectively.

### 4.2.2 Breathing oscillations in the GP regime

We now extend the single-species time-dependent GP approach of Sec. 2.1.2 (valid for low temperatures and weak interactions) to a Bose gas interacting with a single impurity. We assume the impurity-Bose gas system to be in a product state. We write the wavefunction of the single impurity as $\varkappa(\mathbf{r}, t)$. For the condensed part of the Bose gas, we make the Hartree-Fock ansatz. It then follows from Eqs. (4.1) that the impurity and condensate wavefunctions evolve according to

$$\mathrm{i}\hbar\frac{\partial\varkappa}{\partial t} = \left(-\frac{\hbar^2\nabla^2}{2m_a} + \frac{m_a}{2}\Omega_a^2 r^2 + \eta g|\varphi|^2\right)\varkappa\,, \tag{4.2a}$$

$$\mathrm{i}\hbar\frac{\partial\varphi}{\partial t} = \left(-\frac{\hbar^2\nabla^2}{2m_b} + v_b + \eta g|\varkappa|^2 + g|\varphi|^2\right)\varphi, \tag{4.2b}$$

respectively, where $\varphi$ is normalised to the number of particles $N_0$ in the condensate, which we have assumed to be large [48].

We consider the condensate to be in the Thomas-Fermi regime (mentioned in Sec. 2.2.3) in which the kinetic energy of the condensate, represented by the first term on the right hand side of Eq. (4.2b), is negligible. In this regime, if the impurity and Bose gas are decoupled the condensate evolves as $\varphi_0(\mathbf{r}, t) = \sqrt{n_0(\mathbf{r})}\mathrm{e}^{-\mathrm{i}\mu_b t/\hbar}$ (see

Sec. 2.1.2), where the decoupled condensate density is given by

$$n_0(\mathbf{r}) = \frac{\mu_b - v_b(\mathbf{r})}{g}, \tag{4.3}$$

for $\mu_b > v_b(\mathbf{r})$ and zero otherwise, and $\mu_b$ is the chemical potential of the bosons. For non-zero impurity-boson coupling, we assume a similar form for the time-evolution of the condensate $\varphi_\eta(\mathbf{r}, t) = \sqrt{n_\eta(\mathbf{r})}\mathrm{e}^{-\mathrm{i}\mu_b t/\hbar}$. It then follows from Eq. (4.2b) that for attractive (repulsive) interactions the condensate density $n_\eta(\mathbf{r}) = n_0(\mathbf{r}) - \eta|\varkappa(\mathbf{r})|^2$ is enhanced (suppressed) from its decoupled value at the location of the impurity. The self-consistency of our assumption regarding the time-evolution of the condensate requires $n_0(\mathbf{r}) \gg \eta|\varkappa(\mathbf{r})|^2$, i.e., sufficiently weak impurity-boson coupling.

As a consequence of these approximations Eq. (4.2a) becomes the self-focusing non-linear Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial\varkappa}{\partial t} = \left(-\frac{\hbar^2\nabla^2}{2m_a} + \eta g n_0 - \eta^2 g|\varkappa|^2 + \frac{m_a}{2}\Omega_a^2 r^2\right)\varkappa. \tag{4.4}$$

We now specialise to the important case of a spherically symmetric system in $\mathtt{D}$ dimensions with the Bose gas trapped in the harmonic potential $v_b(\mathbf{r}) = \frac{1}{2}m_b\Omega_b^2 r^2$ of frequency $\Omega_b$. Solving the equation $N_0 = \int \mathrm{d}\mathbf{r} n_0(\mathbf{r})$, with $n_0(\mathbf{r})$ given by Eq. (4.3), we find the chemical potential to be

$$\mu_b = \hbar\Omega_b\left[\frac{(2+\mathtt{D})\,\Gamma\left(1+\mathtt{D}/2\right)}{2\,(2\pi)^{\mathtt{D}/2}}\frac{N_0 g}{\hbar\Omega_b w_b^{\mathtt{D}}}\right]^{2/(2+\mathtt{D})}, \tag{4.5}$$

where $w_b = \sqrt{\hbar/m_b\Omega_b}$ is the width of the single-particle ground state and $\Gamma(m) = \int_0^\infty \mathrm{d}x\mathrm{e}^{-x}x^{m-1}$ is the gamma function. We have used that for a spherically symmetric function $f(\mathbf{r}) = f(r)$,

$$\int \mathrm{d}\mathbf{r} f(\mathbf{r}) = \frac{\mathtt{D}\pi^{\mathtt{D}/2}}{\Gamma(1+\mathtt{D}/2)}\int \mathrm{d}r r^{\mathtt{D}-1} f(r). \tag{4.6}$$

We next solve Eq. (4.4) within the variational ansatz

$$\varkappa(r,\sigma,\gamma) = \left(\pi\sigma^2\right)^{-\mathtt{D}/4} \exp\left(-\frac{r^2}{2\sigma^2} - \mathrm{i}\gamma r^2\right), \tag{4.7}$$

with time-dependent width $\sigma$ and phase $\gamma$. This ansatz was shown to describe a Bose gas in a harmonic trap accurately (see Ref. [64] and references therein). We find equations of motion for the Gaussian parameters by extremising the action $S = \int \mathrm{d}t\, \mathrm{d}\mathbf{r}\, \mathcal{L}$ with respect to them, where the Lagrangian density is

$$\mathcal{L} = \varkappa^* \left[\mathrm{i}\hbar\frac{\partial}{\partial t} + \frac{\hbar^2\nabla^2}{2m_a} - \eta g n_0 + \frac{\eta^2}{2} g|\varkappa|^2 - \frac{m_a}{2}\Omega_a^2 r^2\right] \varkappa\,. \tag{4.8}$$

We show in App. 4.A.1 that this extremisation results in $\gamma = -m_a\dot{\sigma}/2\hbar\sigma$ with the width of the impurity $\sigma$ obeying the equation of motion

$$m_a\ddot{\sigma} = -\frac{\partial V_{\mathrm{tot}}(\sigma)}{\partial\sigma}, \tag{4.9}$$

with potential $V_{\mathrm{tot}} = V_0 + V_{\mathrm{st}} + V_{\mathrm{inh}}$, given by

$$V_0(\sigma) = \frac{\hbar^2}{2m_a\sigma^2} + \frac{m_a}{2}\Omega_a^2\sigma^2\,, \tag{4.10a}$$

$$V_{\mathrm{st}}(\sigma) = -\frac{\eta^2 g}{\mathtt{D}(2\pi)^{\mathtt{D}/2}\sigma^{\mathtt{D}}}\,, \tag{4.10b}$$

$$V_{\mathrm{inh}}(\sigma) = \eta\mu_b\left[\frac{2}{\mathtt{D}}\tilde{\Gamma}\left(\frac{\mathtt{D}}{2}, \frac{R^2}{\sigma^2}\right) - \frac{\sigma^2}{R^2}\tilde{\Gamma}\left(1+\frac{\mathtt{D}}{2}, \frac{R^2}{\sigma^2}\right)\right]. \tag{4.10c}$$

Here, $R = \sqrt{2\mu_b/m_b\Omega_b^2}$ is the Thomas-Fermi radius of the condensate and $\tilde{\Gamma}(m,z) = [\Gamma(m)]^{-1}\int_0^z \mathrm{d}x \mathrm{e}^{-x}x^{m-1}$ is the normalised lower incomplete gamma function.

The two terms in $V_0$ describe the quantum pressure and the harmonic trapping of the impurity. The contribution $V_{\mathrm{st}}$ is second order in $\eta$ and represents self-trapping effects, which result from the deformation of the condensate due to its interaction

with the impurity. The inhomogeneity in the condensate density due to its trapping gives rise to the potential $V_{\rm inh}$, which is first order in $\eta$. It has a simple form despite its complicated mathematical expression. The expression in the square brackets of Eq. (4.10c) is a monotonically decreasing function in $\sigma/R$, taking the value $2/\mathtt{D}$ at $\sigma/R = 0$ and asymptotically approaching zero as $\sigma/R \to \infty$. Its gradient is largest for widths $\sigma \sim R$.

We see that Eq. (4.9), governing the evolution of the width $\sigma$, has the form of Newton's second law for the motion of a fictitious particle with position $\sigma$ and mass $m_a$ in the potential $V_{\rm tot}(\sigma)$. This analogy allows us to understand the behaviour of the width $\sigma$ as the impurity undergoes the procedure from the experiment by Catani *et al.* (described in Sec. 4.1). Translating this procedure into the language of our model, we assume the impurity initially feels a tight trap with frequency $\Omega_1$ and the coupled impurity-Bose gas system is cooled to the ground state (although in the experiment such low temperatures were not achieved). At time $t = 0$ the trap frequency is reduced to $\Omega_2$ and the subsequent evolution of the width for $t \geq 0$ is observed. In our model this corresponds to the width of the impurity sitting at the minimum of $V_{\rm tot}$ with frequency $\Omega_a = \Omega_1$ for times $t < 0$. Then for $t \geq 0$ the width moves in the potential $V_{\rm tot}$ with a lower trapping frequency $\Omega_a = \Omega_2$.

Our approach with $V_{\rm tot} = V_0$ exactly captures the correct evolution of the impurity wavefunction for zero impurity-boson coupling. The square of the width oscillates harmonically with frequency $2\Omega_2$ according to

$$\sigma^2(t) = \sigma_+^2 \sin^2(\Omega_2 t) + \sigma_-^2 \cos^2(\Omega_2 t), \tag{4.11}$$

where $\sigma_\pm^2 = w_a^2(\Omega_1/\Omega_2)^{\pm 1}$ and $w_a = \sqrt{\hbar/m_a\Omega_2}$ is the width of the harmonic oscillator ground state. Note that this harmonic solution exists for $\sigma^2$ despite its non-linear equation of motion.

### Homogeneous condensate

For non-zero impurity-boson coupling we first consider the special case of a homogeneous condensate whose decoupled density $n_0$ is uniform, corresponding to the regime of a flat trap $R \gg \sigma$. In this case the potential $V_{\rm inh}$ is also uniform and can be neglected, leaving the potential $V_{\rm tot}$ symmetric with respect to attractive and repulsive interactions.

The self-trapping $V_{\rm st}$ imparts a force on the width of the impurity towards smaller values of $\sigma$, shown in Fig. 4.1(a), and this has several effects. First the minimum of the potential $V_{\rm tot}$ is shifted to smaller widths. Second it shortens the distance between the minimum of $V_{\rm tot}$ for $t < 0$ and its minimum for $t \geq 0$. Third the curvature of potential $V_{\rm tot}$ near its minimum is increased.

These effects lead to an increased localisation of the impurity in the ground state at times $t < 0$. Further at times $t \geq 0$ the amplitudes of the breathing oscillations are reduced, along with an increase in the frequency of the oscillations. This is shown in Figs. 4.1(b) and 4.1(c), where we have plotted the maximum and minimum widths, and the frequency of the breathing oscillations, respectively. The frequencies $\omega$ are found from the best fitting curve of the form $\sigma^2(t)/w_a^2 = \kappa_1 \cos(\omega t) + \kappa_2$ to the oscillations, where $\kappa_1$ and $\kappa_2$ are fitting parameters.

### Condensate in a harmonic trap

In the general case the decoupled condensate density $n_0$ depends on the radial position $r$ and the potential $V_{\rm tot}$ is augmented by $V_{\rm inh}$. The total potential $V_{\rm tot}$ resulting from a trapped condensate is plotted in Fig. 4.2(a). Since the contribution of $V_{\rm inh}$ is first order in $\eta$ the symmetry between attractive and repulsive impurity-boson coupling is broken.

For repulsive coupling the force imparted by $V_{\rm inh}$ acts to increase the width of

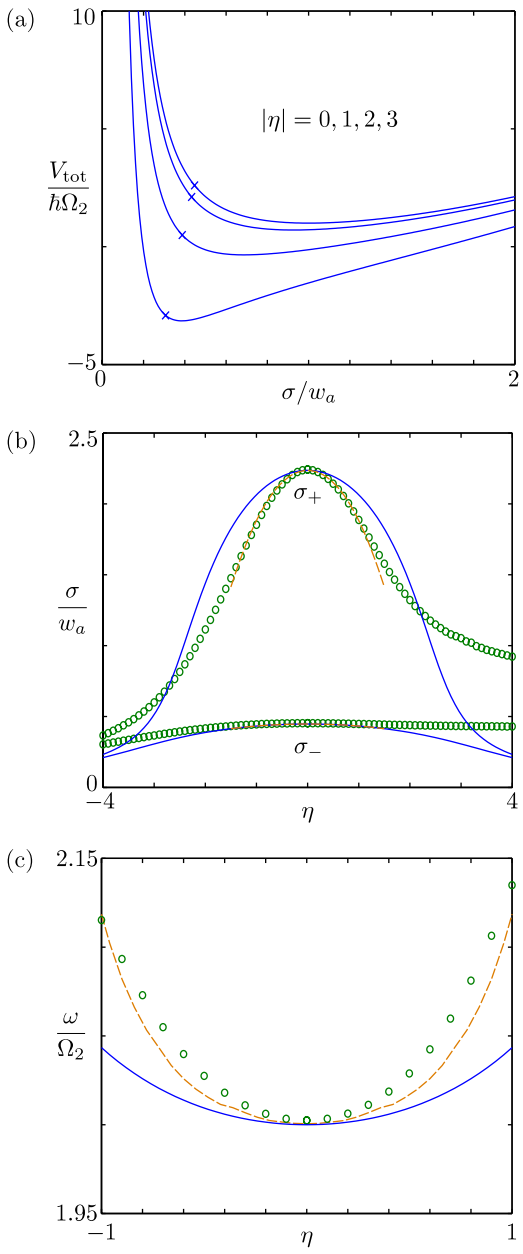


**Figure 4.1**. One-dimensional homogeneous Bose gas. (a) The potential $V_{\rm tot}$ governing the evolution of the width $\sigma$ of an impurity trapped with frequency $\Omega_2$ for various interactions strengths $\eta$. The crosses mark the width $\sigma$ at the minimum of $V_{\rm tot}$ for tight trapping with frequency $\Omega_1 = 5\Omega_2$. (b) The maximum $\sigma_+$ and minimum $\sigma_-$ widths during breathing oscillations as a function of $\eta$. The analytical results of the GP approach predict that self-trapping strongly suppresses breathing oscillations (full blue lines). Additional suppression occurs in the Bogoliubov approach due to dissipation into phonon modes (dashed orange lines), matching the numerical results (green dots). (c) The frequency of the breathing oscillations $\omega$ as a function of $\eta$. The parameters are set to $\mathtt{D} = 1$, $m_b = 2m_a$, $n_0 = \sqrt{m_b\Omega_2/\hbar}$ and $g = \hbar\Omega_2\sqrt{\hbar/m_b\Omega_2}$.

the impurity, opposing self-trapping. Specifically, in the regime $\eta \lesssim 1$ the inhomogeneity of the Bose gas flattens the potential $V_{\rm tot}$ near its minimum. In contrast, for sufficiently large $\eta$ the self-trapping term dominates and the potential develops a narrow trough at a small width, as for the homogenous condensate. If interactions are attractive $V_{\rm inh}$ represents a force on the width of the impurity towards smaller values of $\sigma$ and enhances the self-trapping seen for the homogeneous condensate.

The effects of $V_{\rm inh}$ on the dynamics of the impurity following a sudden decrease in $\Omega_a$ are captured in Figs. 4.2(b) and 4.2(c). For weak repulsive interactions $\eta \lesssim 1$, the effect of the flattening of the potential is to increase the amplitude of oscillations and decrease the frequency. For strong repulsive interactions $\eta \gg 1$, the creation of a second narrow minimum in the potential $V_{\rm tot}$ results in small amplitude and high frequency oscillations. Our results predict a sudden transition between these two regimes, causing a sharp drop in the amplitude of oscillations and an increase in frequency. For attractive impurities the effect of $V_{\rm inh}$ is to enhance the effects seen for a homogeneous condensate.

### 4.2.3 Damping and memory effects

Up until now we have not included a mechanism for the transfer of energy between the impurity and the Bose gas. We now show, for a homogeneous Bose gas, that the oscillations of the impurity create Bogoliubov phonons and thereby lead to the exchange of energy with the Bose gas. We work in a D-dimensional volume $\mathcal{V}$ with periodic boundaries.

To describe the interaction of the impurity with the elementary excitations of the Bose gas we take a similar approach to that in Sec. 2.1.2. We expand $\hat{\phi} = \varphi_0 + \delta\hat{\phi}$, where the decoupled mode $\varphi_0 = \sqrt{n_0}$ introduced previously describes the Bose gas in the ground state if $\eta = 0$. Therefore, in this picture, interactions with the impurity

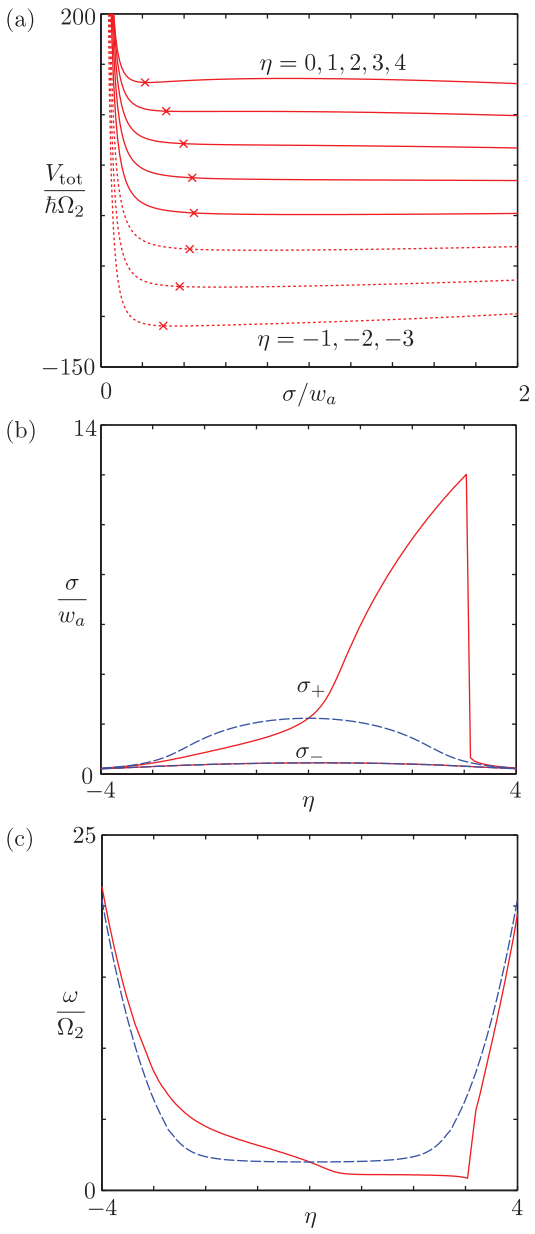


**Figure 4.2**. One-dimensional Bose gas in a harmonic trap. (a) The potential $V_{\rm tot}$ governing the evolution of the width $\sigma$ of an impurity trapped with frequency $\Omega_2$ for positive (full lines) and negative (dotted lines) interaction strengths $\eta$. The crosses mark the width $\sigma$ at the minimum of $V_{\rm tot}$ for tight trapping with frequency $\Omega_1 = 5\Omega_2$. (b) The maximum $\sigma_+$ and minimum $\sigma_-$ widths during breathing oscillations as a function of $\eta$ for a trapped (full red lines) and homogeneous (dashed blue lines) Bose gas. For a trapped Bose gas, breathing oscillations increase in amplitude with $\eta$ and at a critical value, here $\eta \approx 3.1$, undergo a sudden quench. (c) The frequency of the breathing oscillations $\omega$ as a function of $\eta$. The parameters are set to $\mathtt{D} = 1$, $m_b = 2m_a$, $N_0 = 200$, $\Omega_b = \Omega_2/\sqrt{2}$ and $g = \hbar\Omega_2\sqrt{\hbar/m_b\Omega_2}$.

will generate excitations beyond this condensate mode even in the ground state of the impurity-Bose gas system. Terms in $\hat{H}$ higher than second order in either the deformation $\delta\hat{\phi}$ or the coupling parameter $\eta$ are neglected [24]. Then expressing the deformation $\delta\hat{\phi} = \sum_{\mathbf{k}}[u_{\mathbf{k}}(\mathbf{r})\hat{d}_{\mathbf{k}} + v^*_{\mathbf{k}}(\mathbf{r})\hat{d}^\dagger_{\mathbf{k}}]$ in terms of Bogoliubov modes $\hat{d}_{\mathbf{k}}$ with momenta $\hbar\mathbf{k}$, the Hamiltonian of the Bose gas reads [71]

$$\hat{H}_b - \mu_b\hat{N} = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}}\hat{d}^\dagger_{\mathbf{k}}\hat{d}_{\mathbf{k}}\,, \tag{4.12}$$

up to a constant. Here $\hat{N}$ is the boson total number operator, and $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ solve the Bogoliubov-de Gennes Eqs. (2.16) with the phonon energies $\hbar\omega_{\mathbf{k}} = \sqrt{\varepsilon_{\mathbf{k}}\left(\varepsilon_{\mathbf{k}} + 2gn_0\right)}$ and single-particle energies $\varepsilon_{\mathbf{k}} = \hbar^2k^2/2m_b$. Under the same approximations the interaction Hamiltonian simplifies to

$$\hat{H}_{ab} = \eta g n_0 + \eta g \sum_{\mathbf{k}\neq 0}(\hat{f}^\dagger_{\mathbf{k}}\hat{d}^\dagger_{\mathbf{k}} + \hat{f}_{\mathbf{k}}\hat{d}_{\mathbf{k}})\,, \tag{4.13}$$

with the coefficients

$$\hat{f}^\dagger_{\mathbf{k}} = \sqrt{\frac{n_0\varepsilon_{\mathbf{k}}}{\mathcal{V}\hbar\omega_{\mathbf{k}}}}\int \mathrm{d}\mathbf{r}\hat{\varkappa}^\dagger(\mathbf{r})\hat{\varkappa}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}\,. \tag{4.14}$$

The first and the second terms in Eq. (4.13) represent the interaction of the impurity with the decoupled condensate mode $\varphi_0$ and the Bogoliubov phonons $\hat{d}_{\mathbf{k}}$, respectively. If we assume a product state for the impurity-Bose gas system, as we will, this interaction describes the creation of Bogoliubov phonons due to a classical driving force $f_{\mathbf{k}} = \sqrt{n_0\varepsilon_{\mathbf{k}}/\mathcal{V}\hbar\omega_{\mathbf{k}}}\int \mathrm{d}\mathbf{r}|\varkappa|^2\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$. The force $f_{\mathbf{k}}$ depends parametrically on the oscillating width $\sigma$ and is thus approximately periodic.

We solve for both the ground state and the evolution of the total system variationally within the ansatz $|\,\Psi\rangle = |\,\sigma,\gamma\rangle\otimes|\,\{\alpha_{\mathbf{k}}\}\rangle$. For the impurity, $|\,\sigma,\gamma\rangle$ corresponds

to the Gaussian ansatz in Eq. (4.7). The Bose gas is restricted to a product of coherent states $|\{\alpha_{\mathbf{k}}\}\rangle = \bigotimes_{\mathbf{k}} |\alpha_{\mathbf{k}}\rangle$, where the coherent state for each phonon mode is defined as $|\alpha_{\mathbf{k}}\rangle = \mathrm{e}^{-|\alpha_{\mathbf{k}}|^2/2}\mathrm{e}^{\alpha_{\mathbf{k}}\hat{d}^\dagger_{\mathbf{k}}}|\,\mathrm{vac},\mathbf{k}\rangle$. Each state $|\,\mathrm{vac},\mathbf{k}\rangle$ is the vacuum of the Bogoliubov operator $\hat{d}_{\mathbf{k}}$. The coherent product ansatz above is known to exactly describe phonons coupled to a classical field such as a density and approximately describe Bogoliubov phonons coupled to an impurity [24].

We derive equations of motion for the total system by extremising the action $S = \int \mathrm{d}t\, L$ with respect to the ansatz parameters (treating each $\alpha_{\mathbf{k}}$ and its complex conjugate as independent parameters), where

$$L = \langle\Psi|\left(\mathrm{i}\hbar\frac{\partial}{\partial t} - \hat{H} + \mu_b\hat{N}\right)|\Psi\rangle\,, \tag{4.15}$$

is the Lagrangian. Working in the thermodynamic limit $\mathcal{V}^{-1}\sum_{\mathbf{k}} \to 1/(2\pi)^{\mathrm{D}}\int \mathrm{d}\mathbf{k}$ we evaluate the integrals analytically using the small-momentum approximation to the Bogoliubov dispersion relation $\omega_{\mathbf{k}} = ck$, with the speed of sound given by $c = \sqrt{gn_0/m_b}$. This approximation is accurate in the regime $k \ll m_b c/\hbar$. We found no qualitative differences when proceeding numerically without making this approximation (though we do not present those results in this thesis).

Assuming the system is in the ground state at $t < 0$ and applying the methods described in the preceding paragraphs results in a Newtonian equation of motion describing the width of the impurity $\sigma$ moving in the potential $V_0$ augmented by a time-dependent force $F_{\mathrm{phon}}$. The details of the derivation are presented in App. 4.A.2, where we show that the force representing the interaction of the impurity with the

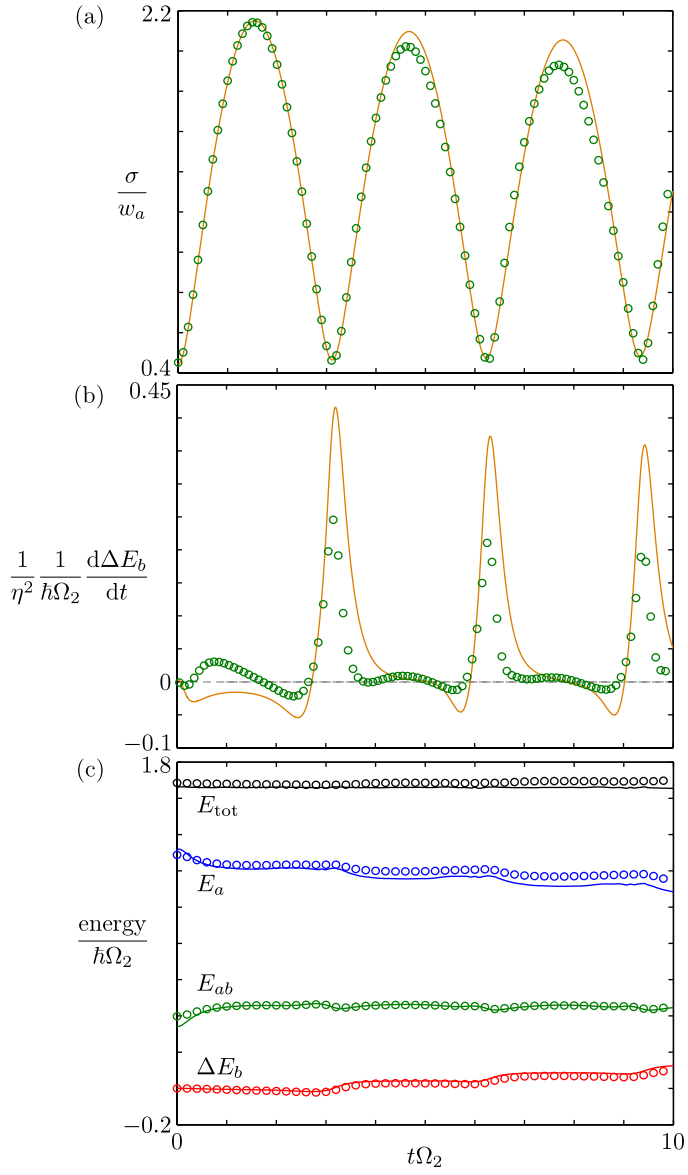


**Figure 4.3**. Damping of oscillations due to loss of energy, for weak coupling. (a) Damped oscillations of the width $\sigma$ of an impurity in a trap with frequency $\Omega_2$ obtained analytically (full orange) and numerically (green dots). The trap frequency for $t < 0$ is $\Omega_1 = 5\Omega_2$. (b) The rate of energy flow $\mathrm{d}\Delta E_b/\mathrm{d}t$ into the Bose gas. Positive (negative) values correspond to energy gained (lost) by the Bose gas. The dashed line marks $\mathrm{d}\Delta E_b/\mathrm{d}t = 0$. (c) The energy of the impurity $E_a$ (blue), interaction $E_{ab}$ (green) and Bose gas $\Delta E_b$ (red), together with their sum $E_{\mathrm{tot}}$ (black), as a function of time, obtained analytically (full line) and numerically (dots). The parameters are set to $\mathtt{D} = 1$, $\eta = 0.5$, $m_b = 2m_a$, $n_0 = \sqrt{m_b\Omega_2/\hbar}$ and $g = \hbar\Omega_2\sqrt{\hbar/m_b\Omega_2}$.

phonon bath, for $\mathtt{D} = 1$, is given by

$$F_{\rm phon}(t) = -\frac{\eta^2 g^2 n_0 \sigma(t)}{m_b c^2 \Gamma(3/2)} \Bigg\{ 2c \int_0^t \mathrm{d}t' \frac{c(t-t')\left[3\Sigma^2(t,t') - 2c^2(t-t')^2\right] \mathrm{e}^{-\frac{c^2(t-t')^2}{\Sigma^2(t,t')}}}{\Sigma^7(t,t')} + \frac{\left[\Sigma^2(t,0) - 2c^2t^2\right] \mathrm{e}^{-\frac{c^2t^2}{\Sigma^2(t,0)}}}{\Sigma^5(t,0)} \Bigg\}, \tag{4.16}$$

with compound width

$$\Sigma(t,t') = \left[\sigma^2(t) + \sigma^2(t')\right]^{1/2}. \tag{4.17}$$

The first and second terms in Eq. (4.16) correspond to the interaction of the impurity with phonons created during the evolution and those present in the initial state, respectively. These interactions introduce non-Markovian effects with a memory that decays exponentially on a timescale $\sigma/c$. As a consequence, an experimentalist can in principle control the degree of non-Markovianity by altering $g$, $n_0$ and $w_a$.

Due to the form of $F_{\rm phon}$ the equation of motion for the width of the impurity $\sigma$ is an integro-differential equation, which we solve using an iterative procedure. We initially evaluate $F_{\rm phon}$ assuming a trajectory for $\sigma$ and integrate the equations of motion. Subsequently, we use the resulting trajectory to evaluate $F_{\rm phon}$ in the next iteration. This procedure converges after a few iterations to a self-consistent evolution for $\sigma$.

The evolution of the width of the impurity after a sudden decrease in $\Omega_a$ is plotted in Fig. 4.3(a). We find that the breathing oscillations are damped out due to interactions with phonons. In Figs. 4.1(b) and 4.1(c) we plot the maximum and minimum widths, and frequencies as functions of the coupling parameter $\eta$. The oscillation frequency $\omega$ is found from the best fitting curve $\sigma^2(t)/w_a^2 = \kappa_1 \mathrm{e}^{-\kappa_3 t}\cos(\omega t) + \kappa_4 t + \kappa_2$ to the data, where $\kappa_1$, $\kappa_2$, $\kappa_3$ and $\kappa_4$ are fitting parameters. Dissipation before the first

peak leads to a stronger suppression of maximum width than found when considering self-trapping only. We find that dissipation also results in a greater dependence of frequency on coupling.

In App. 4.A.3, using the same approximations as for the equation of motion, we analytically calculate the energy deposited into the Bose gas

$$\Delta E_b(t) = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}}(|\alpha_{\mathbf{k}}(t)|^2 - |\alpha_{\mathbf{k}}(0)|^2). \tag{4.18}$$

The rate of energy lost by the impurity is

$$\frac{\mathrm{d}\Delta E_b(t)}{\mathrm{d}t} = \frac{\eta^2 g^2 n_0}{m_b c\Gamma(3/2)}\Bigg\{ c\int_0^t \mathrm{d}t' \frac{[\Sigma^2(t,t') - 2c^2(t-t')^2]\,\mathrm{e}^{-c^2(t-t')^2/\Sigma^2(t,t')}}{\Sigma^5(t,t')} - \frac{ct\mathrm{e}^{-c^2t^2/\Sigma^2(t,0)}}{\Sigma^3(t,0)}\Bigg\}, \tag{4.19}$$

which again is divided into two parts representing interactions with phonons created at $t \geq 0$ and $t < 0$, respectively. In Fig. 4.3(b) we plot this dissipation rate during the evolution shown in Fig. 4.3(a). We find that most dissipation occurs when the width of the impurity is near its minimum value and that $\mathrm{d}\Delta E_b(t)/\mathrm{d}t$ can take negative values (the impurity can gain energy). It may at first appear at odds with the Landau criterion that dissipation is largest when the impurity density distribution is stationary. However such arguments do not apply here since the impurity is not simply moving at a constant velocity. The form of $\mathrm{d}\Delta E_b(t)/\mathrm{d}t$ also differs markedly from that predicted by modelling the impurity as a damped harmonic oscillator, as done in [14], whereby dissipation is due to a friction force.

We show that the energy gained by the bosons is indeed deposited by the impurity, by plotting the impurity energy $E_a = \langle\Psi(t)\,|\,\hat{H}_a\,|\,\Psi(t)\rangle$, interaction energy $E_{ab} = \langle\Psi(t)\,|\,\hat{H}_{ab}\,|\,\Psi(t)\rangle$, $\Delta E_b$ (expressions given in Eqs. (4.37) and (4.43), respectively, of App. 4.A.3) and the total energy $E_{\mathrm{tot}} = E_a + \Delta E_b + E_{ab}$ together as a

function of time. This is done in Fig. 4.3(c). We see that the interaction energy stays roughly constant in time and that decrements in the impurity energy coincide with increments in the bosonic energy. The small variations in the total energy are due to the discretisation used when numerically evaluating the integrals in the analytical expressions for energy.

### 4.2.4 Numerical results

To confirm our analytical approach we numerically solve the coupled time-dependent GP Eqs. (4.2) for a homogeneous Bose gas. The ground states are calculated using normalised gradient flow (see Sec. 3.1.1) and the time-evolved states are calculated by applying the time-splitting sine-spectral technique (see Sec. 3.1.2). We assume for numerical ease that the condensate is spherically symmetric and enclosed in an infinite spherical box with a radius $(M-1)a$ that is large compared to the width $\sigma$ and the healing length of the condensate. More precisely, we choose $(M-1)a = 40\sqrt{2}w_a$, and solve for $\varphi(r)$ and $\varkappa(r)$ on a spatial grid of $M = 10^4$ points using a time-step $\delta t = 10^{-3}\Omega_2^{-1}$.

To calculate the width of the impurity $\sigma$ at a given time, we find the best fitting parameters to $|\varkappa(r)| = \kappa_1 \exp(-r^2/2\sigma^2)$. Maximum and minimum widths are calculated from this data.

For small $|\eta|$ we find good agreement on the frequency of oscillations and the width at the first maximum while the analytics underestimates the heights of later peaks, as exemplified in Figs. 4.3(a). This leads to a close match of the numerically and analytically calculated maximum and minimum widths, and frequencies, as shown in Figs. 4.1(b) and 4.1(c). Also the numerics shown in these two figures demonstrate that the symmetry between attractive and repulsive interactions is broken for large coupling. An attractive impurity can infinitely enhance the condensate

density at the origin, while an infinitely repulsive impurity only depletes the density of the Bose liquid at its location to zero (the Moses effect) [165, 167].

Also, the qualitative features of our analytic results for the dissipation of energy into the Bose gas are confirmed by numerical simulations, as shown in Figs. 4.3(b) and (c). Their similarity is due to the close relationship between the quantised phononic excitations of a Bose gas and the dynamical collective excitations of a condensate (cf. Secs. 2.1.2 and 2.1.2). Quantitative deviations arise from two approximations. We have neglected high order phonon effects and the impurity density does not always have a Gaussian form. Finally, note that the agreement of our analytics, derived by breaking number conservation, and our numerics, which may be derived while ensuring number conservation, provides a demonstration of the similarity of the predictions made using either approach for weak interactions.

### 4.2.5 Discussion

Our results reproduce the main features of the experiment by Catani *et al.* [14] even though we assumed zero temperature. For example, our model naturally explains the decoupled frequency of oscillations and the ratio of maximum to minimum width, along with the basic aspects of damping and amplitude suppression arising from impurity-boson interactions. However our zero temperature calculations predict a much richer dependence of amplitude on coupling, including a sudden quench resulting from the interplay of self-trapping and the inhomogeneity induced by the Bose gas trapping. We also predict a dependence of oscillation frequency on the impurity-Bose gas coupling, indicating that the lack of a dependence observed by Catani *et al.* is a finite temperature effect. Our intuitive model explains the differences between repulsive and attractive impurities, and highlights the way in which the non-Markovianity of the system arises and how it may be controlled. This

complements another work, not presented in this thesis, in which we study the transition between Markovian and non-Markovianity dynamics for a static impurity in a phonon bath [7].

In general, we expect the zero temperature approximation to be valid if the thermal energy of the initial state $k_{\mathrm{B}}T$ is small compared to the relevant energy scales of the system at $t < 0$. Let us now see how large these energy scales may be whilst keeping the typical impurity width on the scale of micrometers $w_a = 1\ \mu$m and the oscillations visible $\Omega_1/\Omega_2 = 10$. One energy scale $\hbar\Omega_1$ is given by the initial frequency of the harmonic trap of the impurity. A second, the initial energy of the phonons (see App. 4.A.3), is given by $\eta^2 g/4\sqrt{2}\Gamma(3/2)\sigma(0)$ for $\mathtt{D} = 1$. To maximise the minimum of these two energy scales, we set them equal to each other for a typical relative interaction strength $\eta = 1$. Assuming that $\sigma(0) \approx w_a\sqrt{\Omega_2/\Omega_1}$ as we found in our analysis, this gives $\hbar\sqrt{\Omega_1\Omega_2} = \hbar\Omega_1/\sqrt{10} = g/4\sqrt{2}\Gamma(3/2)w_a$. Using the value $g = 2.36 \times 10^{-37}\ \mathrm{Jm}^{-1}$ from the Catani *et al.* experiment we find the two energy scales of interest to be approximately $\hbar\Omega_1 \approx g/4\sqrt{2}\Gamma(3/2)\sigma(0) \approx k_B \times 11$ nK. This could be increased by increasing the radial confinement of the tubes, and thus the effective one-dimensional interaction strength $g$, or developing higher resolution imaging techniques so that $w_a$ may be reduced.

In the experiment by Catani *et al.*, the thermal energy of the Bose gas and impurity prior to equilibration were 350 nK and 50 nK, respectively, thus our analysis is not strictly valid for the temperatures of the experiment. However, mechanisms governing the evolution of the impurity at zero temperature, such as self-trapping, might still be relevant up to the temperatures realised in the experiment and could explain, for instance, the reduction in oscillation amplitude with increasing $\eta$.

## 4.3 Subsequent work

Since the completion of the work presented in this chapter, three relevant theoretical treatments [168, 169, 170] of the same system have been published. All focus on the Catani *et al.* experiment and the quasi one-dimensional case. The first two [168, 169] are a pair by Bonart and Cugliandolo in which they describe the Bose gas as a Luttinger liquid and assume an initial state for the impurity that is only valid at high temperatures. The system is mapped onto coupled harmonic oscillators and solved analytically using a path integral approach. The other [170] by Peotta *et al.* places the impurity-Bose gas system in a lattice and uses the time-dependent density-matrix renormalisation group method (closely related to the time-evolving block decimation method of Sec. 3.2) to simulate the system numerically at zero temperature. The aim of this subsection is to briefly review these works and compare them against what we have done in this chapter.

We begin by looking at the treatment of Peotta *et al.* [170] since it is the most similar to our own. As mentioned above, they place the whole system in a lattice but ensure that the spacing between the sites and the average filling are small enough that the continuum limit is well approximated. Thus they essentially near-exactly simulate the same Hamiltonian, shown in Eq. (4.1), as we have considered here. Furthermore, their simulations are at zero temperature, as we have also assumed. The trapping of the bosons is shallow so that the Bose gas is approximately homogeneous.

Their simulation results confirm what we have obtained. For a weakly-interacting Bose gas and weak impurity-boson coupling their numerics reproduce the effects of self-trapping that we predicted, namely an increase in the frequency and decrease in the amplitude of the breathing oscillations due to self-trapping. However for a more strongly-interacting Bose gas they find that the frequency decreases with increased

impurity-boson coupling. We hypothesise that this is a polaronic effect associated with an increased effective mass for the impurity resulting from its dressing by the bosons [24]. Our product-state treatment must underestimate this effect for strongly-interacting bosons.

Peotta *et al.* also give an analytical treatment. They phenomenologically describe the bosons by a Luttinger liquid, which is equivalent to a bath of linearly dispersing phonons. For the impurity-Luttinger liquid interaction they replace the contact interaction $\delta(r - r')$, where $r$ and $r'$ are the positions of two bosons, by some short-range interaction proportional to $k_c/\pi(k_c^2(r - r')^2 + 1)$, which is physically very similar for a large cut-off wave-vector $k_c$. So long as the impurity moves slowly they obtain a quantum Langevin equation with Abraham-Lorentz friction, which predicts an oscillation frequency that decreases with increasing impurity-boson coupling. Notably this Langevin equation does not predict the increase in frequency seen by their simulations and derived by us for weakly-interacting bosons. Further, their analytic treatment fails to capture the asymmetry between attractive and repulsive impurity-Luttinger liquid coupling, something which again they witnessed in their simulations and we found in ours.

Their conclusion is that their Luttinger liquid description of the bath only holds in a small part of parameter space. Similarly, the conclusion we may draw from their simulations is that our treatment also holds in a small and mutually exclusive part of parameter space. While their analytical treatment underestimates self-trapping, our description of the impurity and the Bose gas does not correctly account for the renormalisation of the motion of the impurity.

Moving on to the work of Bonart and Cugliandolo, they again describe the bosons phenomenologically by a Luttinger liquid with the same short-range impurity-Luttinger liquid interaction. The difference between their work and the analytical treatment of Peotta *et al.* is in the initial state of the system. The initial state consid-

ered by Bonart and Cugliandolo is a finite-temperature Gibbs state, rather than the ground state that we and Peotta *et al.* considered. In fact, Bonart and Cugliandolo do not quite consider the Gibbs state, rather the state that would be obtained if the Gibbs state is subjected to a position measurement (the result of which is unknown) of the impurity. They argue this is a valid approximation for the temperatures of the Catani *et al.* experiment. This specially chosen initial state, and the simple form of the Luttinger liquid and impurity-boson interaction then allows them to map the whole system into a set of coupled harmonic oscillators whose evolution is solved analytically using path integral methods. Treating the Luttinger liquid as homogeneous they, in Ref. [168], find oscillations in the width of the impurity that qualitatively match those in the experiment (the Luttinger parameter and cut-off wave-vector are used as fitting parameters).

In a related work [169] the same authors take into account the inhomogeneity of the Luttinger liquid, using semi-classical arguments to justify a renormalisation of the mass and trapping of the impurity for small times. This improves their fit with the experimental results, although they must interpolate in an *ad hoc* manner between short and long times. Unfortunately, as with the experiment of Catani *et al.*, they do not consider attractive impurity-boson coupling.

## 4.A Derivation steps

### 4.A.1 Equations of motion in the GP treatment

Here we evaluate the Lagrangian $L = \int \mathrm{d}\mathbf{r}\mathcal{L}$ for the Lagrangian density $\mathcal{L}$ given in Eq. (4.8). Integrating each term in $\mathcal{L}$ in turn we find that $L$ comprises the following

terms

$$\int \mathrm{d}\mathbf{r}\varkappa^* \mathrm{i}\hbar\frac{\partial}{\partial t}\varkappa = \frac{\mathtt{D}\hbar\dot{\gamma}\sigma^2}{2}, \tag{4.20a}$$

$$\int \mathrm{d}\mathbf{r}\varkappa^* \frac{\hbar^2\nabla^2}{2m_a}\varkappa = -\frac{\mathtt{D}\hbar}{4m_a\sigma^2} - \frac{\mathtt{D}\hbar^2\gamma^2\sigma^2}{m_a}, \tag{4.20b}$$

$$\int \mathrm{d}\mathbf{r}\varkappa^* \left(-\tfrac{1}{2}m_a\Omega_a^2 r^2\right)\varkappa = -\frac{\mathtt{D}\hbar}{4m_a\Omega_a^2\sigma^2}, \tag{4.20c}$$

$$\int \mathrm{d}\mathbf{r}\varkappa^* \left(-\eta g n_0\right)\varkappa = \frac{\eta\mu_b\mathtt{D}}{2}\left[\frac{\sigma^2}{R^2}\tilde{\Gamma}\left(1+\frac{\mathtt{D}}{2},\frac{R^2}{\sigma^2}\right) - \frac{2}{\mathtt{D}}\tilde{\Gamma}\left(\frac{\mathtt{D}}{2},\frac{R^2}{\sigma^2}\right)\right], \tag{4.20d}$$

$$\int \mathrm{d}\mathbf{r}\varkappa^* \left(\tfrac{1}{2}\eta^2 g|\varkappa|^2\right)\varkappa = \frac{\eta^2 g}{2(2\pi)^{\mathtt{D}}\sigma^{\mathtt{D}}}. \tag{4.20e}$$

Equations (4.20a–c) and (4.20e) follow simply from the ansatz of Eq. (4.7). To obtain Eq. (4.20d) we have used the Thomas-Fermi expression for the density in Eq. (4.3), the simplification in Eq. (4.6) for spherically symmetric functions and the expression for the chemical potential given in Eq. (4.5). Summing the terms in Eqs. (4.20) gives the Lagrangian $L$.

The next step is to extremise the action $S = \int \mathrm{d}t\, L$. It is well-known that this is equivalent to solving the Euler-Lagrange (EL) equations

$$\frac{\partial L}{\partial p} = \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{p}}, \qquad p = \sigma, \gamma. \tag{4.21}$$

The EL equations for $\sigma$ and $\gamma$, respectively, are

$$\hbar\sigma\dot{\gamma} - \frac{2\hbar^2\gamma^2\sigma}{m_a} - \frac{1}{2}\frac{\partial}{\partial\sigma}V_{\mathrm{tot}}(\sigma) = 0, \tag{4.22a}$$

$$\gamma = -m_a\dot{\sigma}/2\hbar\sigma, \tag{4.22b}$$

where the terms making up $V_{\mathrm{tot}}(\sigma)$ are given in Eq. (4.10). Inserting Eq. (4.22b) into Eq. (4.22a) and performing some cancellation results in the equation of motion given in Eq. (4.9).

### 4.A.2 Equations of motion in the Bogoliubov treatment

The first step is to evaluate the Lagrangian of Eq. (4.15) by substituting in the Hamiltonian (see Eqs. (4.1a), (4.12) and (4.13)). The expected value of the Hamiltonian is

$$\langle \Psi \,|\, (\hat{H} - \mu_b \hat{N}) \,|\, \Psi \rangle = \langle \sigma, \gamma \,|\, \hat{H}_a \,|\, \sigma, \gamma \rangle + \langle \{\alpha_{\mathbf{k}}\} \,|\, (\hat{H}_b - \mu_b \hat{N}) \,|\, \{\alpha_{\mathbf{k}}\} \rangle + \langle \Psi \,|\, \hat{H}_{ab} \,|\, \Psi \rangle \,, \tag{4.23}$$

where

$$\langle \sigma, \gamma \,|\, \hat{H}_a \,|\, \sigma, \gamma \rangle = \frac{\mathtt{D}\hbar^2}{4 m_a \sigma^2} + \frac{\mathtt{D}\hbar^2 \gamma^2 \sigma^2}{m_a} + \frac{\mathtt{D} m_a \Omega_a^2 \sigma^2}{4}, \tag{4.24a}$$

$$\langle \{\alpha_{\mathbf{k}}\} \,|\, (\hat{H}_b - \mu_b \hat{N}) \,|\, \{\alpha_{\mathbf{k}}\} \rangle = \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} |\alpha_{\mathbf{k}}|^2, \tag{4.24b}$$

$$\langle \Psi \,|\, \hat{H}_{ab} \,|\, \Psi \rangle = \eta g n_0 + \eta g \sum_{\mathbf{k}} (\alpha_{\mathbf{k}} + \alpha_{\mathbf{k}}^*) \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2 k^2/4}. \tag{4.24c}$$

This gives us half of the terms making up the Lagrangian. The others arise from

$$\begin{aligned} \langle \Psi \,|\, \mathrm{i}\hbar \frac{\partial}{\partial t} \,|\, \Psi \rangle = \mathrm{i}\hbar \Big( & \dot{\sigma} \, \langle \sigma, \gamma \,|\, \frac{\partial}{\partial \sigma} \,|\, \sigma, \gamma \rangle + \dot{\gamma} \, \langle \sigma, \gamma \,|\, \frac{\partial}{\partial \gamma} \,|\, \sigma, \gamma \rangle \\ & + \sum_{\mathbf{k}} \left[ \dot{\alpha}_{\mathbf{k}} \, \langle \alpha_{\mathbf{k}} \,|\, \frac{\partial}{\partial \alpha_{\mathbf{k}}} \,|\, \alpha_{\mathbf{k}} \rangle + \dot{\alpha}_{\mathbf{k}}^* \, \langle \alpha_{\mathbf{k}} \,|\, \frac{\partial}{\partial \alpha_{\mathbf{k}}^*} \,|\, \alpha_{\mathbf{k}} \rangle \right] \Big). \end{aligned} \tag{4.25}$$

Starting from Eq. (4.7), some integration and algebra leads to $\langle \sigma, \gamma \,|\, \partial/\partial\sigma \,|\, \sigma, \gamma \rangle = 0$ and $\langle \sigma, \gamma \,|\, \partial/\partial\gamma \,|\, \sigma, \gamma \rangle = -\mathrm{i}\mathtt{D}\sigma^2/2$. For the bosonic part we use the properties of coherent states $\hat{d} \,|\, \alpha_{\mathbf{k}} \rangle = \alpha_{\mathbf{k}} \,|\, \alpha_{\mathbf{k}} \rangle$ and $\langle\, \alpha_{\mathbf{k}} \,|\, \alpha_{\mathbf{k}'} \,\rangle = \exp\left(-\frac{1}{2}\left[|\alpha_{\mathbf{k}}|^2 + |\alpha_{\mathbf{k}'}|^2 - 2\alpha_{\mathbf{k}'}\alpha_{\mathbf{k}}^*\right]\right)$ together with $\langle \alpha_{\mathbf{k}} \,|\, \partial/\partial\alpha_{\mathbf{k}} \,|\, \alpha_{\mathbf{k}} \rangle = \partial/\partial\alpha_{\mathbf{k}'} \langle\, \alpha_{\mathbf{k}} \,|\, \alpha_{\mathbf{k}'} \,\rangle|_{\mathbf{k}'=\mathbf{k}}$ to obtain

$$\langle \alpha_{\mathbf{k}} \,|\, \frac{\partial}{\partial \alpha_{\mathbf{k}}} \,|\, \alpha_{\mathbf{k}} \rangle = \frac{\alpha_{\mathbf{k}}^*}{2}, \tag{4.26a}$$

$$\langle \alpha_{\mathbf{k}} \,|\, \frac{\partial}{\partial \alpha_{\mathbf{k}}^*} \,|\, \alpha_{\mathbf{k}} \rangle = -\frac{\alpha_{\mathbf{k}}}{2}. \tag{4.26b}$$

Combining these results gives the Lagrangian

$$\begin{aligned} L =& \frac{\hbar \mathtt{D} \sigma^2 \dot{\gamma}}{2} + \frac{\mathrm{i}\hbar}{2} \sum_{\mathbf{k}} [\dot{\alpha}_{\mathbf{k}} \alpha_{\mathbf{k}}^* - \alpha_{\mathbf{k}} \dot{\alpha}_{\mathbf{k}}^*] - \frac{\mathtt{D}\hbar^2}{4 m_a \sigma^2} - \frac{\mathtt{D}\hbar^2 \gamma^2 \sigma^2}{m_a} - \frac{\mathtt{D} m_a \Omega_a^2 \sigma^2}{4} \\ & - \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} |\alpha_{\mathbf{k}}|^2 - \eta g n_0 - \eta g \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2 k^2/4}. \end{aligned} \tag{4.27}$$

We next extremise the action $S = \int \mathrm{d}t\, L$ by solving the EL equations

$$\frac{\partial L}{\partial p} = \frac{\partial}{\partial t} \frac{\partial L}{\partial \dot{p}}, \qquad p = \sigma, \gamma, \alpha_{\mathbf{k}}, \alpha_{\mathbf{k}}^*. \tag{4.28}$$

The EL equations for $\sigma$, $\gamma$, $\alpha_{\mathbf{k}}$ and $\alpha_{\mathbf{k}}^*$ (in that order) are

$$\hbar \sigma \dot{\gamma} = -\frac{\hbar^2}{2 m_a \sigma^3} + \frac{2\hbar^2 \gamma^2 \sigma}{m_a} + \frac{m_a \Omega_a^2 \sigma}{2} - \frac{\eta g \sigma}{2\mathtt{D}} \sum_{\mathbf{k}} k^2 \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2 k^2/4} [\alpha_{\mathbf{k}} + \alpha_{\mathbf{k}}^*], \tag{4.29a}$$

$$\gamma = -m_a \dot{\sigma} / 2\hbar\sigma, \tag{4.29b}$$

$$\mathrm{i}\hbar \dot{\alpha}_{\mathbf{k}} = \hbar \omega_{\mathbf{k}} \alpha_{\mathbf{k}} + \eta g \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2 k^2/4}, \tag{4.29c}$$

$$-\mathrm{i}\hbar \dot{\alpha}_{\mathbf{k}}^* = \hbar \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^* + \eta g \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2 k^2/4}. \tag{4.29d}$$

By integrating we get $\alpha_{\mathbf{k}}(t)$ in terms of $\sigma$ and $\alpha_{\mathbf{k}}(0)$,

$$\alpha_{\mathbf{k}}(t) = \mathrm{e}^{-\mathrm{i}\omega_{\mathbf{k}} t} \left[ \alpha_{\mathbf{k}}(0) - \frac{\mathrm{i}\eta g}{\hbar} \sqrt{\frac{n_0 \varepsilon_{\mathbf{k}}}{\mathcal{V} \hbar \omega_{\mathbf{k}}}} \int_0^t \mathrm{d}t' \mathrm{e}^{\mathrm{i}\omega_{\mathbf{k}} t'} \mathrm{e}^{-\sigma^2(t') k^2/4} \right]. \tag{4.30}$$

Substituting this and Eq. (4.29b) into Eq. (4.29a) gives the equation of motion

$$m_a \ddot{\sigma}(t) = -\frac{\partial V_0(\sigma(t))}{\partial \sigma(t)} + F_{\mathrm{phon}}(t), \tag{4.31}$$

where $V_0$ is given in Eq. (4.10a) and the force due to phonons is

$$F_{\text{phon}}(t) = -\frac{2\eta^2 g^2 n_0 \sigma}{\mathtt{D}\mathcal{V}} \sum_{\mathbf{k}} \frac{k^2 \varepsilon_{\mathbf{k}}}{\hbar\omega_{\mathbf{k}}} \mathrm{e}^{-\sigma^2(t)k^2/4} \Bigg[ -\frac{\cos(\omega_{\mathbf{k}} t)\alpha_{\mathbf{k}}(0)}{\eta g} \sqrt{\frac{\mathcal{V}\hbar\omega_{\mathbf{k}}}{n_0\varepsilon_{\mathbf{k}}}} + \frac{1}{\hbar}\int_0^t \mathrm{d}t' \sin(\omega_{\mathbf{k}}(t-t'))\mathrm{e}^{-\sigma^2(t')k^2/4} \Bigg]. \tag{4.32}$$

We next specify that at $t = 0$ we are in the ground state of the Hamiltonian $\hat{H}$ with $\Omega_a = \Omega_1$. In the variational approach we find the parameters for this ground state by minimising the expected value $\langle \Psi \,|\, (\hat{H} - \mu_b \hat{N}) \,|\, \Psi \rangle$. It is verified through differentiation that this results in the EL Eqs. (4.29) with the time-derivatives set to zero, i.e., with some substitution we have initial conditions

$$-\frac{\mathtt{D}\hbar^2}{2m_a\sigma^3(0)} + \frac{\mathtt{D}m_a\Omega_1^2\sigma(0)}{2} - \frac{\eta g \sigma(0)}{2} \sum_{\mathbf{k}} k^2 \sqrt{\frac{n_0\varepsilon_{\mathbf{k}}}{\mathcal{V}\hbar\omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2(0)k^2/4} [\alpha_{\mathbf{k}}(0) + \alpha^*_{\mathbf{k}}(0)] = 0, \tag{4.33a}$$

$$\alpha_{\mathbf{k}}(0) = -\frac{\eta g}{\hbar\omega_{\mathbf{k}}} \sqrt{\frac{n_0\varepsilon_{\mathbf{k}}}{\mathcal{V}\hbar\omega_{\mathbf{k}}}} \mathrm{e}^{-\sigma^2(0)k^2/4}. \tag{4.33b}$$

Substituting Eq. (4.33b) into Eq. (4.33a) gives an equation for the initial width

$$-\frac{\hbar^2}{m_a\sigma^3(0)} + m_a\Omega_1^2\sigma(0) - \frac{2\eta^2 g^2 \sigma(0)}{\mathtt{D}} \sum_{\mathbf{k}} \frac{n_0\varepsilon_{\mathbf{k}} k^2}{\mathcal{V}(\hbar\omega_{\mathbf{k}})^2} \mathrm{e}^{-\sigma^2(0)k^2/2} = 0. \tag{4.34}$$

Also, substituting Eq. (4.33b) into Eq. (4.32) gives an expression for $F_{\text{phon}}$ purely in terms of $\sigma$

$$F_{\text{phon}}(t) = -\frac{2\eta^2 g^2 n_0 \sigma(t)}{\mathtt{D}\mathcal{V}} \sum_{\mathbf{k}} \frac{k^2 \varepsilon_{\mathbf{k}}}{\hbar\omega_{\mathbf{k}}} \mathrm{e}^{-\sigma^2(t)k^2/4} \Big[ \frac{\cos(\omega_{\mathbf{k}} t)}{\hbar\omega_{\mathbf{k}}} \mathrm{e}^{-\sigma^2(0)k^2/4} + \frac{1}{\hbar}\int_0^t \mathrm{d}t' \sin(\omega_{\mathbf{k}}(t-t'))\mathrm{e}^{-\sigma^2(t')k^2/4} \Big]. \tag{4.35}$$

The final step is to assume a linear spectrum $\omega_{\mathbf{k}} = ck$, $\mathtt{D} = 1$ and consider the

thermodynamic limit by making the replacement $\mathcal{V}^{-1}\sum_{\mathbf{k}} \to 1/(2\pi)^{\mathtt{D}}\int \mathrm{d}\mathbf{k}$. It is then just a matter of integrating Eq. (4.35) to show that $F_{\mathrm{phon}}$ is equal to the expression given in Eq. (4.16). Additionally, following from the integration of Eq. (4.34), the initial width is given by

$$-\frac{\hbar^2}{m_a\sigma^3(0)} + m_a\Omega_1^2\sigma(0) - \frac{\eta^2 g^2 n_0}{2\sqrt{2}m_b c^2\Gamma(3/2)\sigma^2(0)} = 0. \tag{4.36}$$

### 4.A.3 Energies

From the details of the variational treatment presented in the previous two subsections we may also derive expressions for the impurity, bosonic and interaction energies in terms of $\sigma$. The simplest is the impurity energy

$$E_a(t) = \langle\Psi(t)\,|\,\hat{H}_a\,|\,\Psi(t)\rangle = \frac{\mathtt{D}\hbar^2}{4m_a\sigma^2(t)} + \frac{\mathtt{D}m_a\dot{\sigma}^2(t)}{4} + \frac{\mathtt{D}m_a\Omega_a^2\sigma^2(t)}{4}, \tag{4.37}$$

which follows from Eqs. (4.24a) and (4.29b). The three terms correspond to the quantum pressure, kinetic energy and potential energy, respectively.

For the initial bosonic energy, combining Eqs. (4.24a) and (4.33b) gives

$$\begin{aligned} E_b(0) &= \langle\Psi(0)\,|\,(\hat{H}_b - \mu_b\hat{N})\,|\,\Psi(0)\rangle \\ &= \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}}|\alpha_{\mathbf{k}}(0)|^2 \\ &= \frac{\eta^2 g^2 n_0}{\mathcal{V}}\sum_{\mathbf{k}} \frac{\varepsilon_{\mathbf{k}}}{(\hbar\omega_{\mathbf{k}})^3}\mathrm{e}^{-\sigma^2(0)k^2/2}. \end{aligned} \tag{4.38}$$

For the subsequent change in bosonic energy

$$\Delta E_b(t) = \langle\Psi(t)\,|\,(\hat{H}_b - \mu_b\hat{N})\,|\,\Psi(t)\rangle - E_b(0), \tag{4.39}$$

Eqs. (4.24a), (4.30) and (4.33b) result in

$$\begin{aligned}\Delta E_b(t) =& \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} \left(|\alpha_{\mathbf{k}}(t)|^2 - |\alpha_{\mathbf{k}}(0)|^2\right), \\ =& \frac{\eta^2 g^2 n_0}{\hbar\mathcal{V}} \sum_{\mathbf{k}} \int_0^t \mathrm{d}t' \Bigg\{ \int_0^t \mathrm{d}t'' \frac{\varepsilon_{\mathbf{k}}}{\hbar} \mathrm{e}^{\mathrm{i}\omega_{\mathbf{k}}(t''-t')} \mathrm{e}^{-\Sigma^2(t',t'')k^2/2} \\ &+ \frac{2\varepsilon_{\mathbf{k}}}{\hbar\omega_{\mathbf{k}}} \sin(\omega_{\mathbf{k}} t') \mathrm{e}^{-\Sigma^2(t',0)k^2/4} \Bigg\}, \end{aligned} \tag{4.40}$$

where we have used the compound width defined in Eq. (4.17). By rearranging the integrals, we write $\Delta E_b(t) = \int_0^t \mathrm{d}t' \mathrm{d}\Delta E_b(t')/\mathrm{d}t'$, where

$$\begin{aligned}\frac{\mathrm{d}\Delta E_b(t)}{\mathrm{d}t} = \frac{2\eta^2 g^2 n_0}{\hbar\mathcal{V}} \sum_{\mathbf{k}} \Bigg\{ & \int_0^t \mathrm{d}t' \frac{\varepsilon_{\mathbf{k}}}{\hbar} \cos(\omega_{\mathbf{k}}(t'-t)) \mathrm{e}^{-\Sigma^2(t,t')k^2/2} \\ &+ \frac{\varepsilon_{\mathbf{k}}}{\hbar\omega_{\mathbf{k}}} \sin(\omega_{\mathbf{k}} t) \mathrm{e}^{-\Sigma^2(t,0)k^2/4} \Bigg\}. \end{aligned} \tag{4.41}$$

For the interaction energy $E_{ab}(t) = \langle \Psi(t) \,|\, \hat{H}_{ab} \,|\, \Psi(t) \rangle$, combining Eqs. (4.24c) and (4.30) gives

$$\begin{aligned}E_{ab}(t) = \eta g n_0 - \frac{2\eta^2 g^2 n_0}{\mathcal{V}} \sum_{\mathbf{k}} \frac{\varepsilon_{\mathbf{k}}}{(\hbar\omega_{\mathbf{k}})^2} \Bigg[ & \frac{1}{\hbar} \int_0^t \mathrm{d}t' \sin(\omega_{\mathbf{k}}(t'-t)) \mathrm{e}^{-\Sigma^2(t,t'')k^2/4} \\ &+ \cos(\omega_{\mathbf{k}} t) \mathrm{e}^{-\Sigma^2(t,0)k^2/4} \Bigg]. \end{aligned} \tag{4.42}$$

Assuming a linear dispersion relation and taking the thermodynamic limit the sum over $\mathbf{k}$ may be evaluated with $\mathtt{D} = 1$ to give $E_b(0) = \eta^2 g/4\sqrt{2}\Gamma(3/2)\sigma(0)$, Eq. (4.19) for $\mathrm{d}\Delta E_b(t)/\mathrm{d}t$ and

$$\begin{aligned}E_{ab}(t) = \eta g n_0 - \frac{\eta^2 g^2 n_0}{2 m_b c^2 \Gamma(3/2)} \Bigg\{ & 2c \int_0^t \mathrm{d}t' \frac{c(t-t') \mathrm{e}^{-c^2(t-t')^2/\Sigma^2(t,t')}}{\Sigma^3(t,t')} \\ &+ \frac{\mathrm{e}^{-c^2 t^2/\Sigma^2(t,0)}}{\Sigma(t,0)} \Bigg\}. \end{aligned} \tag{4.43}$$

### 4.A.4 Checking the derivations

There are many ways to check that our analytical and numerical calculations are sensible. Most of them involve verifying that nothing changes for $\Omega_1 = \Omega_2$, i.e., if we start in the ground state of the Hamiltonian and the Hamiltonian does not change then the state of the system should be stationary. For example, to verify aspects of our expression for $F_{\text{phon}}$ we reassure ourselves that $\sigma(t) = \sigma(0)$ is a solution of the equation of motion for $\sigma$ in the case $\Omega_1 = \Omega_2$. In this special case we evaluate $F_{\text{phon}}(t)$ and find it is equal to the third term in Eq. (4.36). Thus the right hand side of the equation of motion Eq. (4.31) equals the left hand side of Eq. (4.36) and so is zero at all times. This in turn means $\ddot{\sigma}(t) = 0$, demonstrating that $\sigma(t) = \sigma(0)$ is a consistent solution, as must be the case. Similar checks have been performed on the analytically derived energies, the code that takes analytical expressions and uses them to produce the plots appearing in this chapter, and the numerical simulations of the GP equations.

# Chapter 5

# Impurity transport through a Bose lattice gas

Following our investigation of the motion of a harmonically-trapped impurity in a Bose gas, we now move on to discuss the motion of an impurity in a Bose gas under different circumstances[1].

We will initially be interested in mimicking the transport of an electron in a metal [95]. Thus we subject our impurity, representing the electron, to a tilted optical lattice potential that represents the lattice potential together with an applied voltage. Immersing this system in a Bose gas and allowing collisions between the impurity and bosons represents electron-phonon relaxation processes. These give rise to a net impurity/electron current.

Theoretical treatments have shown that the current of an impurity in a tilted lattice interacting with a Bose gas exhibits both ohmic and negative differential conductance (NDC) [23, 130, 171]. NDC is the phenomenon of a current that decreases with increasing forcing, in this case tilt/voltage. Particles in tilted lattices also feature several fundamental quantum mechanical processes related to the non-equilibrium transport of particles, such as Landau-Zener tunnelling [172], Bloch

[1]This chapter presents results that we first demonstrated in Refs. [2, 3].

oscillations [173, 174] and processes analogous to photon-assisted tunnelling [175].

There are crucial advantages to a cold atom implementation over a solid state system [95, 58]. First, there are minimal imperfections in the system and so impurity-boson collisions, which can be fully controlled, are the only relaxation process [177, 178]. Second, large lattice tilts can be achieved [179]. As mentioned above this allows us to observe NDC [23, 130, 171], which is realised in a solid state system with difficulty by resorting to semiconductor superlattices [180, 181, 182]. Third, the number of impurity atoms in a bosonic background may be controlled down to a single impurity [17], allowing the study of both single-impurity and many-impurity phenomena as desired. Last, the parameters of our system can be chosen to ensure that Landau-Zener tunnelling to higher impurity bands — often a significant effect in high field transport [182] — is negligible despite the large lattice tilts [130]. This justifies the use of a tight binding framework.

In the present chapter we naturally extend the above mentioned works. Specifically we confine the bosons to a separate non-tilted optical potential. The initial motivation is that the spectrum of phononic excitations in this case is more akin to that in solids. In particular, there is a band edge. We begin in this vein with Sec. 5.1, analysing the impurity current when the Bose gas is described by a superfluid. We find that the change in the phonon spectrum due to the bosonic lattice causes clear alterations to the dependence of the impurity current on the tilt. Following this we change tack and consider how effective the impurity is at probing the bosonic system. In Sec. 5.2 we go beyond a weakly-interacting Bose gas and consider how well we can distinguish the various strongly- and weakly-interacting regimes of the Bose gas by observing the impurity dynamics.

## 5.1 Atomic currents and phonon resonances

We start by showing that an impurity-boson mixture in different species-specific optical lattices, with that of the impurity lattice tilted, is highly appropriate for exploring the effects of electron-phonon interactions on non-equilibrium electronic transport through a conductor. Specifically we demonstrate that the electron-phonon resonances predicted to exist in solids nearly 40 years ago [183, 181, 184], yet for which there seems to be no conclusive experimental evidence [185, 182], can be realised using cold atoms. They arise as the crystal momentum of the phonon emitted in the electron-phonon relaxation process approaches a so-called Van Hove singularity [186] at the upper edge of the phonon band.

Before going into the details, we introduce our system and the Hamiltonian describing it, including a basic description of an impurity moving in a periodic potential subject to an external forcing.

### 5.1.1 Impurities in a tilted lattice

We consider one or a few impurities confined to a tilted optical lattice that allows them to move along one spatial dimension only, depicted in Fig. 5.1(a). Such an optical lattice can be created for the impurities by subjecting them to both an optical lattice and a static forcing, e.g., gravity [13], or by chirping the frequency difference between the counter-propagating lasers making up the standing wave of the optical lattice along the direction of motion [179]. For sufficiently deep lattices and low temperatures solely the lowest Bloch band is occupied and only tunnelling between nearest-neighbour sites must be considered (see Sec. 2.3.1). The Hamiltonian of the impurities is of the form

$$\hat{H}_a = -J_a \sum_{\langle \ell\ell' \rangle} \hat{a}_\ell^\dagger \hat{a}_{\ell'} + \hbar\omega_B \sum_\ell \ell\, \hat{a}_\ell^\dagger \hat{a}_\ell \,, \tag{5.1}$$

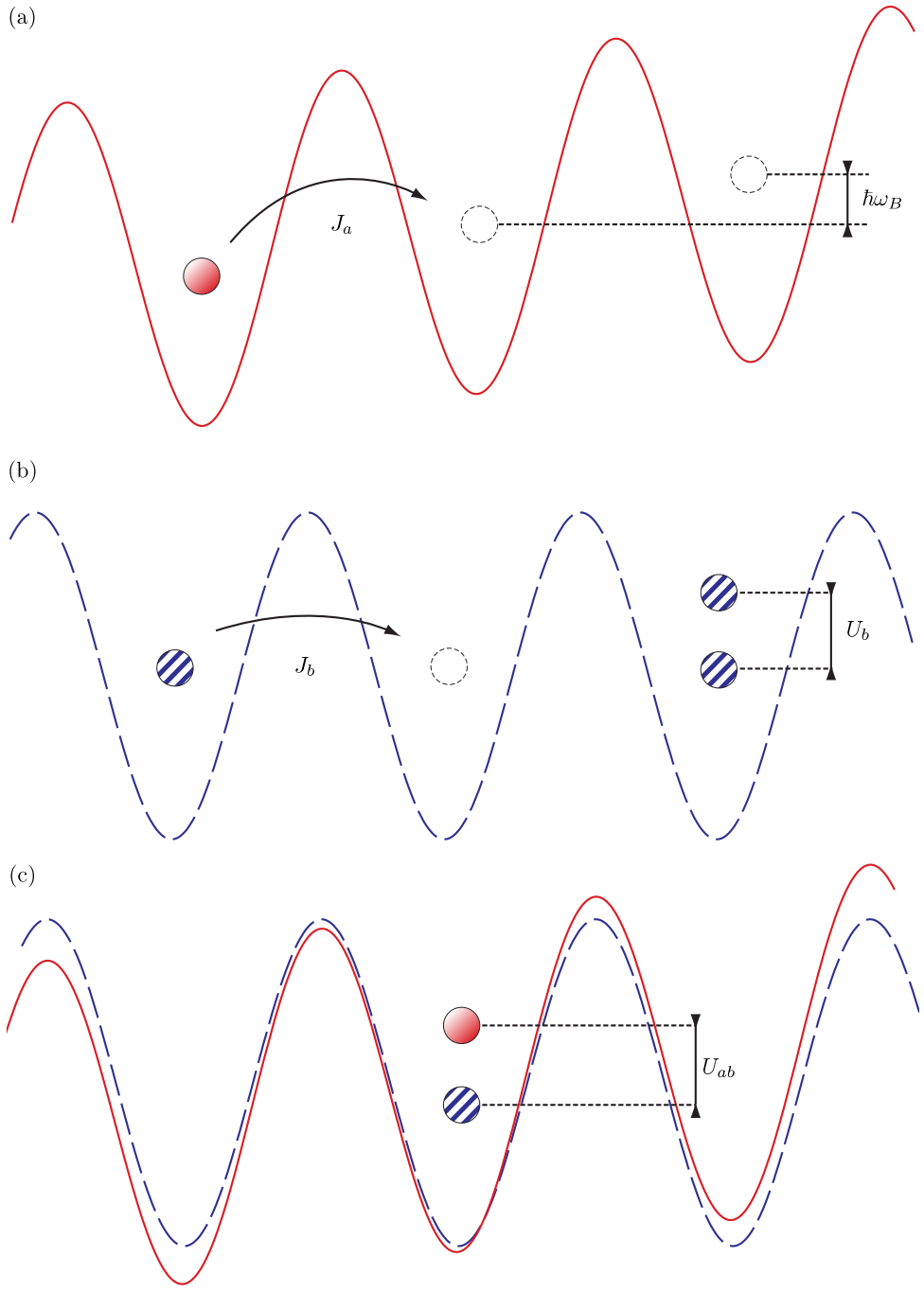


**Figure 5.1**. Schematic diagram for the (a) impurities, (b) Bose gas and (c) interaction Hamiltonians.

where $\langle \ell\ell' \rangle$ denotes the sum over nearest-neighbour sites $\ell$ and $\ell'$. Here $\hat{a}_\ell^\dagger$ ($\hat{a}_\ell$) is the creation (annihilation) operator for an impurity in a Wannier state localised at site $\ell$, separated in energy and distance from its neighbours by the Bloch energy $\hbar\omega_B$ and the lattice constant $a$, respectively. An additional condition for the above Hamiltonian to hold is that the separation of the Bloch bands must be much greater than the Bloch energy. Note that the impurities could be either bosonic or fermionic; we assume them to be sufficiently dilute so that impurity-impurity interactions and their quantum statistics have negligible effects.

Without a further interaction with an environment it is not possible for an impurity to dissipate energy and therefore alter its expected position. This is evident from the eigenstates of the Hamiltonian in Eq. (5.1), which are centred at each of the lattice sites $\ell$ and have energies $\hbar\omega_B\ell$. They are said to form the so-called Wannier-Stark ladder of states [187, 188]. To show this, we introduce the width $\Lambda = 2J_a/\hbar\omega_B$ and the creation operator for a Wannier-Stark state

$$\hat{c}_\ell^\dagger = \sum_{\ell'} \mathcal{J}_{\ell'-\ell}(\Lambda)\hat{a}_{\ell'}^\dagger, \tag{5.2}$$

where $\mathcal{J}_m(\Lambda)$ is a Bessel function of the first kind of order $m$ [189]. The identities $\sum_m \mathcal{J}_{m-m'}(z)\mathcal{J}_{m-m''}(z) = \delta_{m',m''}$ and $(2m/z)\,\mathcal{J}_m(z) = \mathcal{J}_{m+1}(z) + \mathcal{J}_{m-1}(z)$ ensure that $\hat{H}_a$ is diagonalised as

$$\hat{H}_a = \hbar\omega_B \sum_\ell \ell \hat{c}_\ell^\dagger \hat{c}_\ell \,. \tag{5.3}$$

It follows from the inverse of the transformation in Eq. (5.2) that an impurity initially localised at a single lattice site undergoes coherent Bloch oscillations with frequency $\omega_B$ [173]. We show this in Fig. 5.2(a).

This Wannier-Stark picture is consistent with and rigorously connected to [190] the solution of the semi-classical equations for an impurity in a lattice described

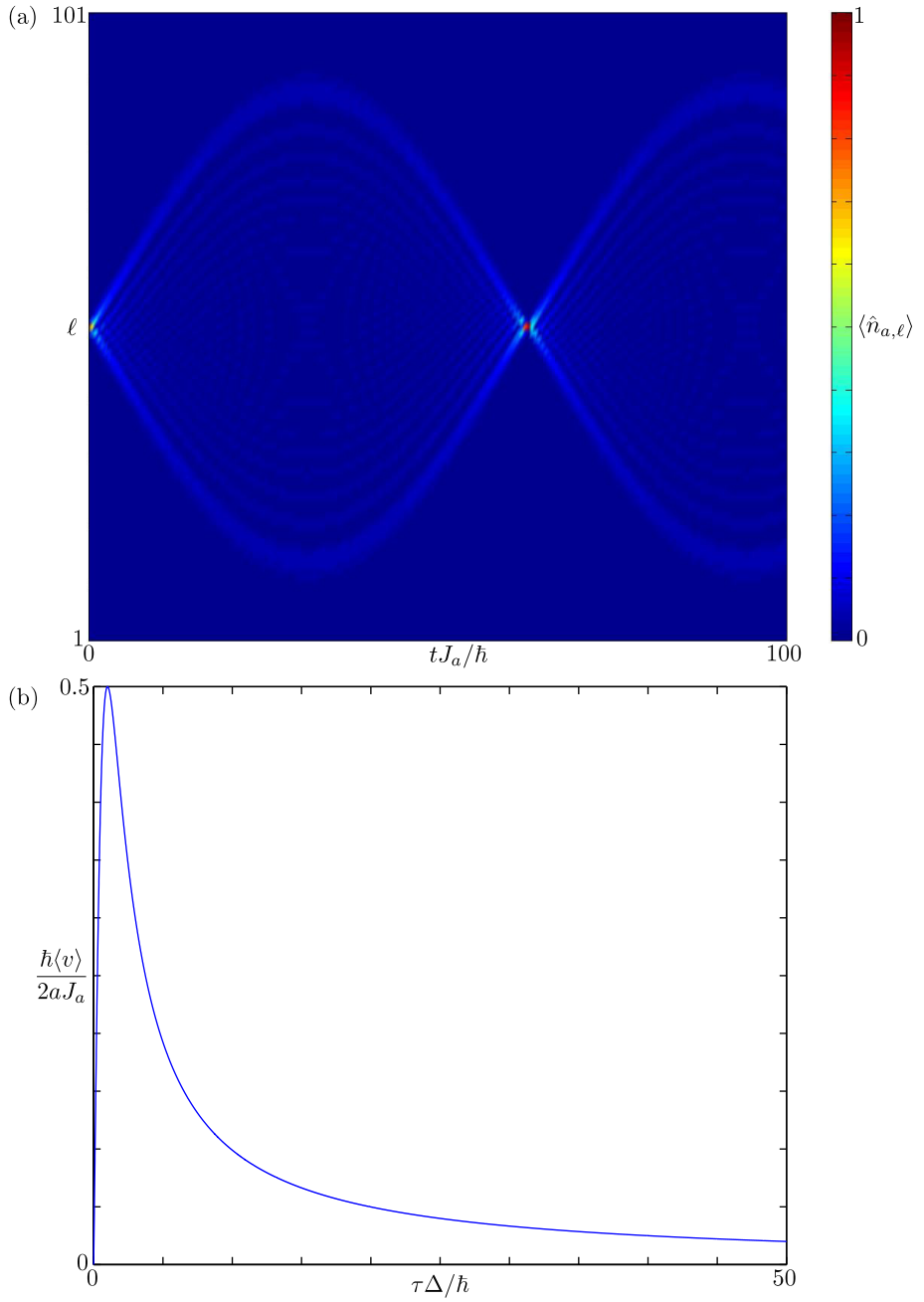


**Figure 5.2**. The motion of a single impurity in a tilted lattice with and without a scattering process. (a) The impurity decoupled from any environment will coherently Bloch oscillate at frequency $\omega_B$ but its expected position will not drift [284, 153]. Here we plot the evolution in time $t$ for just over one and a half Bloch oscillations. The impurity is initially localised in a Wannier state at the central site of the 101-site lattice. We set the tilt to be $\hbar\omega_B/J_a = 0.1$ and $\langle \hat{n}_{a,\ell} \rangle$ is the impurity density at lattice site $\ell$. (b) Introducing a scattering mechanism under the relaxation-time approximation results in the Esaki-Tsu dependence of drift on tilt in Eq. (5.5). We plot this dependence here.

by a wave-packet of well defined crystal momentum (therefore spatially it must be spread over many lattice sites) [96]. In such a framework the effect of the tilt is a constant drift of the crystal momentum of each of the impurities at rate $\hbar\dot{k} = \hbar\omega_B/a$. Together with the semi-classical equation of motion for the group velocity of a wave-packet

$$v_g(k) = \frac{1}{\hbar}\frac{\partial E_a}{\partial k}, \tag{5.4}$$

and the single-particle energy spectrum in a lattice $E_a(k) = 4J_a\sin^2(ka/2)$ this results in a group velocity that is sinusoidal in time. Thus an impurity undergoes harmonic motion with frequency $\omega_B$ but there is no net drift down the lattice.

We note that Bloch oscillations in semi-conductors are usually obscured by large scattering rates, though they have been observed in superlattices where the Bloch frequencies $\omega_B$ achievable are much higher due to larger lattice spacings [182]. Bloch oscillations are also observable in analogous optical lattice setups, as we find here, due to the much smaller scattering rates.

### 5.1.2 Impurities coupled to an environment

If the impurities are coupled to an environment then this introduces a mechanism through which an impurity can dissipate energy and fall down the lattice. Esaki and Tsu [180], using the semi-classical picture, derived an expression for the average current in the relaxation-time approximation. The main assumption is that an impurity suffers collisions at some rate $1/\tau$, which reset its crystal momentum distribution to that of thermal equilibrium (we assume zero temperature here). The resulting expression for the average velocity of each of the impurities is then

$$\langle v\rangle = \frac{2aJ_a}{\hbar}\frac{\tau\omega_B}{(\tau\omega_B)^2+1}, \tag{5.5}$$

This describes a linear (ohmic) dependence on forcing at small $\tau\omega_B \ll 1$ followed by NDC as $(\tau\omega_B)^{-1}$ when $\tau\omega_B \gg 1$. The dependence is plotted in Fig. 5.2(b). In this semi-classical framework the NDC arises if the timescale of the collisions is considerably larger than the period of Bloch oscillations, in which case the velocity of the impurity is not always in the same direction and averages out between collisions. NDC, as with Bloch oscillations, is usually obscured in semiconductors but due to the much larger values of $\tau\omega_B$ achievable it has been observed in optical lattices [174] and superlattices [191].

Clearly the relaxation-time approximation is not always valid and the actual current will depend on the details of the environment and its interaction with the impurity. However it has been found that at least qualitatively an Esaki-Tsu-like dependence of current on forcing is obeyed in many cases: for density-density interactions between an impurity and a Bose lattice gas at finite [192] and infinite temperatures [23, 171]; when an impurity scatters from a set of different fixed impurities with scattering resulting in a Fermi distribution in crystal momentum space [193]; and where an impurity scatters on phonons [183].

Here we are initially interested in the deviations to the Esaki-Tsu dependence for an impurity interacting with an environment consisting of a phonon bath, which we now introduce.

### 5.1.3 Cold atom phonon bath

Our phonon bath is provided by a Bose gas trapped in a horizontal (non-tilted) optical lattice in which the bosons are confined to move in the same single dimension as the impurities. As for the impurities, we consider temperatures to be low enough that only the lowest Bloch band need be considered. In such a case the bosons are

described by the Bose-Hubbard model (cf. Sec. 2.3.1)

$$\hat{H}_b = -J_b \sum_{\langle \ell\ell' \rangle} \hat{b}_\ell^\dagger \hat{b}_{\ell'} + \frac{U_b}{2} \sum_\ell \hat{b}_\ell^\dagger \hat{b}_\ell^\dagger \hat{b}_\ell \hat{b}_\ell \,, \tag{5.6}$$

where the operator $\hat{b}_\ell^\dagger$ ($\hat{b}_\ell$) creates (annihilates) a boson in a Wannier state localised at site $\ell$ and the lattice parameter $a$ is the same as that for the impurities. As discussed in Sec. 2.3.1, $J_b$ and $U_b$ are parameters that determine the hopping between neighbouring sites and the on-site interaction, respectively, and can be tuned experimentally by adjusting the laser parameters or using Feshbach resonances [25, 58]. We depict the bosonic Hamiltonian in Fig. 5.1(b).

At low temperatures and for sufficiently small boson-boson interactions $U_b/J_b \lesssim 1$ most bosons are in the superfluid phase and are accurately described by phononic excitations (see Sec. 2.3.2)

$$\hat{H}_b = \sum_k \hbar\omega_k \hat{d}_k^\dagger \hat{d}_k \,. \tag{5.7}$$

Here $k$ is a wave-vector running over the first Brillouin zone, i.e., $-\pi/a < k \leq \pi/a$ and the bosonic operators $\hat{d}_k^\dagger$ ($\hat{d}_k$) create (annihilate) a Bogoliubov phonon. Restating what we wrote in Sec. 2.3.2, the excitation spectrum of the phonons is given by the Bogoliubov dispersion relation

$$\hbar\omega_k = \sqrt{\varepsilon_k(\varepsilon_k + 2U_b n_b)}, \tag{5.8a}$$

$$\varepsilon_k = 4J_b \sin^2(ka/2), \tag{5.8b}$$

where $\varepsilon_k$ are the single-particle eigenenergies and $n_b = N/M$ is the bosonic occupation number, with $N$ the number of bosons and $M$ the number of lattice sites. The dispersion of the Bogoliubov phonons, albeit not identical, is remarkably similar to the dispersion of acoustic phonons in a solid state system with $\hbar\omega_k \sim |\sin(ka/2)|$,

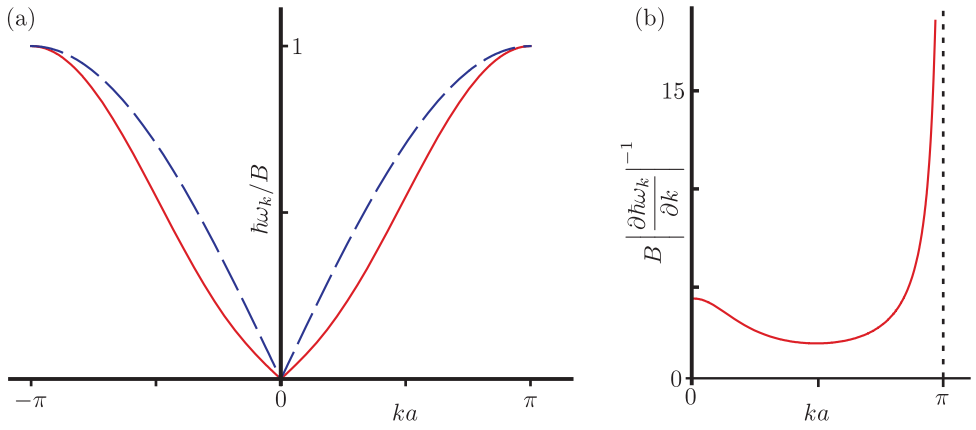


**Figure 5.3**. (a) The dispersion relation of acoustic phonons in a solid with $\hbar\omega_k \sim |\sin(ka/2)|$ (blue dashed line) and of Bogoliubov phonons in a superfluid confined to an optical lattice (red full line) in units of the bandwidth $B$ (for Bogoliubov phonons $B = 4J_b\sqrt{1 + U_b n_b/2J_b}$). The parameters used satisfy $U_b n_b/J_b = 0.5$. (b) The corresponding density of states $|\partial\hbar\omega_k/\partial(ka)|^{-1}$ (red full line) of the Bogoliubov phonons is approximately constant for small wave-vectors and exhibits a Van Hove singularity near the edge of the first Brillouin zone (black dashed line).

as shown in Fig. 5.3(a). The common features are a linear dispersion in the limit $k \to 0$ and the band structure with a gap at the boundary of the first Brillouin zone. It should also be pointed out that the accuracy of the Bogoliubov description of phonons has been demonstrated experimentally [178], thus we have perfect knowledge of and control over the phonons in our system.

What is left is to describe the process by which impurities dissipate energy into the phonon environment. The interaction that naturally arises in cold atom experiments is the contact, or density-density, interaction. Within the same approximations as the other parts of the Hamiltonian, e.g., low energies, the interaction takes the explicit form

$$\hat{H}_{ab} = U_{ab} \sum_{\ell} \hat{a}_\ell^\dagger \hat{a}_\ell \hat{b}_\ell^\dagger \hat{b}_\ell, \tag{5.9}$$

where $U_{ab}$ is its characteristic energy. This is illustrated in Fig. 5.1(c).

To summarise, the full Hamiltonian $\hat{H} = \hat{H}_a + \hat{H}_b + \hat{H}_{ab}$ describing our system is

given by the impurity Hamiltonian $\hat{H}_a$, the Bose-Hubbard model $\hat{H}_b$ and the density-density interaction $\hat{H}_{ab}$, as shown schematically in Fig. 5.1. The system is the same as that considered in Ref. [23] where for low tilts $\hbar\omega_B/J_b \ll 1$ the authors used that the relevant correlation functions of the Bose gas decay quickly enough to allow the dynamics of the impurity to be described by a Markovian master equation. For the same system, the authors of Ref. [194] use the chaotic system approach along with Kubo's formalism to derive the current in the ohmic region $\omega_B\tau \ll 1$.

Before continuing we re-express the Hamiltonian $\hat{H}_{ab}$ in a more relevant form and then review the physical properties we expect to see in the system described by this Hamiltonian. Specifically, we rewrite the impurity-boson interaction in terms of Bogoliubov phonons and Wannier-Stark states by applying the same transformations to $\hat{H}_{ab}$ as for the impurity and bosonic parts $\hat{H}_a$ and $\hat{H}_b$. We find the interaction Hamiltonian to be

$$\hat{H}_{ab} = U_{ab} n_b \sum_{\ell} \hat{c}_\ell^\dagger \hat{c}_\ell + \sum_{\ell,\Delta\ell,k} \left[ f_{k,\Delta\ell} \hat{d}_k + \text{h.c.} \right] \hat{c}_{\ell+\Delta\ell}^\dagger \hat{c}_\ell, \tag{5.10}$$

where $\Delta\ell$ is an integer and the matrix elements are

$$f_{k,\Delta\ell} = U_{ab} \left( \frac{n_b \varepsilon_k}{M\hbar\omega_k} \right)^{1/2} \mathcal{J}_{\Delta\ell} \left( 2\Lambda \sin\frac{ka}{2} \right) \mathrm{i}^{\Delta\ell} \mathrm{e}^{-\mathrm{i}ka(\ell+\Delta\ell/2)} \,. \tag{5.11}$$

We note that $\hat{H}_{ab}$ coincides exactly with the description of the electron-phonon interactions in a solid state system [187] and is thus perfectly adequate for investigating related effects.

The phenomena described by the impurity-phonon interaction $\hat{H}_{ab}$ include the creation of a Bogoliubov phonon out of the superfluid phase and the reverse process, both caused by the hopping process of the impurity. These incoherent processes involving a single phonon enable impurities to make transitions between Wannier-

Stark states separated by $\Delta\ell$ lattice sites. On the other hand, coherent processes resulting from $\hat{H}_{ab}$ in which a virtual phonon is emitted and reabsorbed lead to a renormalised impurity hopping, a phonon-mediated impurity-impurity interaction, and a mean-field energy shift of the impurities [24, 130].

Our theoretical treatment below captures only incoherent processes. Therefore we will have to add in the coherent effects mentioned above when comparing our theoretical model to experimental or numerical results. It is expected from theoretical considerations [24, 130] and confirmed by our numerical results in Sec. 5.1.7 that the renormalisation reduces the bare hopping $J_a$. The effective impurity-impurity interaction, being short-range, can be safely neglected if we assume a filling factor of the impurities much lower than 1. Last, the mean-field energy shift is explicitly given by $U_{ab}n_b + \sum_k |f_{k,0}|^2/\hbar\omega_k$, which is readily determined in the regime $U_b n_b/J_b \ll 1$ by taking the continuum limit[2]. For a single delocalised impurity with $J_a/\hbar\omega_B \gg 1$ the shift is simply $U_{ab}n_b$, whereas for a localised impurity with $J_a/\hbar\omega_B \ll 1$ we find $U_{ab}n_b(1 - aU_{ab}/2\hbar c)$. Here $c = (a/\hbar)\sqrt{2J_bU_bn_b}$ the speed of sound in the superfluid. We restrict our analysis to impurities moving slower than the superfluid critical velocity, which is $c$ according to Landau's criterion [195], to avoid excitations other than those caused by the hopping transitions. More precisely we consider the parameter regime $\zeta = 2aJ_a/\hbar c \ll 1$, where $2aJ_a/\hbar$ is the maximal group velocity of the impurities in the lowest Bloch band.

### 5.1.4 Phonon resonances

Since the number of impurities is significantly smaller than the number of bosons we effectively treat the superfluid as a phonon bath at temperature $T = 1/\beta k_B$. Accordingly we describe the dynamics of the impurities by a master equation for

[2]Explicitly, the continuum limit of the sum over all crystal momenta in the lowest band is $(Ma)^{-1}\sum_k \to \int_{-\pi/a}^{\pi/a} \mathrm{d}k/2\pi$.

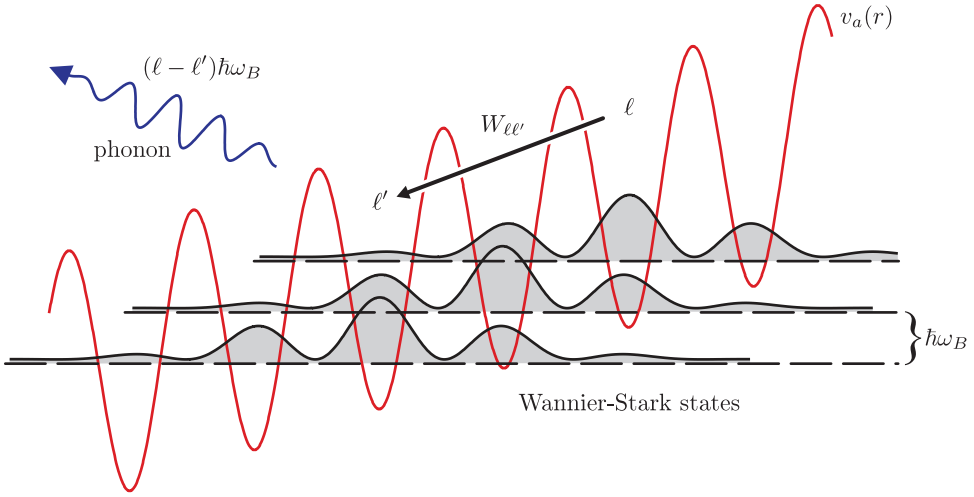


**Figure 5.4.** Impurities confined to a tilted but oscillating potential $v_a(r)$ form a Wannier-Stark ladder with a constant energy separation $\hbar\omega_B$ between adjacent lattice sites. The impurity-boson interaction enables phonon-assisted transitions from site $\ell$ to $\ell'$ at rate $W_{\ell\ell'}$. The released energy $(\ell - \ell')\hbar\omega_B$ is dissipated in the form of a phonon emitted into the superfluid.

the probabilities $P_\ell$ that an impurity occupies the Wannier-Stark states centred at sites $\ell$. The master equation reads

$$\partial_t P_\ell = \sum_{\ell' \neq \ell} \left[ P_{\ell'} W_{\ell'\ell} - P_\ell W_{\ell\ell'} \right] , \tag{5.12}$$

where $W_{\ell\ell'}$ is the rate of a phonon-assisted transition from site $\ell$ to site $\ell'$. The probabilities $P_\ell$ either describe the occupation of a single impurity or the distribution of an ensemble of impurities. In both cases we choose $\sum_\ell P_\ell = 1$. The average displacement $\langle \hat{r}_a \rangle$ of the impurities at time $t$ relative to site $\ell_0$ is $\langle \hat{r}_a \rangle = a \sum_\ell (\ell - \ell_0) P_\ell$ and the average drift velocity $\langle v \rangle$, which quantifies the atomic current along the lattice, is given by $\langle v \rangle = \partial_t \langle \hat{r}_a \rangle = a \sum_\ell (\ell - \ell_0) \partial_t P_\ell$.

To obtain a more useful expression for the drift velocity $\langle v \rangle$ we exploit the fact that the system is homogeneous so that the transition rate $W_{\ell\ell'}$ only depends on the jump distance $\Delta\ell = \ell' - \ell$. A jump with $\Delta\ell > 0$ is defined to be up the tilted

lattice, as depicted in Fig. 5.4. By using Fermi's golden rule based on the interaction Hamiltonian $\hat{H}_{ab}$, i.e., to second order in the coupling $U_{ab}$, we find

$$W^E_{|\Delta\ell|} = \frac{2\pi}{\hbar} \sum_k |f_{k,|\Delta\ell|}|^2 (N_k + 1)\, \delta\left(\hbar\omega_k - |\Delta\ell|\hbar\omega_B\right) , \tag{5.13a}$$

$$W^A_{|\Delta\ell|} = \frac{2\pi}{\hbar} \sum_k |f_{k,|\Delta\ell|}|^2 N_k\, \delta\left(\hbar\omega_k - |\Delta\ell|\hbar\omega_B\right) , \tag{5.13b}$$

for the rate $W^A_{|\Delta\ell|}$ ($W^E_{|\Delta\ell|}$) of jumping $|\Delta\ell|$ lattice sites up (down) the ladder by absorbing (emitting) a phonon of energy $|\Delta\ell|\hbar\omega_B$. Here we allow for the finite temperature of the bath through the mean occupation of the phonon modes $N_k = (\mathrm{e}^{\beta\hbar\omega_k} - 1)^{-1}$. The expressions for the jump rates $W^E_{|\Delta\ell|}$ and $W^A_{|\Delta\ell|}$ are valid as long as heating effects caused by the emitted phonons are negligible.

To determine $\langle v\rangle$ in terms of transition rates we substitute the expression in Eq. (5.12) for $\partial_t P_j$ and find after some rearrangement that

$$\langle v\rangle = -a \sum_{|\Delta\ell|>0} |\Delta\ell| W^0_{|\Delta\ell|}, \tag{5.14}$$

with

$$W^0_{|\Delta\ell|} = W^E_{|\Delta\ell|} - W^A_{|\Delta\ell|} = (2\pi/\hbar) \sum_k |f_{k,|\Delta\ell|}|^2\, \delta\left(\hbar\omega_k - |\Delta\ell|\hbar\omega_B\right), \tag{5.15}$$

the phonon emission rate at zero temperature. To arrive at this expression we have made use of the normalisation $\sum_\ell P_\ell = 1$. Note that, within our approximations, the current does not depend on the temperature of the bath.

Taking the continuum limit of the sum over all phonon crystal momenta and integrating over the first Brillouin zone we finally obtain

$$\langle v\rangle = -\frac{a^2 M}{\hbar} \sum_{|\Delta\ell|=1}^{|\Delta\ell|_{\max}} |\Delta\ell| \left|\frac{\partial\hbar\omega_k}{\partial k}\right|^{-1} |f_{k,|\Delta\ell|}|^2 \Big|_{k=k_{|\Delta\ell|}} , \tag{5.16}$$

with $k_{|\Delta\ell|}$ implicitly defined by $\varepsilon_{k_{|\Delta\ell|}} = \sqrt{(|\Delta\ell|\hbar\omega_B)^2 + (U_b n_b)^2} - U_b n_b$. The latter relation also yields the maximum jump distance $|\Delta\ell|_{\max}$ that is compatible with energy conservation in the impurity-phonon relaxation process. We see from Eq. (5.16) that the total drift velocity $\langle v\rangle$ is determined by the sum over all admissible jump distances $|\Delta\ell|$ weighted by the phonon density of states $|\partial\hbar\omega_k/\partial k|^{-1}$ and the impurity-phonon coupling $|f_{k,|\Delta\ell|}|^2$.

We first determine the mobility $\mu_{\mathrm{mob}}$ of the impurities in the ohmic regime, which we define by the relation $\langle v\rangle = 4\mu_{\mathrm{mob}}\hbar\omega_B/J_b$ in the limit $\hbar\omega_B/J_b \to 0$. We find that $\langle v\rangle J_b/\hbar\omega_B \propto \sum_{|\Delta\ell|=1}^{\infty} |\Delta\ell|^2\, \mathcal{J}^2_{|\Delta\ell|}\left(|\Delta\ell|\,\zeta\right)$ for $\hbar\omega_B/J_b \to 0$, with $\zeta = 2aJ_a/\hbar c$ as previously defined. This sum is a Kapteyn series of the second kind, for which a closed form is known [189]. It yields the mobility

$$\mu_{\mathrm{mob}} = \frac{aU_{ab}^2}{8\sqrt{2}\,U_b\hbar}\left(\frac{J_b}{U_b n_b}\right)^{1/2}\frac{\zeta^2(\zeta^2+4)}{(1-\zeta^2)^{7/2}}\,. \tag{5.17}$$

The mobility diverges in the limit $\zeta \to 1$, however, it is well defined in the regime we consider $\zeta \ll 1$. For the practically important case, where the potential depth for the impurities and hence the hopping parameter $J_a$ is varied, the mobility scales as $\mu_{\mathrm{mob}} \sim J_a^2$ in the limit $J_a \to 0$.

The dependence of the drift velocity $\langle v\rangle$ given in Eq. (5.16) on the full range of accessible lattice tilts $\hbar\omega_B/J_b$ is shown in Fig. 5.5. We see that the atomic current makes a sharp transition from ohmic conductance to NDC. One way of understanding this (a semiclassical interpretation has been given in Sec. 5.1.2) is that the width $\Lambda = 2J_a/\hbar\omega_B$ of the Wannier-Stark states, hence the overlap between states at different sites, decreases as the tilt $\hbar\omega_B/J_b$ is increased for fixed $J_a/J_b$. This reduced overlap results in a low drift velocity $\langle v\rangle$. The drift velocity $\langle v\rangle$ depends on the tilt as $(J_a/\hbar\omega_B)^{2|\Delta\ell|}$ for each jump distance $|\Delta\ell|$ in the limit of small impurity bandwidth $J_a/\hbar\omega_B \ll 1$, which can be readily achieved in cold atom systems. As

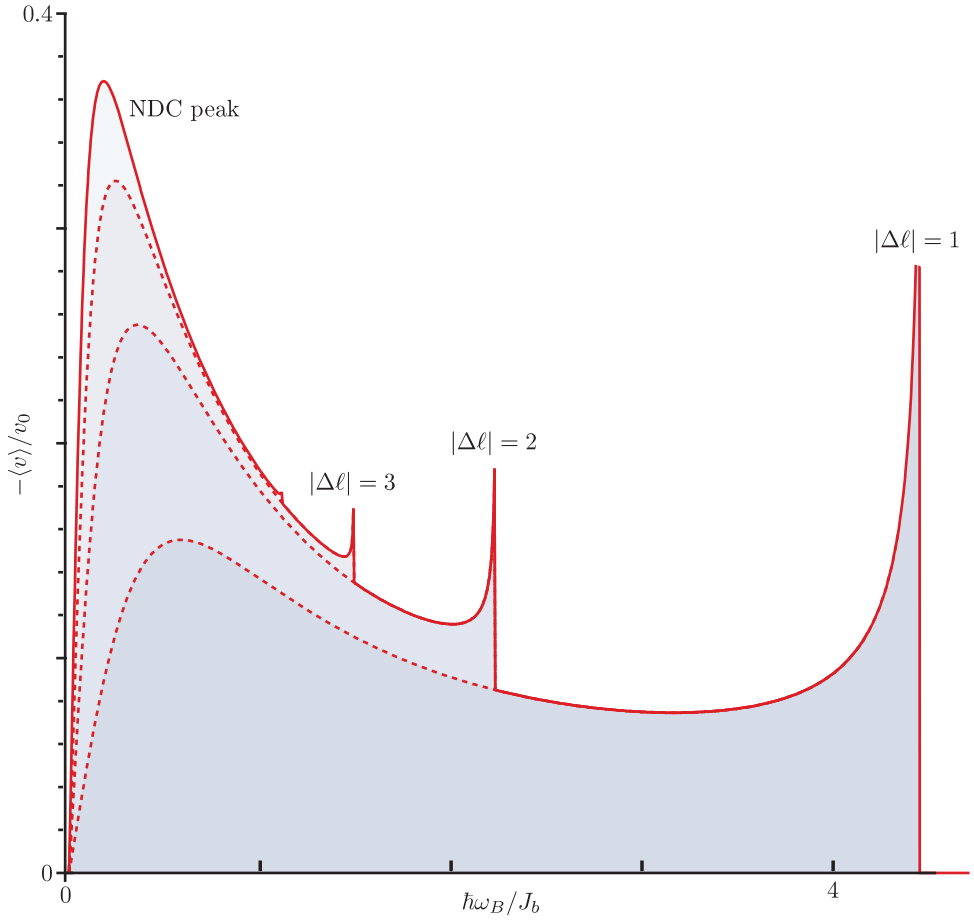


**Figure 5.5**. The total drift velocity $\langle v\rangle$ (full line) as a function of the lattice tilt $\hbar\omega_B$. The contributions from different jump distances $|\Delta\ell|$ (dashed lines) result in a distinct NDC peak and a series of phonon resonances at the tilts $4\hbar\omega_B/J_b \approx 1/|\Delta\ell|$. The system parameters are $U_b n_b/J_b = 0.5$ and $J_a/J_b = 0.4$, which yields $B = 1.12 \times 4J_b$ for the phonon bandwidth and $\zeta = 0.5$. Here the prefactor is $v_0 = a n_b U_{ab}^2/4J_b\hbar$.

can be seen in Fig. 5.5 the NDC is particularly pronounced due to the superposed contributions from different jump distances $|\Delta\ell|$ to the total current. These findings are in accordance with predictions for a free homogeneous superfluid [23, 130, 171], however, the result for $\langle v\rangle$ in Eq. (5.16) describes additional features stemming from the influence of the phonon density of states.

Specifically, the drift velocity exhibits anomalies in the form of sharp peaks that correspond to the anticipated electron-phonon resonances in a solid state system [183, 184]. As shown in Fig. 5.3(b), the phonon density of states $|\partial\hbar\omega_q/\partial q|^{-1}$ is approximately constant for small crystal momenta but exhibits a Van Hove singularity [186] at the edge of the first Brillouin zone. Therefore phonon resonances arise as the crystal momentum of the emitted phonons approaches the edge of the phonon band. It can be found from Eq. (5.16) that the resonance condition is given by $4|\Delta\ell|\hbar\omega_B/J_b = \sqrt{1+U_b n_b/2J_b}$ with $|\Delta\ell| = 1, 2, \ldots, |\Delta\ell|_{\max}$. Thus the current displays a peak at $4\hbar\omega_B/J_b \approx 1/|\Delta\ell|$ corresponding to each admissible jump distance $|\Delta\ell|$. Since the impurity-phonon interaction provides the only relaxation process in the system these anomalies are directly reflected in the tilt dependence of the current. Further, the current vanishes as the lattice tilt exceeds the phonon bandwidth because there are no phonon states available in order to dissipate energy into the superfluid (we discuss this in more detail in Sec. 5.2).

### 5.1.5 Outline of experimental procedure

The theoretical results derived in the previous section are, strictly speaking, valid for stationary atomic currents in an infinite size homogeneous system. We now show that the predicted NDC and the phonon resonances are observable in a system of finite size under the conditions of a realistic measuring procedure.

The procedure is as follows. The $N$ bosons are cooled to their ground state

with the few impurities (see Ref. [17] on injecting a controlled number of impurities) fixed in highly-localised states at their initial locations (for example, using a tight optical trap [14]). In this configuration the impurities are automatically cooled by the surrounding superfluid [196]. The system reaches equilibrium for times $t < 0$ and the presence of the impurities leads to density variations in the initial state of the bosons as compared to the ground state of the Bose-Hubbard Hamiltonian of Eq. (5.6) (cf. the self-trapping [165] effects discussed in Ch. 4). An example of this can be seen in Fig. 5.6(a). Then the impurities are released at $t = 0$, their lattice is tilted and the system is left to evolve. An *in situ* image [104] is taken a time $t$ later. From the difference between the initial and final density distributions it should be possible to extract a reliable estimate for the atomic current after multiple realisations. Alternatively, the crystal momentum distribution and hence the drift velocity of the impurities may be determined directly by a time-of-flight measurement [58]. Note that we considered an alternative initial state in which the bosons were prepared in equilibrium without the impurities and then the impurities were added at $t = 0$. This produced the same qualitative results despite this being a more energetic initial state (though we do not have space to show those results in this thesis).

For simplicity, here we consider a single impurity initially localised in a Wannier state at a central lattice site $\ell_0 = (M + 1)/2$ (for odd $M$). The expected impurity displacement $\langle \hat{r}_a \rangle = \sum_\ell a(\ell - \ell_0)\langle \hat{n}_{a,\ell} \rangle$ can be used to calculate the average impurity velocity $\langle v \rangle = \langle \hat{r}_a \rangle / t$. Below (see Fig. 5.6(b) for example), we will find that the displacement does not increase perfectly linearly in time but oscillates slightly at the frequency $\omega_B$ of the Bloch oscillations. For all but the lowest tilt energies $\hbar\omega_B$ the duration of evolution considered is sufficient that by the maximum time $t_f$ the displacement has averaged over enough Bloch oscillations that $\langle v \rangle$ gives a good representation of the long-time average drift.

The maximum time is given by $t_f = M\hbar/4J_b$, which is approximately how long it takes the disturbances in the bath caused by the impurity to reach the boundaries of the system. This precise expression is determined by the maximum group velocity $v_g$ of the bosons in the band which can be derived in the superfluid limit $U_b/J_b \to 0$ (and also in the strongly-interacting $U_b/J_b \to \infty$ limit, as will be useful for Sec. 5.2) by differentiating $\hbar\omega_k$ (or $\varepsilon_k$) — cf. Eqs. (5.8) — with respect to $k$. The time taken for disturbances in the Bose gas to travel a distance $Ma/2$ to the edge of the system is then $Ma/2v_g$ from which the previous expression for $t_f$ follows. Restricting ourselves to times $t$ smaller than this thus minimises finite size effects.

### 5.1.6 Outline of numerical simulations

We now numerically realise the above experimental procedure by simulating the nearest-neighbour Hamiltonian $\hat{H}_a + \hat{H}_b + \hat{H}_{ab}$, using the forms given in Eqs. (5.1), (5.6) and (5.9). To make the Hilbert space a finite size we restrict the bosonic occupancy of each lattice site to a maximum of 4, a common approximation that has been extensively and successfully used in the study of the Bose-Hubbard model with low filling. Since we only have a single impurity the local Hilbert space for each lattice site is $d = 5 \times 2 = 10$. This is the tensor product of two Hilbert spaces, the first for the bosons is spanned by 5 Fock states and the second for the impurity is spanned by 2 Fock states.

With the size of the local Hilbert space restricted and the Hamiltonian to be simulated in nearest-neighbour form, the time-evolving block decimation (TEBD) algorithm (cf. Sec. 3.2) is applicable. We first use TEBD to simulate imaginary time-evolution and find the ground state of the bosonic part of the system, with an on-site potential $U_{ab}$ added at the site $\ell_0$ corresponding to the initial location of an impurity. The initial state of the whole system is then the tensor product of this

bosonic state with the initial state of the impurity, which consists of a product of Fock states with occupancies 0 or 1. Next, TEBD is used to simulate the unitary evolution of this state under the full Hamiltonian $\hat{H} = \hat{H}_a + \hat{H}_b + \hat{H}_{ab}$. Expected values such as bosonic and impurity densities are then easily calculated from the evolved matrix product state (MPS), as discussed in Sec. 3.2.3. The largest MPS dimension we use for our simulations is $\chi = 120$. The timestep is $\delta t = 5 \times 10^{-3} \hbar / J_b$. We found these parameters were sufficient; increasing $\chi$ and decreasing $\delta t$ resulted in no significant changes to the observables computed.

### 5.1.7 Numerical results

As a demonstration of the system and numerical method we show the evolution of the densities of both the impurity and bosons in Fig. 5.6(a) for a weak boson-boson interaction strength $U_b/J_b = 0.1$ and a medium tilt $\hbar\omega_B/J_b = 0.57$. The impurity undergoes Bloch oscillations as it would if the bosons were not present (see Fig. 5.2(a)), however, non-zero impurity-boson interactions cause the impurity to dissipate energy into the Bose gas and drift down the lattice. The same interactions also create density variations in the Bose gas that spread out and reach the system boundary near the end of the simulation $t_f$.

Examining the system for a range of tilts $\hbar\omega_B$ allows us to find the dependence of the impurity displacement $\langle \hat{r}_a \rangle$ on the tilt. This dependence is shown in Fig. 5.6(b), where we have also included the dependence on time. From this time-dependence we can see that while the displacement does not grow exactly linearly in $t$, after a duration $t \gtrsim \omega_B^{-1}$ the average velocity $\langle v \rangle = \langle \hat{r}_a \rangle / t$ converges to the long-time drift velocity, meaning that our value of $\langle \hat{r}_a \rangle$ at time $t_f$ is a good indicator of this drift velocity for all but the smallest tilts $\hbar\omega_B$.

Therefore in Fig. 5.7 we plot the displacement $\langle \hat{r}_a \rangle$ of the impurity at the fi-

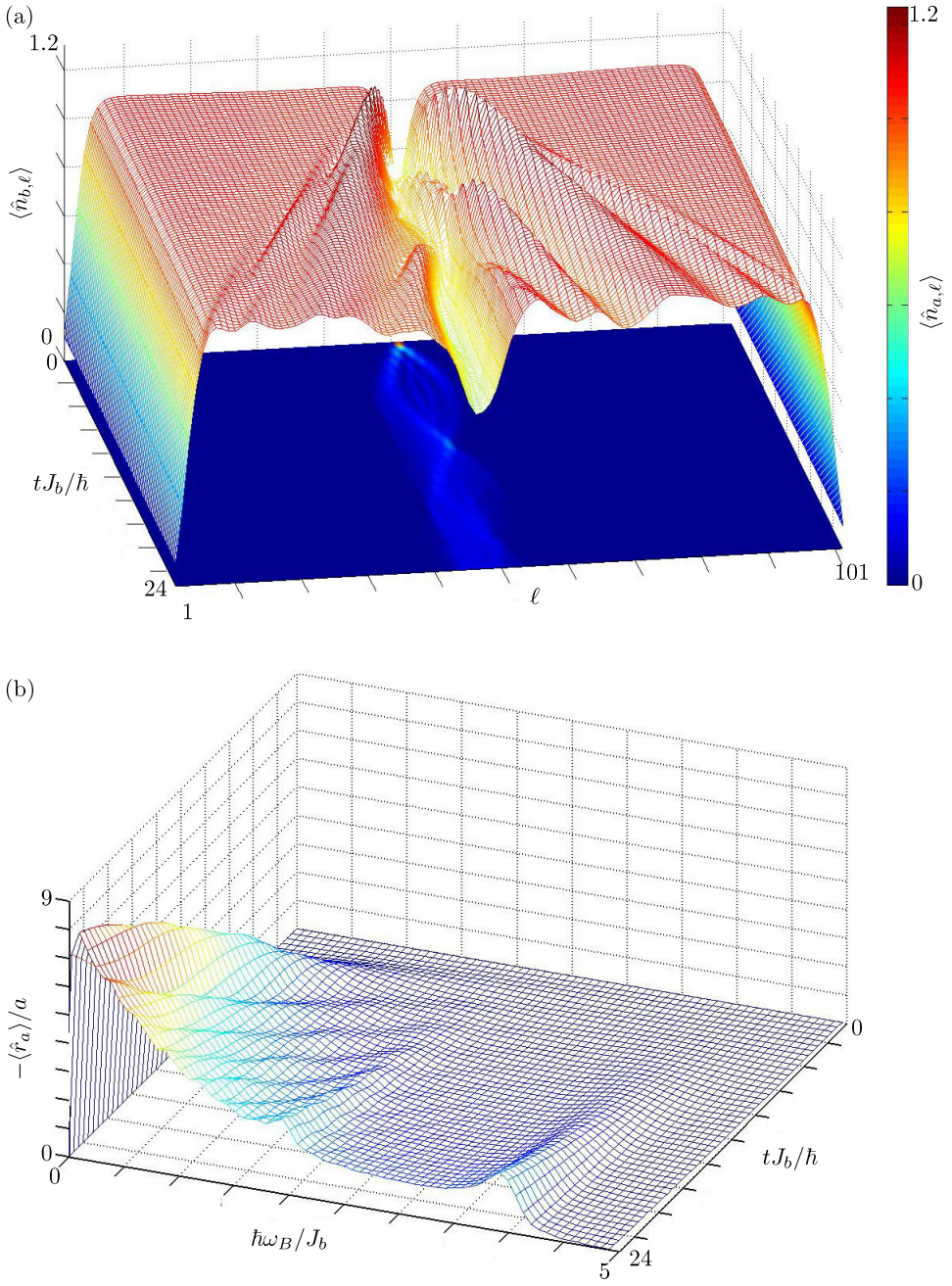


**Figure 5.6**. The transport of an impurity through a bosonic bath. (a) An example of the evolution of the impurity and boson densities for the case $J_a/J_b = 0.8$, $\hbar\omega_B/J_b = 0.57$, $U_b/J_b = 0.1$, $U_{ab}/J_b = 0.5$, $M = 101$ and $N = 101$. The colour map in the $xy$-plane shows the expected impurity density at each lattice site $\langle \hat{n}_{a,\ell} \rangle$, and similarly for the surface plot and boson density $\langle \hat{n}_{b,\ell} \rangle$. (b) The displacement $\langle \hat{r}_a \rangle$ of the impurity as a function of time $t$ and lattice tilt $\hbar\omega_B$.

nal time $t_f$ as a function of the tilt $\hbar\omega_B/J_b$. We do this for three values of $J_a/J_b$. The main features predicted by our theoretical analysis in Sec. 5.1.4 can be clearly recognised, namely the dominant NCD peak, the phonon resonance at $4\hbar\omega_B/J_b \approx 1$ corresponding to the jump distance $|\Delta\ell| = 1$ and the suppression of the current for $\hbar\omega_B/J_b \gg 1$. In addition, for sufficiently large values of $J_a/J_b$ the phonon resonance at $4\hbar\omega_B/J_b \approx 1/2$ corresponding to the jump distance $|\Delta\ell| = 2$ is visible. For lower values of $J_a$ the phonon resonance for $|\Delta\ell| = 2$ may still lead to a characteristic drop in the current once the lattice tilt exceeds half the phonon bandwidth. The resonances for higher values of the jump distance $|\Delta\ell|$ seem to be masked by broadening and finite-time effects for the parameter regimes tested.

The ratio between the height of the $|\Delta\ell| = 1$ phonon resonance and the NDC peak, i.e., their relative visibility $V$, depends nontrivially on the impurity hopping $J_a$. The hopping $J_a$ enters the expression for $\langle v\rangle$ in Eq. (5.16) through the matrix elements $f_{k,|\Delta\ell|}$ in the form of $\mathcal{J}_{|\Delta\ell|}(s|\Delta\ell|J_a/J_b)$, where the parameter $s$ depends on the lattice tilt. Close to resonance we have $s \approx 1$, whereas for small tilts $|\Delta\ell|\hbar\omega_B \ll U_b n_b$ we obtain $s \approx \sqrt{2J_b/U_b n_b}$ and hence $s \gg 1$ in the superfluid regime. Accordingly, the height of the $|\Delta\ell| = 1$ phonon resonance ($s \approx 1$) varies only slowly with $J_a$ as opposed to the height of the NDC peak ($s \gg 1$), which is characterised by the oscillatory nature of the Bessel functions. Thus a careful choice of $J_a$ allows us to optimise the visibility $V$, i.e., to minimise the height of the NDC peak, by tuning $s|\Delta\ell|J_a/J_b$ close to a zero of the relevant Bessel functions. For the increasing values of $J_a$ considered in our simulation the first zero of the Bessel function $\mathcal{J}_2(z)$ is approached, which explains the improved visibility $V$ for higher $J_a$ shown in Fig. 5.7.

The general broadening of the phonon resonances may be caused by finite-size effects and multi-phonon processes. The finite-size effects include reflections of phonons from the boundary of the system that in turn affect the motion of the impurity. Multi-phonon processes, which have been neglected in our theoretical

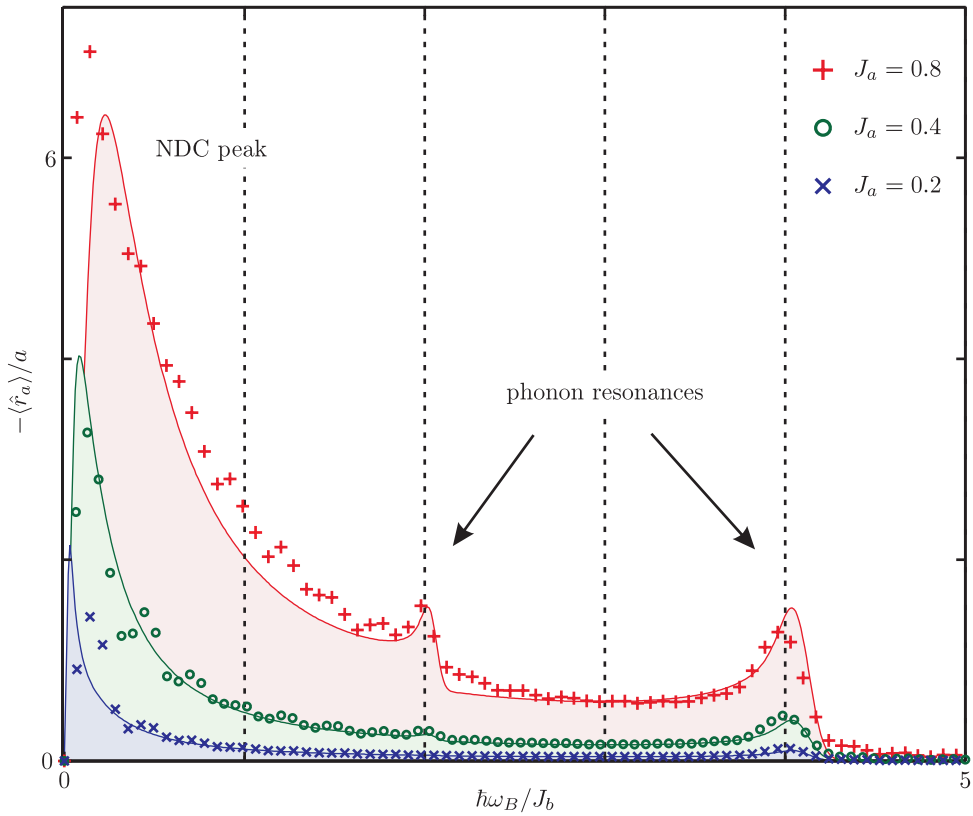


**Figure 5.7**. The displacement $\langle \hat{r}_a \rangle$ of the impurity after time $t_f$ as a function of the lattice tilt $\hbar\omega_B/J_b$, determined numerically (symbols $+$, $\circ$, $\times$), and the corresponding analytical fits (solid lines). The dominant NDC peak and the $|\Delta\ell| = 1$ phonon resonance at the boundary of the phonon band can be clearly recognised. The ratio $V$ between the height of the $|\Delta\ell| = 1$ resonance and the NDC peak is $V = 0.09, 0.14, 0.18$ for increasing $J_a$. The $|\Delta\ell| = 2$ resonance is visible for sufficiently large impurity hopping $J_a$. The system parameters are $J_a/J_b = 0.2, 0.4, 0.8$, $U_b/J_b = 0.1$, $U_{ab}/J_b = 0.5$ and the fitting parameters are $\varepsilon_{\text{broad}}/J_b = 0.1$, $\tilde{J}_a/J_a = 0.7$, $|\Delta\ell|_{\max} = 2$.

analysis, also cause some broadening of the single-phonon resonances. In particular, the continuous drop in the current to zero can be explained by the emission of several low-energy phonons during a jump process, which is allowed even if the lattice tilt exceeds the phonon bandwidth.

In order to compare the numerical and theoretical results, at least qualitatively, we do two things. First, we replace the bare impurity hopping $J_a$ in our theoretical results by a renormalised value $\tilde{J}_a < J_a$ as discussed in Refs. [24, 130]. This is in direct analogy to the increased effective mass of polarons due to the drag of the phonon cloud [197]. Second, we include broadening effects characterised by an energy $\varepsilon_{\text{broad}}$ into our theory. Explicitly, we calculate the average drift velocity by using the expression $\langle v\rangle^{\varepsilon_{\text{broad}}} = -a\sum_{|\Delta\ell|=1}^{|\Delta\ell|_{\max}} |\Delta\ell|\, W_{|\Delta\ell|}^{\varepsilon_{\text{broad}}}$ with the rates

$$W_{|\Delta\ell|}^{\varepsilon_{\text{broad}}} = (2\pi/\sqrt{\pi}\varepsilon_{\text{broad}}\hbar)\sum_{q} |f_{q,|\Delta\ell|}|^2 \exp[-(\hbar\omega_q - |\Delta\ell|\hbar\omega_B)^2/\varepsilon_{\text{broad}}^2]. \tag{5.18}$$

The rates reduce to the previous expression $W_{|\Delta\ell|}^0$ in absence of broadening, i.e., $\lim_{\varepsilon_{\text{broad}}\to 0} W_{|\Delta\ell|}^{\varepsilon_{\text{broad}}} = W_{|\Delta\ell|}^0$.

Using $\langle \hat{r}_a\rangle^{\varepsilon_{\text{broad}}} = \langle v\rangle^{\varepsilon_{\text{broad}}} t_f$ we fit this theoretical model to the numerical results in Fig. 5.7 with the broadening $\varepsilon_{\text{broad}}$, the maximal jump distance $|\Delta\ell|_{\max}$ and the effective impurity hopping $\tilde{J}_a$ as free parameters. The first two parameters allow us to extract a quantitative value for the broadening and to determine the dominant hopping processes in the experiment, respectively.

As can be seen in Fig. 5.7 the fit describes the numerical results very accurately with moderate broadening $\varepsilon_{\text{broad}}$ and a minor reduction of the impurity hopping $J_a$. Further, we find that $|\Delta\ell|_{\max} = 2$ and thus the dominant transport processes are nearest- and next-nearest-neighbour hopping; this is consistent with the absence of higher-order phonon resonances noted earlier.

Since we simulate the full dynamics of the system we can extend the range of

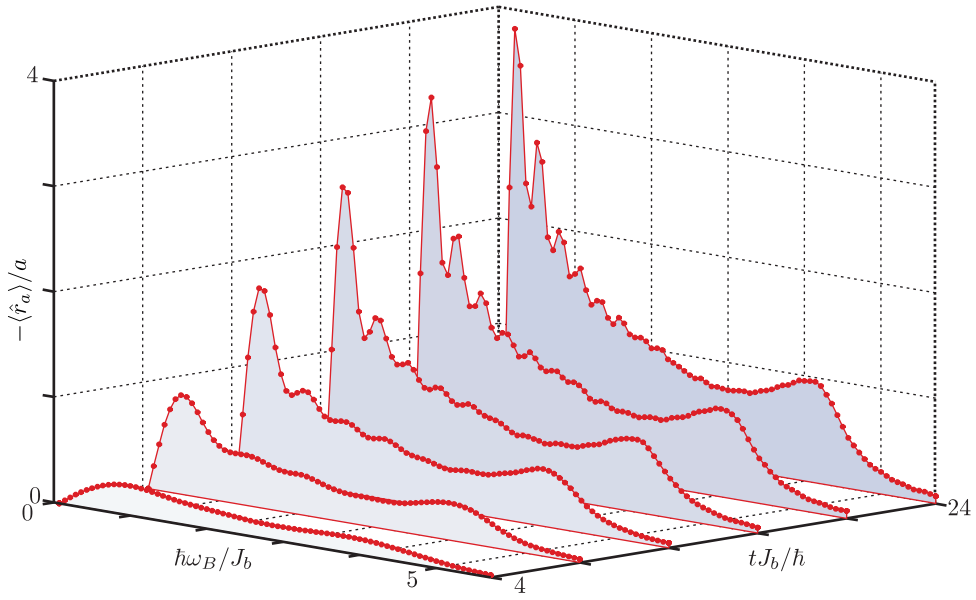


**Figure 5.8**. The numerically determined displacement of the impurity as a function of the lattice tilt $\hbar\omega_B/J_b$ for different evolution times $t$ (connected data points). The NDC peak and the $|\Delta\ell| = 1$ phonon resonance are clearly visible in presence of strong impurity-boson and boson-boson interactions. The additional peaks between the NDC peak and the phonon resonance are due to finite-time effects. Otherwise the displacement increases approximately linearly with the evolution time $t$. The system parameters are $J_a/J_b = 0.5$, and $U_b/J_b = U_{ab}/J_b = 1$.

parameters considered so far. In particular, we are interested in the atomic current in the presence of stronger impurity-boson and boson-boson interactions. Our theory, assuming a superfluid phase, is not applicable in this regime. Figure 5.8 shows the displacement of the impurity as a function of the lattice tilt $\hbar\omega_B/J_b$ for the interaction strengths $U_b/J_b = U_{ab}/J_b = 1$ at different times $t$. We see that the NDC peak and the $|\Delta\ell| = 1$ phonon resonance are not noticeably affected by the presence of strong interactions. They both seem to be robust features of the atomic current. To understand why, we will go into more depth on the strongly-interacting limit in the next Sec. 5.2.

### 5.1.8 Discussion

In our analytical and numerical investigation of phonon-assisted atomic currents along a tilted potential we have shown that atomic mixtures in optical lattices lend themselves naturally to investigate non-equilibrium transport phenomena present in solid state systems. In more detail, we have formulated an effective model for an impurity-boson mixture describing the two species in terms of Bogoliubov phonons and Wannier-Stark states, respectively, with a generic impurity-phonon interaction of an identical type to the one encountered in solids.

We have studied the dependence of the atomic current on the lattice tilt from first principles and found that our model accommodates NDC and phonon resonances. To demonstrate that these features are observable by using cold atoms in the context of a finite size system and a realistic measuring procedure, we have calculated the atomic current numerically by using the TEBD algorithm including the full dynamics of both the bosons and the impurities. Our numerical results show that resonances due to the Van Hove singularity at the boundary of the phonon band are robust phenomena. They occur over a wide range of system parameters despite broadening,

which might be increased by finite temperature effects [130, 171] not directly taken into account in our simulations.

## 5.2 Impurity transport as a probe for a strongly-correlated Bose gas

In the previous section we studied the effects of an environment's excitation spectrum on the current of an impurity travelling through it, where dissipating energy into this environment is the only way for a non-zero current to arise. Specifically we saw that a divergence in the density of environmental excitations results in sharp peaks in the dependence of the current on forcing. In this section, we turn the idea on its head. Rather than using a prior knowledge of the environment to predict the impurity current, we aim to use the impurity current to tell us about the environment, i.e., we consider the effectiveness of using the impurity as a probe.

We discuss generally what information can be revealed about the environment non-destructively by first allowing the impurity of the previous section to interact with it and then observing the impurity dynamics. For example, we demonstrate that the bandwidth and the gap in the excitation spectrum of the environment can be determined by analysing the transport of the impurity. Following this, using simulations, we show this for the specific environment of a one-dimensional Bose lattice gas. This time, however, we move away from the superfluid limit and consider strongly-interacting bosons up to the Mott insulator and Tonks-Girardeau limits whereupon the bosons exhibit fermionic characteristics [198]. In the case of a commensurately filled Bose gas the impurity propagation provides a clear signature of the superfluid to Mott insulator transition in the form of a sharp quench of the expected current. The strong dependence on the filling allows it to deter-

mined, in principle non-destructively, whether the bosons have been prepared in a commensurately filled state, as desired in some applications of quantum information processing.

Our work then adds to other methods for probing cold atoms in optical lattices. For example, the excitation properties of a Bose gas in an optical lattice can be analysed using Bragg spectroscopy via two photon scattering [199] and lattice depth modulation [99, 200] — the experiment in Ref. [99] having been later successfully simulated in Refs. [201, 129]. Studying the transport of a Bose gas after imparting a crystal momentum on it using a magnetic field [200] or tilting the lattice [202] has allowed experimentalists to successfully reveal signatures of phase transitions or determine excitation spectra, respectively. A Bose-Einstein condensate has been used as a source of matter waves to infer the spatial properties of atoms in a Mott insulator through Bragg diffraction [203]. Recently an impurity has been used in experiments as a coherent probe for large quantum systems in order to extract impurity-boson collision parameters [16, 15]. There have also been proposals to use impurities to extract work functions [160, 161] or measure temperature [18, 162] among other things [19, 158, 21, 159].

This work also extends, to the strongly-interacting limit, the large body of theoretical work on the transport of an impurity (partly subject to static forcing) through a superfluid [204, 23, 153, 130, 154, 2, 205], which we discussed in the previous section. Recently impurity transport in a strongly-interacting background was realised by experiment, using gravity to provide a static force for the impurities [13, 206]. The transport of an impurity through a Tonks-Girardeau gas has very recently (after the completion and publication of our work) been studied theoretically in Refs. [207, 22].

### 5.2.1 Transitions between crystal momentum states

We begin by considering the way in which environmental properties manifest themselves in the current of the impurities, approaching the problem using both the semi-classical and Wannier-Stark pictures.

First we use the semi-classical approach where the distribution of impurities in crystal momentum space $P_k$ obeys the Boltzmann equation [96]

$$\frac{\mathrm{d}P_k}{\mathrm{d}k} = \frac{a}{\omega_B}\int \mathrm{d}k' \left[W_{kk'}P_k - W_{k'k}P_{k'}\right], \tag{5.19}$$

which is valid for infinite homogeneous systems of impurities in the stationary state. Here $W_{kk'}$ is the rate of incoherent scattering between crystal momenta $k$ and $k'$. The average velocity of each impurity is then the expected value of the group velocity of wave-packets distributed according to $P_k$,

$$\langle v\rangle = \int \mathrm{d}k P_k v_g(k), \tag{5.20}$$

where the group velocity for a particular crystal momentum is defined in Eq. (5.4). Several choices of scattering rates lead to the expected velocity $\langle v\rangle$ of the impurity exhibiting exactly or approximately an Esaki-Tsu dependence [180, 171].

For small enough tilts the crystal momentum will be a very slowly changing quantity and hence for weak interactions with the environment we may approximate the incoherent scattering rates using Fermi's golden rule as

$$W_{kk'} = \frac{2\pi}{\hbar}\sum_{\nu_b}\left|\langle k' \,|\, \langle \nu_b \,|\, \hat{H}_{ab} \,|\, k\rangle \,|\, 0_b\rangle\right|^2 \delta\left(E_{\nu_b} - E_{0_b} - 2J_a[\cos ka - \cos k'a]\right). \tag{5.21}$$

Here we have assumed the environment to be sufficiently large and cold so that before scattering we may take it to be in the ground state $|\, 0_b\rangle$ of energy $E_{0_b}$. The result of

the interaction of the impurity with the environment, for which the Hamiltonian is given by $\hat{H}_{ab}$, is to scatter the environment to some excited state $|\,\nu_b\rangle$ with energy $E_{\nu_b}$. Here $|\,k\rangle$ is the Dirac representation of the state with crystal momentum $\hbar k$.

For the case of impurities coupled to a many-body environment via a density-density interaction, see Eq. (5.9), the corresponding crystal momentum scattering rates are given by [208]

$$W_{kk'} = \frac{2\pi U_{ab}^2}{\hbar M a} S(k-k', -2J_a[\cos ka - \cos k'a]), \tag{5.22}$$

where $S(k,\omega)$ is the zero-temperature dynamic structure factor of the environment

$$S(k,\omega) = \sum_{\nu_b} |\,\langle \nu_b \,|\, \hat{\rho}_k \,|\, 0_b\rangle\,|^2 \delta(E_{\nu_b} - E_{0_b} - \hbar\omega), \tag{5.23}$$

and $\hat{\rho}_k$ is the Fourier transform of the environment's particle density operator. Hence the impurity current and, more directly, the crystal momentum distribution at small tilts are governed by the dynamic structure factor of the environment. Note that this is quite general. It does not depend on the form of environment, only the form of the interaction.

### 5.2.2 Transitions between Wannier-Stark states

At larger lattice tilts it is more convenient to work in the basis of Wannier-Stark states and to consider the incoherent transitions between these, as we did above in Sec. 5.1.4. There we argued that for an infinite homogeneous system the average velocity would be $\langle v\rangle = \sum_{\ell'} a(\ell'-\ell)W_{\ell\ell'}$, where $W_{\ell\ell'}$ is the incoherent hopping rate from the Wannier-Stark state centred at sites $\ell$ to that at $\ell'$. Again, for weak coupling to the environment, this time making no assumptions about the form of this environment or its interaction with the impurities, the transition rates $W_{\ell\ell'}$ are

given by Fermi's golden rule as

$$W_{\ell\ell'} = \frac{2\pi}{\hbar} \sum_{\nu_b} \left| \langle \ell' \,|\, \langle \nu_b \,|\, \hat{H}_{ab} \,|\, \ell \rangle \,|\, 0_b \rangle \right|^2 \delta \left( E_{\nu_b} - E_{0_b} + [\ell' - \ell]\hbar\omega_B \right). \tag{5.24}$$

Here $|\,\ell\rangle$ is the Dirac representation of the Wannier-Stark state centred at lattice site $\ell$.

From this we immediately see two effects that we expect to be visible in the impurity current. First, consider the case where the environment has a gap $G$ in its excitation spectrum such that $E_{\nu_b} - E_{0_b} \geq G$. This means that the delta function in Eq. (5.24) will be zero unless $\hbar\omega_B \geq G/(\ell - \ell')$ and hence the contributions to the current due to hopping between Wannier-Stark states shifted by $|\Delta\ell|$ sites will only be made at tilts above $\hbar\omega_B = G/|\Delta\ell|$. Second, if the excitation spectrum of the environment has a finite bandwidth such that there are no states above an energy $B$, i.e., $E_{\nu_b} - E_{0_b} \leq B$, then we expect $|\Delta\ell|$-site contributions to the current to be suppressed at tilts above $\hbar\omega_B = B/|\Delta\ell|$. Hence the deviation of the dissipation process from the relaxation-time approximation results in the dependence of the impurity current on $\hbar\omega_B$ exhibiting sharp features that clearly reveal information about the environment; we can infer these environmental properties by observing impurities.

### 5.2.3 Demonstration for a Bose gas environment

Next we seek to demonstrate the general principles of the previous two subsections by simulating the dynamics of an impurity for a specific environment. As in Sec. 5.1, we describe our environment by the Bose-Hubbard model of Eq. (5.6) the relevant properties of which are given in Sec. 2.3.2. As stated previously in this chapter, in the superfluid regime $U_b/J_b \ll 1$ the Bose gas supports excitations in the form of Bogoliubov phonons, whose excitation spectrum is given in Eqs. (5.8). From

this we find that the density of phonon states is zero for energies larger than the bandwidth $B = 4J_b\sqrt{1 + U_b n_b/2J_b}$. In the opposite limit $U_b/J_b \to \infty$ the bosons map to the same number of non-interacting identical fermions with energy spectrum $\varepsilon_k$ given in Eq. (5.8b). Hence in this regime the Bose gas supports particle-hole excitations for $n_b < 1$ within the single band of bandwidth $B = 4J_b$. For the commensurately filled case of $n_b = 1$ an increasing interaction strength $U_b/J_b$ takes the gas through the superfluid to Mott insulator transition, occurring at $U_b/J_b \approx 3.37$ in the thermodynamic limit. In the Mott insulating regime the spectrum is gapped, where for $U_b/J_b \gg 1$ the gap is $G \approx U_b$.

We will see that all of the above properties are apparent from the impurity current $\langle v \rangle$ , which we calculate using the same setup and numerics as in Secs. 5.1.5 and 5.1.6. We now present our results.

### Superfluid and Tonks-Girardeau regimes

Our starting point is the propagation of the impurity through a Bose gas in the superfluid regime with incommensurate filling, similar to what we analysed in Sec. 5.1.7. A typical dependence of the displacement of the impurity $\langle \hat{r}_a \rangle$ on the lattice tilt $\hbar\omega_B$ is given by the red circles in Fig. 5.9(a). As before, this dependence has the characteristic features of the Esaki-Tsu result, shown in Fig. 5.2(b), namely an ohmic region followed by NDC. However the superfluid Bose gas is a prime example of how the value of the bandwidth of its excitation spectrum $B$ manifests itself in the dependence of the current of the impurity. The phononic excitation spectrum has a bandwidth $B = 4J_b\sqrt{1 + U_b n_b/2J_b}$ and, as expected, we indeed see sharp decreases in the current at lattice tilts $\hbar\omega_B = B/|\Delta\ell| \approx 4J_b/|\Delta\ell|$ for integer $|\Delta\ell| \geq 1$.

We expect to see the bandwidth dependent features even with strong boson-boson interactions. This is most clear in the extreme case of infinitely strong interactions $U_b/J_b \to \infty$, i.e., in the Tonks-Girardeau limit. The bosons map to identical

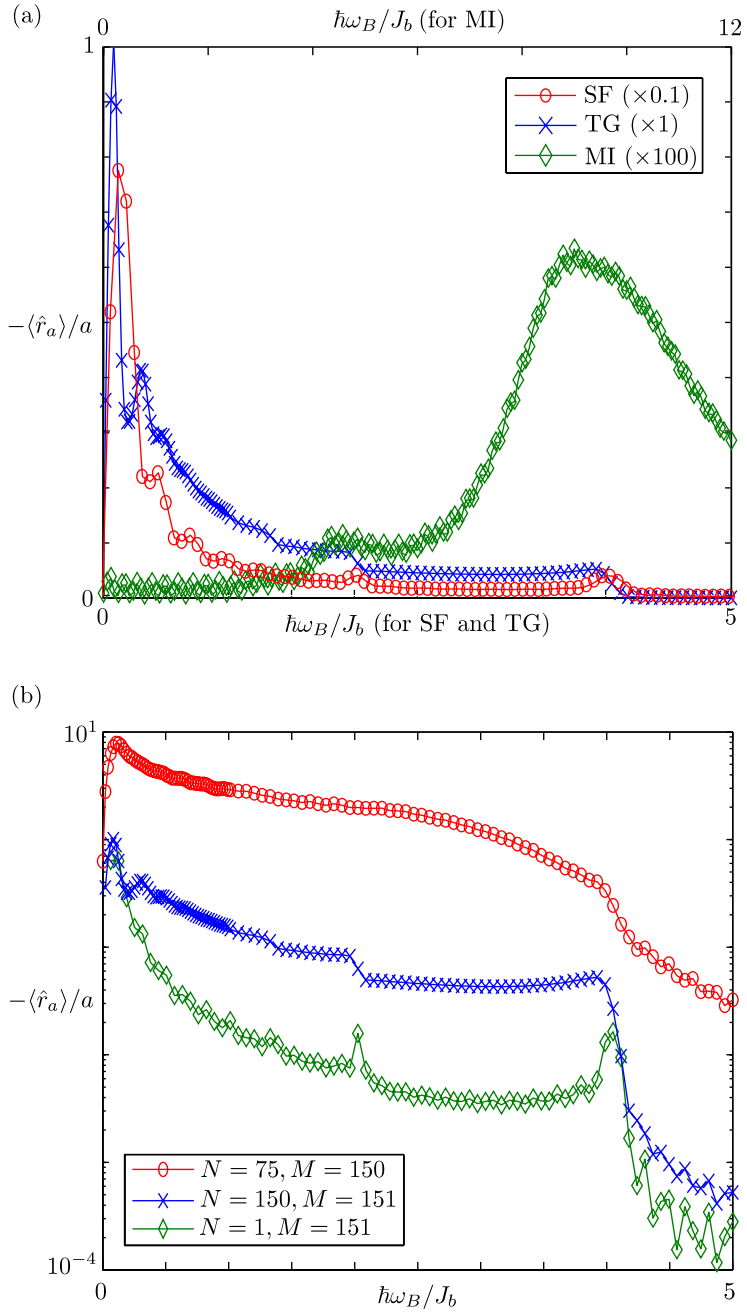


**Figure 5.9**. Determining the bandwidth and gap from the dependence of displacement on tilt. (a) The dependence of the displacement of the impurity $\langle\hat{r}_a\rangle$ at $t_f$ as a function of tilt $\hbar\omega_B$ for three regimes of the Bose gas: the superfluid regime $U_b/J_b = 0.1$ with $N = 50$, $M = 101$ (SF - displacement scaled down by a factor of 0.1); the Tonks-Girardeau limit $U_b/J_b \to \infty$ with $N = 150, M = 151$ (TG); and the Mott insulating regime $U_b/J_b = 10$ with $N = M = 101$ (MI - displacement scaled up by a factor of 100). (b) The same quantity is plotted for several bosonic filling factors in the Tonks-Girardeau limit, demonstrating that the qualitative shape and the positions of the bandwidth suppressions are independent of the filling (so long as it is incommensurate) of the Bose gas.

non-interacting fermions with bandwidth $B = 4J_b$ and so the creation of particle-holes pairs across the Fermi surface provides low-energy excitations. Accordingly we expect a large impurity displacement at low lattice tilts and sharp drops in propagation for tilts above integer divisions of the bandwidth $4J_b/|\Delta\ell|$, in agreement with Fig. 5.9(a).

We find the same behaviour in the Tonks-Girardeau limit for many different bosonic filling factors, as shown in Fig. 5.9(b). The half-filled bosonic system gives rise to a peak impurity current that is an order of magnitude above that for the cases where the bosonic lattice is filled by a single or $M-1$ bosons. This can be attributed to there being a greater number of excitations available in the half-filled bosonic system; using the fermionic mapping we find there are approximately $M^2/4$ single quasi-particle excitations available in the half-filled case as compared to $M-1$ when there is only one boson or $M-1$ present. This reasoning suggests that the peak displacement may be used as a measure of the filling of the bosonic system.

For intermediate interaction strengths we have simulated $M = 25$ site systems over a range of boson-boson interaction strengths $U_b/J_b$ (not shown in this thesis) and found no qualitative change in the behaviour of the dependence of current on tilt. However, as shown in Fig. 5.10, the peak current decreases with increasing interaction strength.

### Superfluid to Mott insulator transition

We now turn to the case of a commensurately filled Bose gas and the continuous superfluid to Mott insulator transition. Since the motion of the impurity though the superfluid at low lattice tilts is a result of the ability of the bosons to accept low-energy excitations we expect impurity propagation to be highly suppressed in the Mott insulator phase, which has a gap in the excitation spectrum of $G \approx U_b$.

The effect of this signature on the propagation can be seen in Fig. 5.9(a), where

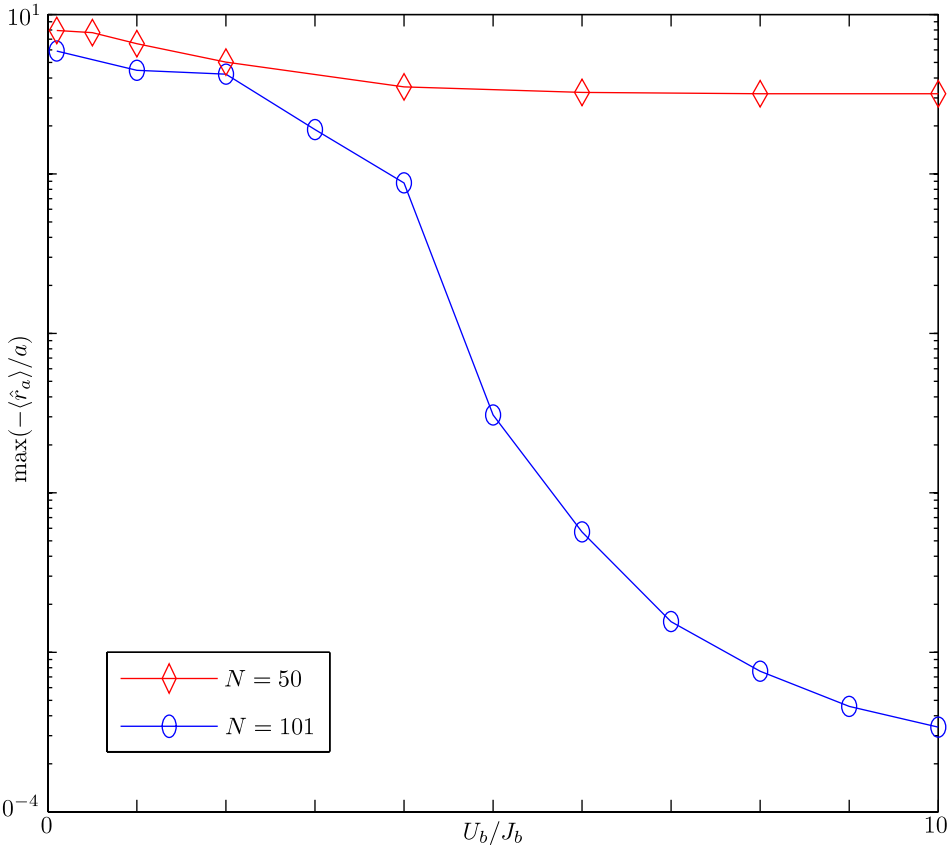


**Figure 5.10**. The suppression of average drift in the strongly-interacting regimes. Here we show the peak displacement (maximised over all tilts) as a function of interaction strength $U_b/J_b$ for commensurate and incommensurate fillings of a 101-site bosonic lattice.

the displacement has a peak at $\hbar\omega_B \approx U_b$. There is also a second peak at $\hbar\omega_B \approx U_b/2$ resulting from second order processes involving two lattice-site jumps, similar to what is seen using Bragg spectroscopy in Ref. [201] and when tilting the bosonic lattice in Ref. [202]. As in those cases the non-zero $J_b$ results in a broadening of the peaks. This behaviour confirms the predictions in Sec. 5.2.2 for the transport of an impurity in a gapped environment. This drastic change from the behaviour predicted by Esaki and Tsu also tells us that the relaxation-time approximation is not at all valid for the interaction of an impurity with a Mott insulator.

Deep in the Mott insulating regime, the motion of the impurity caused by low-frequency excitations is completely suppressed and the ground state of the environment is well approximated by a unit-filled Fock state. The impurity Bloch oscillates as it has no means of dissipating energy into the Bose gas, hence there is no net current across the lattice. As a result we get a startlingly clear signature of the superfluid to Mott insulator transition in our dynamics in the form of a quench of more than 4 orders of magnitude in the propagation of the impurity. This is shown in Fig. 5.10.

As well as the ability to probe the superfluid to Mott insulator transition using the impurity, this quench provides a non-destructive way of establishing whether the bosonic system is commensurately filled or not. If our system is strongly interacting $U_b/J_b \gtrsim 4$ there are several orders of magnitude separating the currents for an impurity in the commensurate and incommensurate cases, shown in Fig. 5.10. Many implementations of quantum information processing involve creating a commensurately filled system. Using an impurity as a probe is one way in which this commensurability could be checked non-destructively.

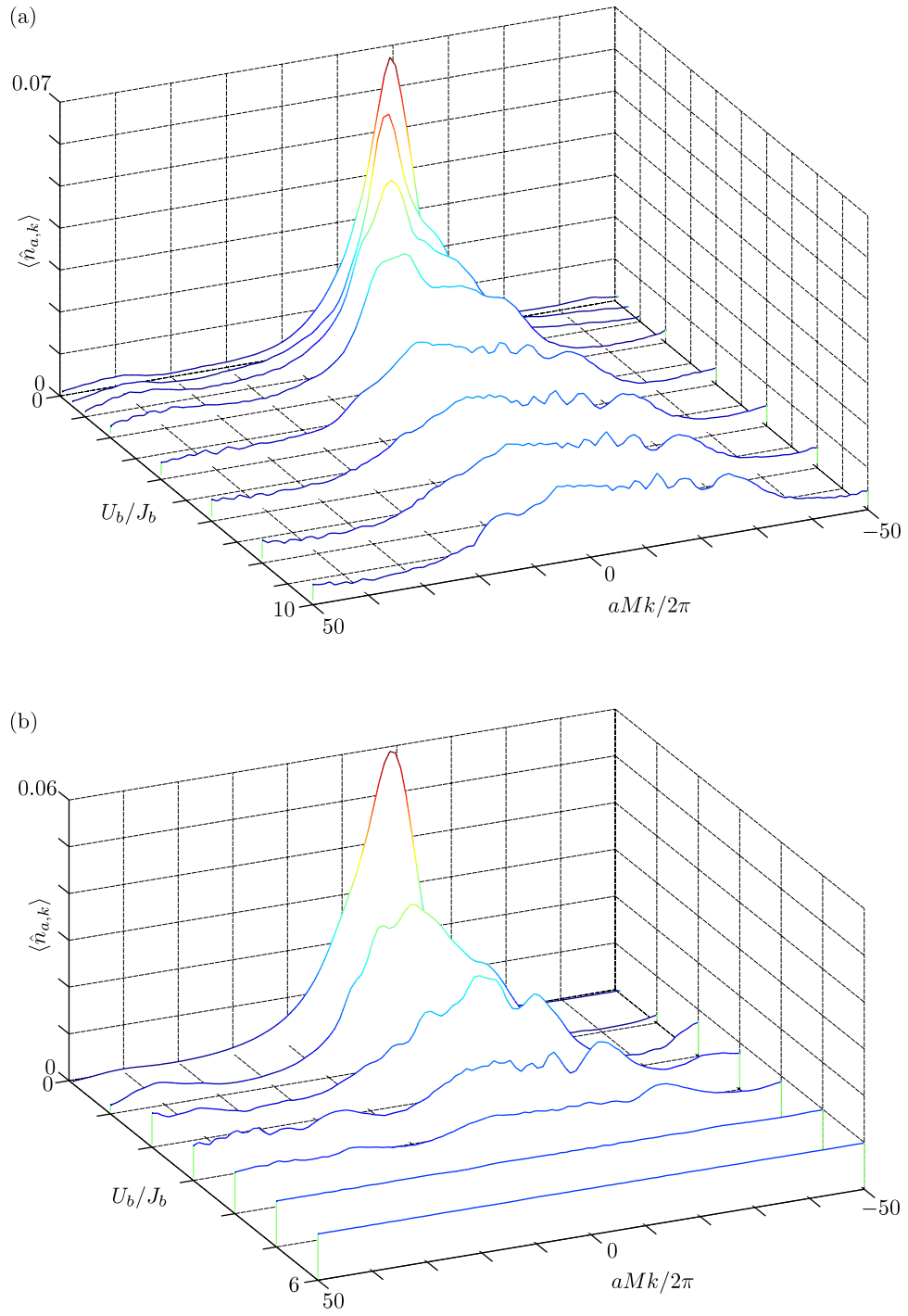


**Figure 5.11**. Crystal momentum distribution of the impurity after travelling through the Bose gas. Here $\hbar\omega_B/J_b = 0.063$, $\hat{n}_{a,k}$ is the number operator for an impurity with crystal momentum $\hbar k$, and $M = 101$. (a) The crystal momentum distribution of the impurity at $t_f$ for a range of interaction strengths $U_b/J_b$ and incommensurate filling $N = 50$. (b) The same as (a) but for commensurate filling $N = 101$.

**Crystal momentum distribution of the impurity**

Alternatively we can gain information about the impurity by looking at its crystal momentum distribution, which is directly accessible in time-of-flight experiments. With the impurity prepared in a Wannier state, each crystal momentum state is equally occupied. However we expect interactions with the Bose gas to alter this crystal momentum distribution, resulting in the drift seen in the previous two subsections. The current $\langle v \rangle$ can be calculated from the distribution in the semi-classical picture using Eq. (5.20).

As can be seen in Figs. 5.11(a) and 5.11(b) the result of interactions with the Bose gas is the enhancement of the distribution near the centre of the Brillouin zone. Moreover the interaction gives rise to an asymmetry with the distribution skewed towards negative crystal momentum, necessary for the impurity to have a current.

Our simulations show that a well-defined crystal momentum emerges in the superfluid regime by the time $t_f$, supporting the accuracy of the semi-classical interpretation. However, for stronger bosonic interaction strengths the scattering results in a wide distribution with a negative crystal momentum shoulder. By $U_b/J_b \approx 10$ the distribution already approximates its $U_b/J_b \to \infty$ value. In the case of the Mott insulating regime in Fig. 5.11(b) it is unchanged from its initially flat distribution. Once again, the crystal momentum distribution of the impurity provides a signature of the transition (cross-over) between the superfluid and Mott insulator (Tonks-Girardeau) regimes.

The stationary state crystal momentum distribution is related to the structure factor of the Bose gas, see Eqs. (5.19) and (5.22), but a quantitive comparison with our simulations is difficult since in general they do not reach the stationary state for such low tilts. Since this limitation is due to the finite time of our simulations it would also restrict an experimental realisation.

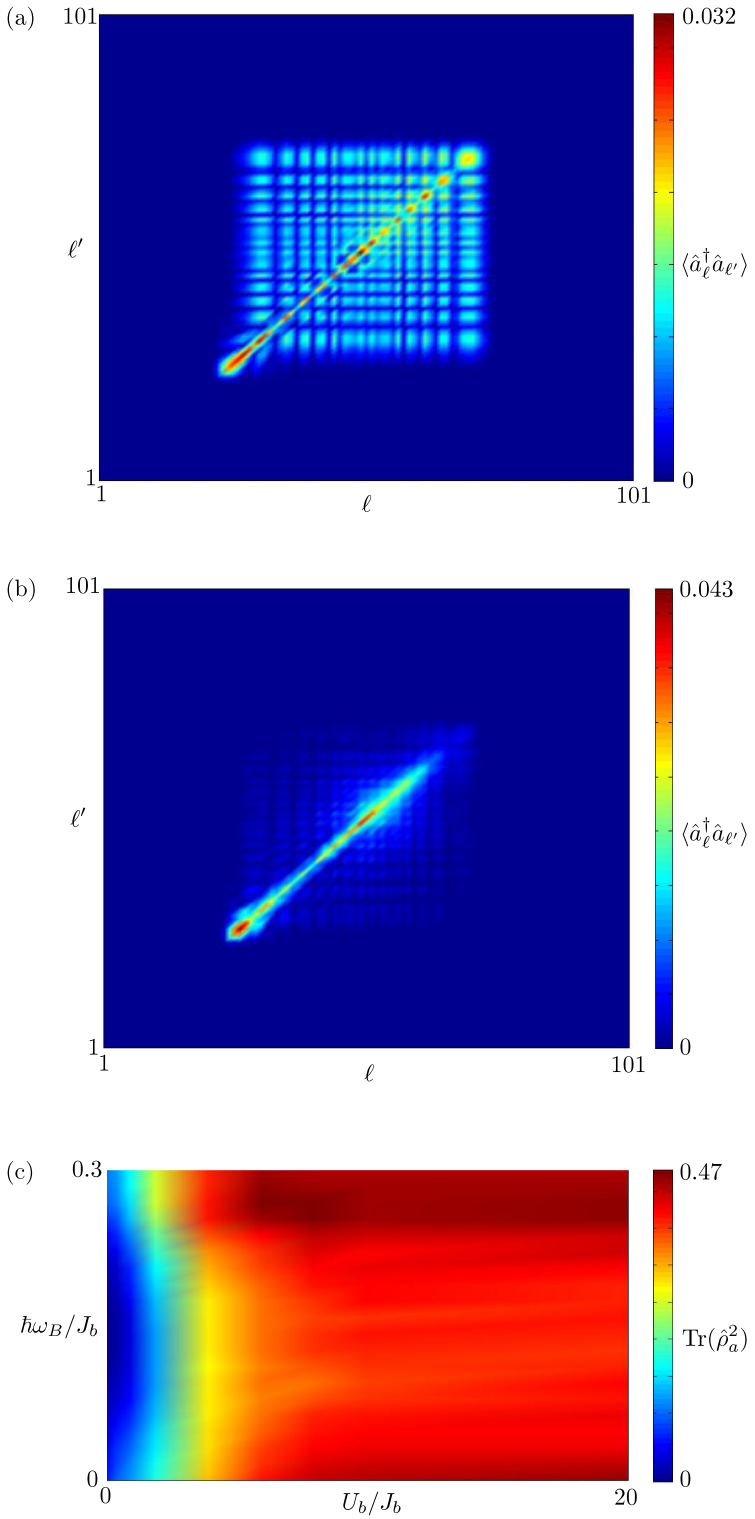


**Figure 5.12**. Decoherence of an impurity travelling through an incommensurately filled system $N = 50, M = 101$. (a) Absolute values $|\langle \hat{a}^\dagger_\ell \hat{a}_{\ell'} \rangle|$ of the elements of the reduced density-matrix for the impurity with $U_b/J_b = 2$ and $\hbar\omega_B/J_b = 0.1$, calculated at $t_f$. (b) The same as (a) but for $U_b/J_b = 0.1$. (c) Also at $t_f$ we have plotted the purity of the density-matrix $\mathrm{Tr}(\hat{\rho}_a^2)$ as a function of tilt and interaction strength.

**Reduced density-matrix of the impurity**

So far we have focused on the evolution of the spatial and crystal momentum distributions of the impurity. However, more information is contained in its reduced density-matrix $\hat{\rho}_a$. Evaluating the elements of the density-matrix in the Wannier basis $\langle \hat{a}_\ell^\dagger \hat{a}_{\ell'} \rangle$ may tell us more about the effect of boson-impurity interactions on the impurity. Another quantity that can be calculated from the reduced density-matrix is the purity $\mathrm{Tr}(\hat{\rho}_a^2)$, which takes the maximal value 1 if $\hat{\rho}_a$ describes a pure state and decreases for increasingly mixed states. The initial state of the total system is a product of pure states for the impurity and bosons, and since we consider coherent evolution it remains pure. However, through interactions of the impurity with the Bose gas the two subsystems are entangled resulting in a loss of purity for both subsystems. Our bosonic environment is sufficiently large that for most parameter values revivals are not observed and so we use purity to quantify the decoherence of the impurity.

To demonstrate the behaviour of the density-matrix $\hat{\rho}_a$, we have plotted in Figs. 5.12(a) and 5.12(b) the absolute values of its Wannier basis elements at $t_f$ for both intermediate and weak bosonic interaction strengths, respectively. In Fig. 5.12(a) we observe a symmetric coherent interference pattern in the off-diagonal elements as the impurity undergoes Bloch oscillations. Additionally, density-density interactions with the bosons result in a slight decay of the off-diagonal coherences and an asymmetric distribution along the diagonal with a higher occupancy of the lower energy lattice sites. This decoherence is much more pronounced for a superfluid environment, as shown in Fig. 5.12(b), and the resulting density-matrix is nearly diagonal.

To make this relationship between the decoherence of the impurity and the internal interaction strength of the environment clearer, Fig. 5.12(c) shows the purity

of the reduced density-matrix of the impurity as a function of both tilt $\hbar\omega_B$ and interaction strength $U_b/J_b$. As can be seen in the figure, the coherence of the impurity is a clear signature of the cross-over between the superfluid and Tonks-Girardeau regimes. Moreover the values of $U_b/J_b$ at which the most significant changes occur are around the same values where the transition occurs in a commensurate Bose gas. We also calculated the entanglement entropy between the impurity and Bose gas (not shown in this thesis), which showed very similar behaviour to purity. If the Bose gas is commensurately filled (also not shown here) the transition between weak and strong interaction strengths makes an even clearer impact on the purity since deep in the Mott insulator phase the impurity does not become at all entangled with the Bose gas.

Finally, comparing Figs. 5.12(c) and 5.6(d) we find that the regimes of high transport coincide with those of the smallest purity; the very same scattering processes that lead to the propagation of the impurity along the lattice also cause its decoherence.

### 5.2.4 Discussion

In this section we have extended our study of the transport of an impurity confined to a tilted optical lattice through a Bose lattice gas by considering strong boson-boson interactions. Our focus was on how deviations to the Esaki-Tsu shape depend on properties of the Bose gas, and thus how information about the gas can be extracted by examining the dynamics of the impurity. In particular we discussed how the impurity could be used to probe the commensurateness of the system, and obtain the gap and bandwidth of its excitation spectrum. Interestingly, these quantities can be probed via the impurity without needing to measure the bosons themselves, i.e., in principle non-destructively.

Specifically we found that the Esaki-Tsu dependence of the current on tilt, along with resonances, is qualitatively obeyed by an impurity travelling through an incommensurately filled bosonic system of any interaction strength including the Tonks-Girardeau limit. Contrasting this a very different dependence was found in the Mott insulating regime. We also analysed the transport of the impurity by considering its crystal momentum distribution, showing that the semi-classical picture of an impurity of well-defined crystal momentum emerges after a short time for bosons in the superfluid regime.

## 5.3 End of chapter discussion

We have studied the transport of an impurity through a Bose lattice gas. While some analytical results were obtained in simplifying limits, a full numerical simulation was required in order to consider the strongly interacting regime, where strongly here refers to both the impurity-boson and boson-boson interaction.

It is worth briefly considering the benefits of performing such a full simulation rather than, for example, a mean-field simulation. A mean-field treatment of a Bose gas neglects any entanglement between the lattice sites (most easily seen with the Gutzwiller ansatz approach [209, 210]) and assumes small fluctuations of quantities around their mean values (most easily seen from the site-decoupling approach [211]). Despite these limitations it is known to quite accurately capture the behaviour of the Bose-Hubbard model in high dimensions (it is exact in infinite dimensions), for example, predicting the value of $U_b/J_b$ at the superfluid-Mott insulator phase transition in three dimensions to within about 20%.

However, for the one-dimensional Bose-Hubbard model we consider, quantum fluctuations are of greater importance and the value of $U_b/J_b$ at the transition point is greatly overestimated (by about 250%). The mean-field treatment also misses

some essential physics. For example a mean-field analysis predicts that on-site boson number fluctuations are zero in the Mott insulator regime, though this is of course only strictly true for infinite $U_b/J_b$. A mean-field approximation also results in uniform off-diagonal elements of the single-particle density-matrix (only valid deep in the Mott insulating regime where they are zero, and in the superfluid limit when they are non-zero). We might then expect differences even between the qualitative descriptions of a mean-field and full simulation for the intermediate values of $U_b/J_b$ we have considered in this chapter.

For the rest of this section we will examine the feasibility of experimentally realising the physics presented in this chapter. To begin, let us analyse the validity of the zero temperature approximation that we have implicitly made when performing our simulations. For this to be accurate the temperature $T$ must be low enough that $k_BT$ is small compared to the other energy scales $J_a$, $J_b$, $U_b$, $U_{ab}$ and $\hbar\omega_B$. In Sec. 2.3.1 we found the temperatures corresponding to $J_b$ and $U_b$ could be around $E_R/10k_B$. For a typical atomic mass $m_b = 50$ amu and wavelength $\lambda = 1000$ nm, the recoil energy is given by $E_R/k_B \approx 200$ nK. For the same reasons as given for $J_b$, the value of $J_a$ will be on the same order. The interaction energy $U_{ab}$ has roughly the same magnitude as $U_b$ when the inter- and intra-species scattering lengths have similar values. For example, the intra-species scattering lengths of $^{87}$Rb and $^{41}$K are $99a_B$ [212] and $65a_B$ [213], respectively, while their inter-species scattering length is $163a_B$ [214, 215], where $a_B$ is the Bohr radius.

What remains is to show that the Bloch energy $\hbar\omega_B$ can be brought up to this same order of magnitude. Perhaps the most flexible way of creating a tilt is to chirp the frequencies of the lasers producing the optical lattice potential [179]. This accelerates the optical lattice potential at an amount $\tilde{a}$. In the frame of the standing wave this appears as an additional term $m_b\tilde{a}r$ in the impurity potential $v_a(r)$, which leads to a Bloch energy $\hbar\omega_B = m_b\tilde{a}a$. In Ref. [179] accelerations $\tilde{a} \approx 1500$ ms$^{-2}$

several hundred times gravitational acceleration were observed. Using $a = \lambda/2$ and the typical parameters given previously, this would result in tilts $\hbar\omega_B \approx 24E_R$. Therefore creating tilts $\hbar\omega_B$ the order of the other energies is realistic.

Possibly the largest barrier for realising this setup experimentally is that, apart from the periodic lattice, we have assumed a flat bottomed potential to describe the confinement of the bosons (though see Ref. [216] for an example of a box trap realised in an experiment). We expect our system to share its bulk behaviour with the experimentally important case where the bosons experience an additional harmonic trap, provided that the trap is sufficiently shallow. To see this, let us briefly discuss the effect of this shallow harmonic trap.

Take the potential due to the confinement (not including the lattice potential) to have the form $v_b(r) = m_b\Omega_b^2 r^2/2$. A typical trap frequency in a cold atom experiment might be $\Omega_b = 2\pi \times 10$ Hz [217]. We will work in the grand canonical ensemble and make the local density approximation (LDA) [218, 219, 220]. The LDA consists of replacing the bosonic chemical potential $\mu_b$ at each point $r$ by an effective position-dependent chemical potential $\mu_b(r) = \mu_b - v_b(r)$. The properties of the Bose gas at each position $r$ will be those of a homogeneous system with the chemical potential $\mu_b(r)$.

The LDA is expected to hold on occasions where the trapped condensate would be well-described by the Thomas-Fermi approximation in the absence of a lattice [219], i.e., the density changes slowly compared to the healing length. In our case this is satisfied provided the harmonic oscillator length $w_b = \sqrt{\hbar/m_b\Omega_b}$ obeys the relationship $w_b \gg Na$ [218]. For the typical parameters above we find that $w_b \approx 3$ $\mu$m and so a weaker potential is needed for this criterion to be strictly met.

Nevertheless let us now, in the framework of the LDA, say something about how the three regimes focused on in this chapter could be realised with a harmonic trap. These regimes are the weakly-interacting superfluid, the strongly-interacting incom-

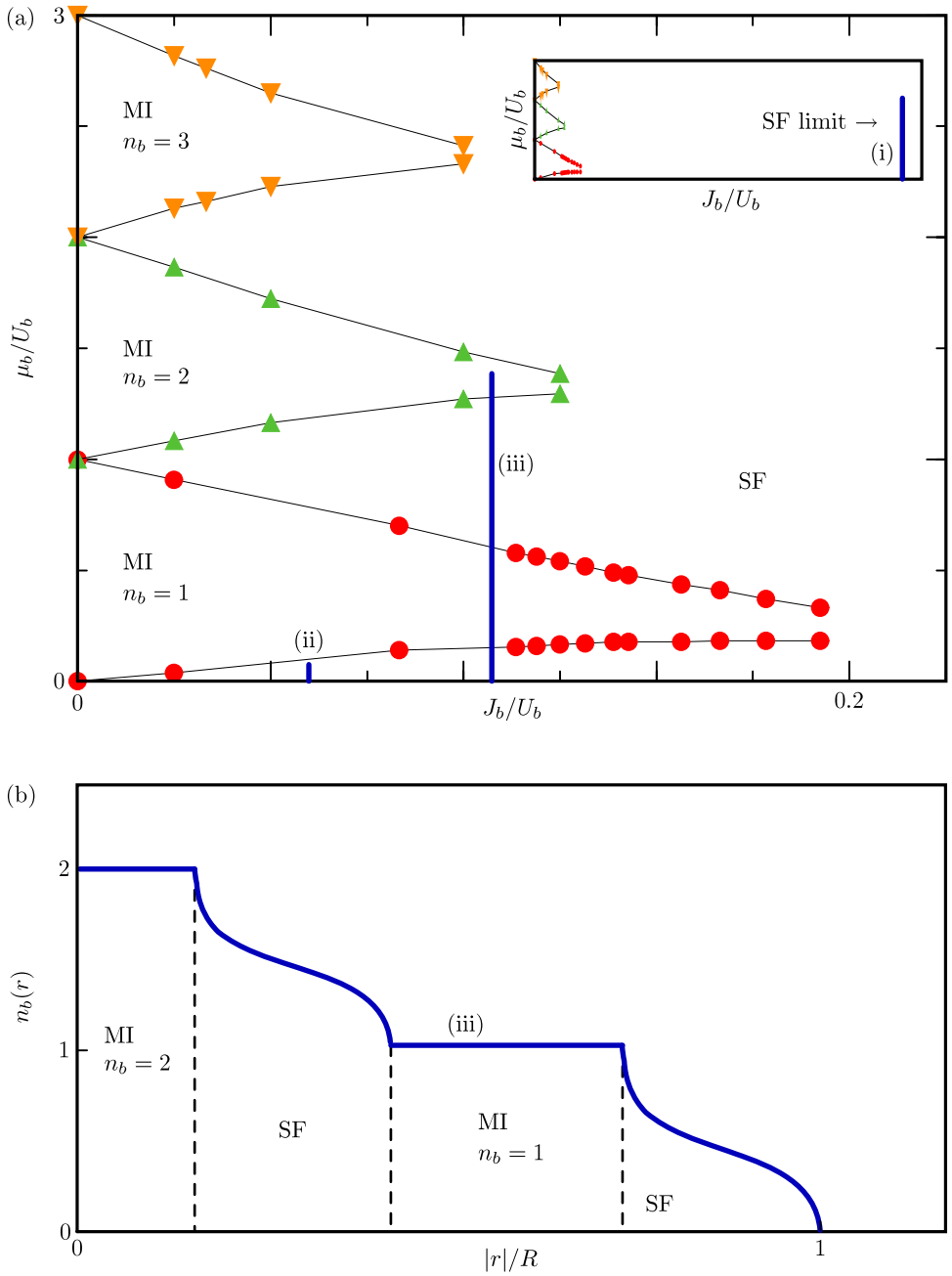


**Figure 5.13**. The $T = 0$ phase diagram of the Bose-Hubbard model. (a) The phase diagram with points on the lobe calculated using Monte Carlo simulations (this part of the figure is borrowed from Ref. [221] with the permission of the authors). Also drawn are three lines of constant $J_b/U_b$, showing the phases of a trapped Bose gas at different positions within the LDA. Each line corresponds to a different chemical potential $\mu_b$ and interaction strength $J_b/U_b$. The top of the lines correspond to the centre of the trap. Line (i) corresponds to a superfluid, (ii) to a strongly-interacting incommensurately filled system and (iii) to a Bose gas that alternates between a Mott insulator and a superfluid at different positions in the trap. (b) A sketch of the density as a function of position in the trap, corresponding to line (iii) in figure (a).

mensurately filled ($n_b < 1$) Bose gas, and the Mott insulator ($n_b = 1$). To guide the discussion we present the $J_b/U_b$ vs. $\mu_b/U_b$ phase diagram of the homogeneous Bose gas in Fig. 5.13(a) [221]. In the LDA the phase of the Bose gas as a function of $r$ corresponds to a vertical line of constant $J_b/U_b$ and decreasing $\mu_b/U_b$. The top of this line corresponds to the centre of the trap.

First, consider the case where $U_b/J_b \ll 1$ such that the vertical line is always in the superfluid region. At all positions the bosons form a superfluid. Further choose the filling at the centre to be small enough and the height of the line to be small enough that $U_b n_b(0)/J_b \ll 1$. This corresponds to the regime where the bosons form a weakly-interacting superfluid $U_b n_b(r)/J_b \ll 1$ for all $r$. Since the position of the phonon resonances $4\hbar\omega_B/J_b = |\Delta\ell|^{-1}\sqrt{1 + U_b n_b/2J_b}$ and more generally $\langle v \rangle$ (except for its prefactor $v_0$) depend on $n_b$ only through the small quantity $U_b n_b/J_b \ll 1$ the effects of the trap on these quantities will themselves be small.

Second, consider a much larger $U_b/J_b$. However let $\mu_b(r)/U_b$ be small such that all positions $r$ are expected to lie below the first Mott insulator lobe. See the line labelled (ii) in Fig. 5.13(a). From Fig. 5.13(a) it seems that $J_b/U_b \sim 0.05$ and $\mu_b(0)/U_b \sim 0.05$ satisfy these two conditions. The bosons will then be strongly interacting with a density profile $n_b(r) < 1$ varying over a radius $R = \sqrt{2\mu_b/m_b\Omega_b^2}$. The strongly-interacting limit has been reached in a harmonic trap experimentally [222, 223]. Further for the typical trapping frequency, mass and wavelength given above, $R \sim 20a$. In this way, an impurity could be immersed in an incommensurately filled strongly interacting Bose gas with a nearly constant filling, realising the simulations we presented in the previous section.

Last, if $\mu_b(0)$ is increased (with $U_b/J_b$ still large) then eventually a region of small $r$ will become a Mott insulator with $n_b(r) = 1$, while a superfluid $n_b(r) < 1$ will remain outside of this region. Increasing $\mu_b$ further still will lead to a density structure consisting of Mott insulating regions of integer filling $n_b = n_{\max}, n_{\max} -$

$1, \ldots, 0$ separated by superfluid regions of intermediate and spatially varying $n_b$ [224, 111, 225, 226]. This is sketched in Fig. 5.13(b). An impurity moving within a Mott insulator could be experimentally realised by initialising the impurity at a lattice site within one of these Mott insulating regions, provided the displacement of the impurity is smaller than the size of the region. To estimate the size of the regions, note that in the limit that $J_b/U_b \to 0$ the width of the superfluid regions vanishes and a 'wedding cake' distribution is obtained with the Mott insulator regions of filling $n_b = n$ beginning at positions $r_n = \pm\sqrt{2(\mu_b - nU_b)/m_b\Omega_b^2}$ [227]. The size $\sim \sqrt{U_b/m_b\Omega_b^2}$ of the regions for the typical parameters given above is $\sim 60a$. The bigger problem is that in order for the Mott insulating regions to still exhibit a gap $G \approx U_b$, excitations arising from particles leaving the Mott insulating region [228] must be of a energy greater than the particle-hole excitations within the Mott insulating region. We may estimate the energy of a particle leaving a Mott insulating region to be $\frac{1}{2}m_b\omega_b^2[(|r_n| + a)^2 - r_n^2] \sim \sqrt{U_b m_b \Omega_b^2}a$. Unfortunately for the typical values given above this is small compared to $U_b$ (about a tenth or a hundredth). To ensure that the creation of a particle-hole pair inside the Mott insulator region is the lowest energy excitation available a much larger spacing $a$ would be required or a trap that is more box-like, e.g., $v_b(r) \propto r^4$. Even if this condition is not met, it is expected that the effect of particles exiting the Mott insulating region will be small deep inside the region.

# Part II

# Stochastic processes

# Chapter 6

# Continuous-time Markov processes

In this second part we shift our focus from unitary quantum to strictly classical stochastic dynamics. Out-of-equilibrium dynamics in classical stochastic systems have attracted interest for two key reasons. Firstly, they exhibit a wealth of non-trivial phenomena, not observed in equilibrium systems, such as boundary-induced phase transitions [229] and infinite range correlations [230]. Secondly, classical stochastic systems far away from equilibrium have applications in describing a variety of driven diffusive processes such as vehicular traffic flow [231, 232, 233] and biological transport [234, 235, 236, 237]. There exists a range of analytical approaches with which to address such systems including matrix product-based analytical methods [238, 239] and the Bethe ansatz [240]. However there is no complete analytical framework for non-equilibrium systems, especially away from stationary states, which increases the importance of numerical techniques.

Our main aim will therefore be to develop and evaluate the effectiveness of numerical methods for simulating the dynamics of classical stochastic processes. Our particular novel contribution is to apply a matrix product state (MPS) numerical method, specifically the time-evolving block decimation (TEBD) algorithm (introduced for quantum systems in Sec. 3.2), to stochastic processes and to rigorously evaluate its performance in principle and in practice.

Rather than using well-understood methods to capture novel physical properties as we did in Part I, here we will instead focus more on the method itself and its performance, with the study of the physics playing a secondary role. We will establish whether the TEBD algorithm works for classical systems, how well it works and whether it is competitive with other numerical methods for simulating the dynamics of stochastic processes, in particular, Monte Carlo. The motivation is clear: numerics based on MPS methods have had a transformative effect on the simulation of quantum systems, so it is worth investigating whether they could similarly impact the economically and socially-important area of stochastic processes.

The mathematics underlying this work will be that of Markovian stochastic processes, which we therefore briefly introduce in this first chapter. Then in Ch. 7 we present some results regarding stochastic systems, namely transfer matrix solutions and Jarzynski's equality, which we will make use of in later chapters. In Ch. 8 we introduce the current state-of-the-art method for simulating Markovian processes, dynamical Monte Carlo (DMC). The final two chapters contain our original research. We adapt TEBD to simulate a paradigm transport model in Ch. 9 and show it to be successful. In Ch. 10 we further adapt TEBD and demonstrate an exponential speed-up with respect to system size, compared to basic DMC simulations, when calculating the expected value of an observable with a high variance.

Returning to the goal of this chapter, to introduce Markovian stochastic processes, we first describe a stochastic system at a single time in Sec. 6.1 and give the example of a system in a Gibbs distribution in Sec. 6.2. We then discuss how the system evolves in time in Sec. 6.3. This leads on to the concept of a (possibly non-equilibrium) stationary distribution, which we introduce in Sec. 6.4.

## 6.1 Configuration space and the probability vector

All stochastic processes take place in a configuration space. In this thesis we will be concerned with a configuration space that is a collection of configurations $\mathbf{i} = (i_1, \ldots, i_M)$, where local indices $i_\ell = 0, \cdots, d-1$ can each take one of $d$ values. Our notation is deliberately analogous to the description of the configurations of quantum systems in Sec. 3.2. A (continuous-time) stochastic process is then a random variable $\mathbf{s}$ that is a continuous function of time $t$. At each time $t$ this random variable takes is equal to a configuration $\mathbf{s}(t) = \mathbf{i}$ according to some probability distribution formed of real non-negative elements $P_\mathbf{i}(t)$ that sum to unity $\sum_\mathbf{i} P_\mathbf{i} = 1$. Again in analogy to quantum systems, we represent this distribution by the probability vector

$$|\, P(t)\rangle = \sum_\mathbf{i} P_\mathbf{i}(t)\, |\, \mathbf{i}\rangle\,, \tag{6.1}$$

where we again write $|\, \mathbf{i}\rangle = |\, i_1\rangle \cdots |\, i_M\rangle$. It follows that probability vectors $|\, P\rangle$ are contained in the positive orthant of the vector space $(\mathbb{R}^d)^{\otimes M}$ and are normalised according to the $L_1$-norm as $\|\, |\, P\rangle\, \|_1 = 1$. We call $|\, P(t)\rangle$ the state of the system at time $t$.

## 6.2 Gibbs distributions

Perhaps the most famous type of probability distribution arises as a consequence of assigning some energy $E_\mathbf{i}(\lambda)$ to each configuration $\mathbf{i}$. In anticipation of later chapters, we have introduced a parameter $\lambda$ on which these energies depend. The

energy assignment then specifies a Gibbs distribution

$$P_{\mathbf{i}}(\lambda) = \frac{\mathrm{e}^{-\beta E_{\mathbf{i}}(\lambda)}}{Z(\lambda)}, \tag{6.2}$$

with partition function $Z(\lambda) = \sum_{\mathbf{i}} \mathrm{e}^{-\beta E_{\mathbf{i}}(\lambda)}$ and inverse temperature $\beta$. We denote the corresponding probability vector by $|\, P(\lambda)\rangle$.

The Gibbs distribution has two physical interpretations [241]. It is the canonical distribution arrived at when the system is in thermal equilibrium with a reservoir at inverse temperature $\beta$. Alternatively, subject to the constraint of a fixed expected energy

$$\langle E(\lambda)\rangle = \sum_{\mathbf{i}} E_{\mathbf{i}}(\lambda) P_{\mathbf{i}} = -\frac{\partial}{\partial\beta} \ln Z(\lambda), \tag{6.3}$$

the Gibbs distribution maximises the entropy. The parameter $\beta$ in this case is a Lagrange multiplier whose value is determined by Eq. (6.3). This second interpretation means the Gibbs distribution also plays an important role outside physics (see Sec. 10.1). Specifically whenever some known properties of a stochastic system take the form of Eq. (6.3), the Gibbs distribution represents the optimal (according to maximum entropy arguments) distribution that reproduces these known properties.

Another important class of distributions are stationary distributions of a stochastic process, but to define these we first need to describe the stochastic process.

## 6.3 Markov processes

Describing the correlations between the random variables $\mathbf{s}(t)$ and $\mathbf{s}(t')$ at different times $t$ and $t'$ completes the definition of a stochastic process. We restrict ourselves to Markovian processes [242] with configurations connected by independent Poisson processes. To be more specific, when the system is in configuration $\mathbf{i}$ at time $t$, there is a Poisson transition to each other configuration $\mathbf{i}' \neq \mathbf{i}$ with rate $W_{\mathbf{i}\mathbf{i}'}(t)$. Given

$\mathbf{s}(t) = \mathbf{i}$ the probability of being in $\mathbf{s}(t+\delta t) = \mathbf{i}'$ a time $\delta t$ later is then

$$P^{\delta t}_{\mathbf{i}\mathbf{i}'}(t) = \delta_{\mathbf{i}\mathbf{i}'}\big[1 - \delta t \sum_{\mathbf{i}'\neq\mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t)\big] + (1-\delta_{\mathbf{i}\mathbf{i}'})\delta t W_{\mathbf{i}\mathbf{i}'}(t), \tag{6.4}$$

to first order in $\delta t$. It follows that the probability of the system evolving along a path $\mathbf{s}$ that takes values $\mathbf{s}(t)$ at times $0 \leq t \leq t_f$ is given by the distribution [243]

$$\mathcal{P}[\mathbf{s}] = \lim_{\delta t\to 0} \mathcal{P}^{\delta t}[\mathbf{s}], \tag{6.5a}$$

$$\mathcal{P}^{\delta t}[\mathbf{s}] = P_{\mathbf{s}_0}(0) P^{\delta t}_{\mathbf{s}_0\mathbf{s}_1}(0) P^{\delta t}_{\mathbf{s}_1\mathbf{s}_2}(t_1)\cdots P^{\delta t}_{\mathbf{s}_{n_{\text{steps}}-1}\mathbf{s}_{n_{\text{steps}}}}(t_{n_{\text{steps}}-1}), \tag{6.5b}$$

where we have used the shorthand $\mathbf{s}_j = \mathbf{s}(t_j)$, $t_j = j\delta t$ and set $n_{\text{steps}}\delta t = t_f$. We then find that

$$P_{\mathbf{i}}(t_f) = \lim_{\delta t\to 0} P^{\delta t}_{\mathbf{i}}(t_f), \tag{6.6a}$$

$$P^{\delta t}_{\mathbf{s}_{n_{\text{steps}}}}(t_f) = \sum_{\mathbf{s}_0,\mathbf{s}_1,\ldots,\mathbf{s}_{n_{\text{steps}}-1}} \mathcal{P}^{\delta t}[\mathbf{s}]. \tag{6.6b}$$

We are now in a position to derive the equations governing the time-evolution of the distribution $P_{\mathbf{i}}(t)$. Using Eq. (6.6b) it follows that to first order in $\delta t$ we have

$$P_{\mathbf{i}}(t) - P_{\mathbf{i}}(t-\delta t) = \delta t \sum_{\mathbf{i}'\neq\mathbf{i}} \big[P_{\mathbf{i}'}(t-\delta t) W_{\mathbf{i}'\mathbf{i}}(t-\delta t) - P_{\mathbf{i}}(t-\delta t) W_{\mathbf{i}\mathbf{i}'}(t-\delta t)\big]. \tag{6.7}$$

Dividing this by $\delta t$ and taking the limit $\delta t \to 0$ results in the master equation [239]

$$\frac{\partial P_{\mathbf{i}}(t)}{\partial t} = \sum_{\mathbf{i}'\neq\mathbf{i}} \big(P_{\mathbf{i}'}(t) W_{\mathbf{i}'\mathbf{i}}(t) - P_{\mathbf{i}}(t) W_{\mathbf{i}\mathbf{i}'}(t)\big). \tag{6.8}$$

The first term on the right hand side gives the rate of transitions into configuration $\mathbf{i}$ and the second gives the rate of leaving it.

To highlight the similarity between this master Eq. (6.8) and unitary quantum dynamics, we rewrite it as a stochastic Schrödinger equation [244, 245] of the form

$$\frac{\partial}{\partial t} \mid P\,(t)\rangle = \hat{H}(t) \mid P\,(t)\rangle\,. \tag{6.9}$$

The Hamiltonian $\hat{H}$ is defined by

$$\langle \mathbf{i} \mid \hat{H}(t) \mid \mathbf{i}'\rangle = W_{\mathbf{i}'\mathbf{i}}(t), \qquad \mathbf{i} \neq \mathbf{i}', \tag{6.10a}$$

$$\langle \mathbf{i} \mid \hat{H}(t) \mid \mathbf{i}\rangle = -\sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t). \tag{6.10b}$$

The probability vector describing a stochastic system therefore undergoes non-unitary evolution via a Schrödinger equation in imaginary time with a non-Hermitian Hamiltonian. From Eqs. (6.10) it follows that the Hamiltonian is what we call infinitesimal stochastic [246]. It has non-negative off-diagonal elements, a consequence of the non-negativity of transition rates, and non-positive diagonal elements that satisfy

$$\sum_{\mathbf{i}'} \langle \mathbf{i}' \mid \hat{H} \mid \mathbf{i}\rangle = 0, \tag{6.11}$$

a consequence of the conservation of probability.

## 6.4 Stationary states, ergodicity and equilibration

It has been shown in Ref. [247] that for every infinitesimal stochastic $\hat{H}$ there exists some stationary distribution $\mid P\rangle$, an eigenvector of $\hat{H}$ corresponding to a zero eigenvalue, such that

$$\frac{\partial}{\partial t} \mid P\rangle = \hat{H} \mid P\rangle = 0. \tag{6.12}$$

This follows from applying the Frobenius-Perron theorem. The same work [247] also showed that this vector is unique if the graph of $\hat{H}$ is strongly connected, i.e., all configurations are accessible from all others via some combination of allowed transitions. This implies ergodicity for time-independent Hamiltonians; after long times a system must arrive at the unique stationary state regardless of its initial state.

In later chapters we will take special interest in a (non-unique) Hamiltonian $\hat{H}(\lambda)$ for which the Gibbs distribution corresponding to energies $E_{\mathbf{i}}(\lambda)$ is stationary, i.e.,

$$\hat{H}(\lambda)\,|\,P(\lambda)\rangle = 0. \tag{6.13}$$

One way, but not the only way, to ensure this stationarity is to impose detailed balance. We write the condition of detailed balance in terms of the transition rates as

$$W_{\mathbf{ii'}}(\lambda) = \frac{w_{\mathbf{ii'}}}{1 + \mathrm{e}^{-\beta(E_{\mathbf{i}}(\lambda) - E_{\mathbf{i'}}(\lambda))}}, \tag{6.14}$$

for $\mathbf{i} \neq \mathbf{i'}$, where $w_{\mathbf{ii'}} = w_{\mathbf{i'i}}$ are some symmetric rates. Equation (6.14) implies Eq. (6.13).

If a sufficient number of the symmetric rates $w_{\mathbf{ii'}}$ are non-zero then $\hat{H}$ is strongly connected. The process is then ergodic with a single equilibrium/thermal stationary state and so the process describes equilibration/thermalisation.

# Chapter 7

# Useful results

Having introduced the basic mathematical description of stochastic processes, in this chapter we review two main results that we will make use of later. Specifically in Sec. 7.2 we discuss Jarzynski's equality, which relates the expected values of exponentials of the work done in a non-equilibrium process to equilibrium free energies. Before this, in Sec. 7.1 we describe transfer matrix methods for calculating the free energy of models in statistical physics. Together they will allow us to exactly calculate the expected value of the exponential of the work done in a non-equilibrium process, therefore allowing us to verify the accuracy of approximate numerical approaches for calculating the same quantity in Ch. 10.

## 7.1 Equilibrium: transfer matrices

The idea behind one-dimensional transfer matrix methods is to rewrite the partition function of a statistical mechanical system as a product of matrices. This product can then be efficiently evaluated. It is always possible if the energies of configurations $E_{\mathbf{i}}(\lambda)$ comprise only local terms.

Let us demonstrate this for the case that each energy is of the form

$$E_{\mathbf{i}}(\lambda) = \sum_{\ell=1}^{M-1} E^{[\ell]i_\ell i_{\ell+1}}(\lambda), \tag{7.1}$$

i.e., is a sum of nearest-neighbour terms. In this case the exponential of the energy of a configuration $\mathbf{i}$ may be re-written as a product of matrices (and boundary vectors)

$$\mathrm{e}^{-\beta E_{\mathbf{i}}(\lambda)} = A^{[1]i_1} A^{[2]i_2} \cdots A^{[M]i_M}, \tag{7.2}$$

where the elements of the tensors $A^{[\ell]}$ are given by

$$A^{[1]i_1}_{\mu_1} = \exp\left[-\beta E^{[1]i_1\mu_1}(\lambda)\right], \tag{7.3a}$$

$$A^{[\ell]i_\ell}_{\mu_{\ell-1}\mu_\ell} = \delta_{i_\ell,\mu_{\ell-1}} \exp\left[-\beta E^{[\ell]i_\ell\mu_\ell}(\lambda)\right], \qquad 1 < \ell < M, \tag{7.3b}$$

$$A^{[M]i_M}_{\mu_{M-1}} = \delta_{i_M,\mu_{M-1}}. \tag{7.3c}$$

The indices $\mu_\ell$ take values $0, \ldots, d-1$ and so the matrices $A^{[\ell]i_\ell}$ for $1 < \ell < M$ are $d \times d$ (the boundary vectors $A^{[1]i_1}$ and $A^{[M]i_M}$ are $1 \times d$ and $d \times 1$, respectively).

To obtain the partition function we take the sum

$$Z(\lambda) = \sum_{\mathbf{i}} \mathrm{e}^{-\beta E_{\mathbf{i}}(\lambda)} = T^{[1]} T^{[2]} \cdots T^{[M]}, \tag{7.4}$$

where $T^{[\ell]} = \sum_{i_\ell} A^{[\ell]i_\ell}$ are called transfer matrices. Often the partition function can then be evaluated analytically. This is always possible, for example, when taking the thermodynamic limit in the case that the energy contributions from the bulk are homogeneous, i.e., $T^{[\ell]} = T$ for $1 < \ell < M$. Alternatively, for a system of finite size the partition function may be efficiently computed numerically as it is just the product of $M$ matrices of dimension $d \times d$ (with appropriately sized boundary vectors). All of the thermodynamic properties of the system may then be computed

from the partition function.

Not only will we be interested in the thermodynamic properties that can be found via the transfer matrix approach, but also what the approach tells us about the description of a Gibbs distribution. To see this, it follows from Eq. (7.2) with the efficiently performable rescaling $A^{[\ell]} \mapsto A^{[\ell]} \left[Z(\lambda)\right]^{1/M}$ that the Gibbs distribution satisfies

$$P_{\mathbf{i}} = \frac{\mathrm{e}^{-\beta E_{\mathbf{i}}(\lambda)}}{Z(\lambda)} = A^{[1]i_1} A^{[2]i_2} \cdots A^{[M]i_M}. \tag{7.5}$$

It is therefore exactly represented by a matrix product state (MPS) of dimension $\chi = d$, where the matrices are non-negative. Transfer matrix methods therefore can be thought of as a special case of MPS methods applied to Gibbs distributions. This will be of interest in Chs. 9 and 10 where we study the representation of stochastic systems using MPS. It is also one way, in combination with Sec. 9.3.1 discussing the calculation of classical expected values for distributions represented by MPS, of demonstrating that the expected values of any local observables may be computed efficiently if the system is described by a Gibbs distribution with a transfer matrix solution.

Transfer matrix methods underly many exact solutions to statistical physics models. Examples of one-dimensional systems include the Zimm-Bragg [248] and Lifson-Roig [249] models of helictical coils in linear polymers. Onsager's solution to the two-dimensional Ising model with zero magnetic field uses the generalisation of the transfer matrix approach to two dimensions, where an exact analytical expression is obtained in the thermodynamic limit [250]. This is a result of the fact that for energies $E_{\mathbf{i}}(\lambda)$ that comprise only local terms an expression for the partition function (and Gibbs distribution) in terms of a contraction of tensors (the generalisation of matrix multiplication to higher dimensions) is possible in any dimension. However efficiently computing these contractions exactly, numerically or analytically is in

general not possible. Onsager's solution is a special case in which an exact analytical expression may be obtained. Other researchers have sought numerical methods to approximately perform the required contractions, using tensor network methods related to MPS. To give two examples, Nishino's calculated the partition function of a semi-infinite two-dimensional Ising model using density-matrix renormalisation group methods [251] while Orus and Vidal applied the infinite time-evolving block decimation algorithm to the infinite two-dimensional Ising model [252].

## 7.2 Non-equilibrium: Jarzynski's equality

Jarzynski's equality connects non-equilibrium properties of a system driven away from equilibrium to the system's equilibrium properties. Specifically in the case that some parameter is varied, it relates the expected value of the exponential of the work done on the system to the difference in equilibrium free energies of the system for the initial and final values of the parameter. If the system is initially in an equilibrium state of known free energy, then measuring the work done by varying the parameter in many realisations provides an estimate of the free energy of the equilibrium state corresponding to the final value of this parameter. This is despite never reaching this final equilibrium state. Therefore as well as being of interest as an analytical result, Jarzynski's equality underlies experimental and numerical approaches to free energy estimation.

The equality holds for Markovian stochastic processes described by an infinitesimal stochastic Hamiltonian $\hat{H}(\lambda)$ for which the Gibbs distribution $P_{\mathbf{i}}(\lambda)$ is stationary, i.e., $\hat{H}(\lambda)\,|\,P(\lambda)\rangle = 0$. At some time $\tau_i$ the system is assumed to be in equilibrium with the parameter set to $\lambda = \lambda_i$, i.e., $|\,P(\tau_i)\rangle = |\,P(\lambda_i)\rangle$. Then chang-

ing $\lambda$ to $\lambda = \lambda_f$ by a later time $\tau_f$ the following equality holds [243]

$$\mathrm{E}[\mathrm{e}^{-\beta W}, \mathcal{P}] = \mathrm{e}^{-\beta\left[F(\lambda_f) - F(\lambda_i)\right]}, \tag{7.6}$$

expressed in terms of the free energy

$$F(\lambda) = -\frac{1}{\beta} \ln Z(\lambda). \tag{7.7}$$

Here $W[\mathbf{s}]$ is the work done between $\tau_i$ and $\tau_f$ if the system evolves according to path $\mathbf{s}$ and $\mathrm{E}[\mathrm{e}^{-\beta W}, \mathcal{P}]$ represents the expected value over the possible paths[1], i.e.,

$$\mathrm{E}[\mathrm{e}^{-\beta W}, \mathcal{P}] = \int [\mathrm{D}\mathbf{s}] \mathrm{e}^{-\beta W[\mathbf{s}]} \mathcal{P}[\mathbf{s}]. \tag{7.8}$$

Before Jarzynski's proof [243], the result was known in two limits. Firstly if the parameter $\lambda$ is changed infinitely slowly (reversibly) then the system will always remain in equilibrium and the work done $W[\mathbf{s}] = F(\lambda_f) - F(\lambda_i)$ will be equal to the free energy difference for all paths. Secondly if $\lambda$ is changed suddenly at time $\tau_*$ then the work done is equal to $W[\mathbf{s}] = E_{\mathbf{s}(\tau_*)}(\lambda_f) - E_{\mathbf{s}(\tau_*)}(\lambda_i)$ the increase in energy for the configuration $\mathbf{s}(\tau_*)$ of the system at that time. The expected value according to the Gibbs distribution $P_{\mathbf{s}(\tau_*)}(\lambda_i)$ then obeys Eq. (7.6). The importance of Jarzynski's result is that it holds however $\lambda$ is varied. For example, those wishing to perform free energy estimation may choose a balance between considering many short evolutions and fewer longer evolutions, where the variance of $\mathrm{e}^{-\beta W[\mathbf{s}]}$ may be smaller for the latter. The relationship between the variance of an observable and the number of paths that must be sampled will be important when discussing Monte Carlo simulations, which we do next.

[1]Here we of course must have that $\tau_i \geq 0$ and $\tau_f \leq t_f$ so that we are interested in the work done within the period over which our stochastic process is defined.

# Chapter 8

# Dynamical Monte Carlo

In this chapter we present the principal method used to simulate non-equilibrium stochastic systems, called dynamical Monte Carlo (DMC). The application of Monte Carlo to Markovian dynamics has a long history going back to the 1960s [36, 37, 38, 39]. Here we have only the modest aim of reviewing DMC in sufficient detail to understand its workings in principle, and to identify its strengths and weaknesses.

We begin in Sec. 8.1 by outlining a basic DMC simulation of a Markovian process, showing how to sample paths $\mathbf{s}$ given the rates $W_{\mathbf{i},\mathbf{i}'}$ and how to obtain unbiased estimates of expected values. Then in Sec. 8.2 we describe two types of improvements that are made when the observable of interest has a high variance. These are sequential importance sampling and the so-called go with the winners strategy. We conclude the chapter with a discussion of the main implications in Sec. 8.3.

## 8.1 Basic Monte Carlo

### 8.1.1 Sampling

The aim of a Monte Carlo simulation is to estimate the expected value

$$\mathrm{E}\left[O, \mathcal{P}\right] = \int [\mathrm{D}\mathbf{s}] O[\mathbf{s}] \mathcal{P}[\mathbf{s}]. \tag{8.1}$$

Here $\mathcal{P}[\mathbf{s}]$, introduced in Eq. (6.5a), is the element of the path-dependent probability distribution $\mathcal{P}$ associated with the Markov process tracing the path $\mathbf{s}$ between times $t = 0$ and $t = t_f$. Similarly $O[\mathbf{s}]$ is the value an observable $O$ takes for that same path. A basic Monte Carlo algorithm estimates Eq. (8.1) by first randomly selecting $N$ paths according to the probability distribution[1] $\mathcal{P}$. Writing the sampled paths as $\mathbf{s}^1, \cdots, \mathbf{s}^N$, the expected value $\mathrm{E}[O, \mathcal{P}]$ is then estimated as the mean[2]

$$\bar{O}_N = \frac{1}{N} \sum_{n=1}^{N} O[\mathbf{s}^n]. \tag{8.2}$$

Trivially then, the average value of this estimate is

$$\left\langle \bar{O}_N \right\rangle = \frac{1}{N} \sum_{n=1}^{N} \mathrm{E}[O, \mathcal{P}] = \mathrm{E}[O, \mathcal{P}], \tag{8.3}$$

so the Monte Carlo estimate provides an unbiased estimator for the true expected value. The root of the average mean square error $\langle \overline{\Delta O_N} \rangle$ of the Monte Carlo estimate after some manipulation becomes

$$\langle \overline{\Delta O_N} \rangle = \sqrt{\left\langle \frac{1}{N} \sum_{n=1}^{N} (O[\mathbf{s}^n] - \bar{O}_N)^2 \right\rangle} = \sqrt{\frac{\mathrm{Var}[O, \mathcal{P}]}{N}}, \tag{8.4}$$

where the variance is given by

$$\mathrm{Var}[O, \mathcal{P}] = \int [\mathrm{D}\mathbf{s}] \left( O[\mathbf{s}] - \mathrm{E}[O, P] \right)^2 \mathcal{P}[\mathbf{s}]. \tag{8.5}$$

[1] In Part I, $N$ and $n$ were used to denote particle number and density, respectively. In this part $N$ is used to denote the number of Monte Carlo instances and $n$ is used to index the instances.

[2] For clarity, we use mean and a bar $\bar{\cdot}$ when we refer to taking the mean over the sampled paths, average and angled brackets $\langle \cdot \rangle$ when averaging over the possible results the Monte Carlo algorithm could have given, i.e., averaging over all possible random numbers generated in the algorithm, and expected value and $\mathrm{E}[\cdot, \mathcal{P}]$ when taking the expected value of an observable with respect to the distribution of paths $\mathcal{P}$.

Therefore, in the limit $N \to \infty$ the error $\langle\overline{\Delta O_N}\rangle$ goes to zero. More realistically, the sample size $N$ required to ensure a fractional average error $\langle\overline{\Delta O_N}\rangle/\mathrm{E}\left[O, \mathcal{P}\right] = \mathcal{E}$ is

$$N = \frac{\mathrm{Var}\left[O, \mathcal{P}\right]}{\mathcal{E}\mathrm{E}\left[O, \mathcal{P}\right]^2}, \tag{8.6}$$

which increases linearly with the variance $\mathrm{Var}\left[O, \mathcal{P}\right]$. Thus convergence is expected to be slow for observables with a high variance (by which it is meant that the variance over expected value squared is large). This will be important for motivating more advanced Monte Carlo methods in Sec. 8.2 and the use of the time-evolving block decimation method in Ch. 10.

On several occasions we will be interested in the special case of an observable $O$ that depends only on the configuration $\mathbf{s}(t)$ of the system at some time $t$. Writing $O[\mathbf{s}] = O_{\mathbf{s}(t)}$, the expressions for the average estimate and error simplify to

$$\langle\bar{O}_N\rangle = \mathrm{E}\left[O, P\right] = \sum_{\mathbf{i}} O_{\mathbf{i}} P_{\mathbf{i}}(t), \tag{8.7a}$$

$$N\langle\overline{\Delta O_N}\rangle = \mathrm{Var}\left[O, P\right] = \sum_{\mathbf{i}} \left(O_{\mathbf{i}} - \mathrm{E}\left[O, P\right]\right)^2 P_{\mathbf{i}}(t). \tag{8.7b}$$

We next focus on how to sample paths according to the correct probability distribution $\mathcal{P}$. To do this we assume our process is Markovian and use the formalism we introduced in Sec. 6.3. We show how to sample according to $\mathcal{P}$ without knowing its form, only the Poisson rates $W_{\mathbf{ii}'}$. We use the DMC terminology of Ref. [40].

### 8.1.2 Fixed time-step

In Sec. 6.3 we defined our Markovian process by stating that given $\mathbf{s}(t) = \mathbf{i}$ the probability of being in $\mathbf{s}(t+\delta t) = \mathbf{i}'$ a time $\delta t$ later is (cf. Eq. (6.4))

$$P^{\delta t}_{\mathbf{i}\mathbf{i}'}(t) = \delta_{\mathbf{i}\mathbf{i}'}\big[1 - \delta t \sum_{\mathbf{i}'' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}''}(t)\big] + (1 - \delta_{\mathbf{i}\mathbf{i}'})\delta t W_{\mathbf{i}\mathbf{i}'}(t), \tag{8.8}$$

to first order in $\delta t$. The fixed time-step method uses this definition as a starting point for a numerical procedure. Specifically, given $\mathbf{s}(t) = \mathbf{i}$ the configuration $\mathbf{s}(t+\delta t)$ a time $\delta t$ later is sampled according to the probabilities $P^{\delta t}_{\mathbf{i}\mathbf{i}'}(t)$. It therefore relies on using a small enough $\delta t$ that this first order expression for $P^{\delta t}_{\mathbf{i}\mathbf{i}'}(t)$ is a good approximation. This means that each $W_{\mathbf{i}\mathbf{i}'}$ must vary over time-scales much larger than $\delta t$ and $\delta t W_{\mathbf{i}\mathbf{i}'}(t), \delta t \sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}'\mathbf{i}}(t) \ll 1$.

This sampling can be performed by first finding the allowed final configurations $\mathbf{i}' \neq \mathbf{i}$ satisfying $W_{\mathbf{i}\mathbf{i}''} \neq 0$. Then a weighted random choice is made between transition and no transition with weights $\delta t \sum_{\mathbf{i}'} W_{\mathbf{i}\mathbf{i}'}(t)$ and $1 - \delta t \sum_{\mathbf{i}'} W_{\mathbf{i}\mathbf{i}'}(t)$, respectively. If no transition is to occur then $\mathbf{s}(t+\delta t) = \mathbf{i}$. If a transition is to occur then a random choice is made between each allowed final configuration $\mathbf{s}(t+\delta t) = \mathbf{i}' \neq \mathbf{i}$ with relative weight $W_{\mathbf{i}\mathbf{i}'}(t)$.

Alternatively it may be more efficient to take each allowed final configuration $\mathbf{i}'$ in turn and perform a weighted random choice between undergoing a transition to $\mathbf{i}'$ or not with weights $\delta t W_{\mathbf{i}\mathbf{i}'}(t)$ and $1 - \delta t W_{\mathbf{i}\mathbf{i}'}(t)$, respectively. It is crucial in this case that $\delta t$ is small enough that the probability of two different transitions being chosen is negligible.

Repeating this sampling procedure for each time-step in the desired time-interval selects a path $\mathbf{s}$ according to the distribution $\mathcal{P}[\mathbf{s}]$.

### 8.1.3 Variable time-step

The variable time-step method is based on the expression, given in Eq. (6.5), for $\mathcal{P}[\mathbf{s}]$ in the continuous limit $\delta t \to 0$. From this expression, and using the identity $\lim_{m\to\infty} \prod_{j=1}^{m} \{1 + f(t + j[t' - t]/m)/m\}^m = \exp(-\int_t^{t'} \mathrm{d}t'' f(t''))$ [253], it follows that given $\mathbf{s}(t) = \mathbf{i}$ the probability of not changing configuration at all by a later time $t' > t$ is

$$R_{\mathbf{i}}(t, t') = \exp\left( - \int_t^{t'} \mathrm{d}t'' \sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t'') \right). \tag{8.9}$$

Therefore the time of the first transition $t'$ is governed by the probability distribution that is the negative of the derivative of this,

$$-\frac{\mathrm{d}R_{\mathbf{i}}(t, t')}{\mathrm{d}t'} = \sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t') \exp\left( - \int_t^{t'} \mathrm{d}t'' \sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t'') \right). \tag{8.10}$$

Moreover if the transition occurs at time $t'$ then the probability of $\mathbf{s}(t') = \mathbf{i}'$ is

$$T_{\mathbf{i}\mathbf{i}'}(t') = \frac{W_{\mathbf{i}\mathbf{i}'}(t')}{\sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t')}. \tag{8.11}$$

The variable time-step approach is to first generate some time $t'$ according to the distribution in Eq. (8.10) and then randomly choose a final configuration $\mathbf{s}(t') = \mathbf{i}'$ with weight $T_{\mathbf{i}\mathbf{i}'}(t')$. To generate the time $t'$, a random number $x_{\mathrm{rand}}$ is first generated uniformly between 0 and 1. Then the equation

$$R_{\mathbf{i}}(t, t') = x_{\mathrm{rand}}, \tag{8.12}$$

is solved for $t'$. In the simple case that each $W_{\mathbf{i}\mathbf{i}'}$ has no time-dependence

$$t' = t - \frac{\ln x_{\mathrm{rand}}}{\sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}}. \tag{8.13}$$

It is clear for the case of time-independent rates that the variable time-step method is likely to generate paths **s** appreciably faster than the fixed time-step method. In the latter many steps have to be performed for each transition whereas in the former each step corresponds to a transition. For time-independent rates the speed of the variable time-step method depends on how easy it is to solve equations of the form of Eq. (8.12).

There are many variations upon the method we have described above, designed to optimise the processing time and memory required to keep track of the allowed transitions and their rates. It is usually possible to perform each time-step using resources that scale as $\mathcal{O}(1)$ or $\mathcal{O}(\log M)$, where $M$ is the size of the system and $d^M$ the number of configurations. This assumes that each configuration is only connected to a few others by allowed transitions and the number of different values that transition rates take is small. Further this does not take into account the scaling of the resources needed to solve Eq. (8.12) in the case of time-dependent rates.

## 8.2 High variance

As we have seen above, in most cases a single path of a Markov process is sampled efficiently using DMC. Difficulties then only arise when the required sample size given in Eq. (8.6) grows large. This occurs when the variance of the observable of interest is much larger than the square of its expected value. Developing more efficient methods to estimate expected values of high variance observables is a priority in several areas, e.g., where accurate estimates of the probabilities of rare events are sought. One such method is sequential importance sampling [41, 42, 43, 44], whereby paths are sampled according to the distribution arising from a different set of parameters chosen to preferentially sample paths corresponding to large values of the observable. Observable values are then reweighted such that the average value of

their mean is unchanged while the variance is (hopefully) reduced. Other methods, including go with the winners [45, 46] or multi-level splitting [47], involve branching. Paths may be duplicated or discarded as they are sampled, depending on whether they are more or less likely to correspond to large observable values, respectively. Again, possibly in conjunction with importance sampling, carefully assigned weights ensure the average value of the mean is unchanged. In this section we present these two main approaches for estimating high variance observables.

### 8.2.1 Sequential importance sampling

Importance sampling is based on the realisation that sampling paths according to some distribution $\mathcal{Q}$ and assigning values $O'[\mathbf{s}] = O[\mathbf{s}]\mathcal{P}[\mathbf{s}]/\mathcal{Q}[\mathbf{s}]$ to an observable $O'$ leads to an expected value

$$\mathrm{E}[O', \mathcal{Q}] = \int [\mathrm{D}\mathbf{s}] O'[\mathbf{s}]\mathcal{Q}[\mathbf{s}] = \int [\mathrm{D}\mathbf{s}] O[\mathbf{s}]\mathcal{P}[\mathbf{s}] = \mathrm{E}\left[O, \mathcal{P}\right], \tag{8.14}$$

that is exactly the same as that of observable $O$ for distribution $\mathcal{P}$. Crucially, the variance $\mathrm{Var}[O', \mathcal{Q}]$ might be less than that of the original $\mathrm{Var}\left[O, \mathcal{P}\right]$. In fact it is possible that $\mathrm{Var}[O', \mathcal{Q}] = 0$ iff $\mathcal{Q}$ is chosen such that $O[\mathbf{s}]\mathcal{P}[\mathbf{s}]/\mathcal{Q}[\mathbf{s}] = \mathrm{E}\left[O, \mathcal{P}\right]$. This limit is of course unfeasible in practice since it relies on knowing the expected value $\mathrm{E}\left[O, \mathcal{P}\right]$, which after all is what we are trying to estimate. The aim of importance sampling is then to efficiently sample according to a probability distribution $\mathcal{Q}$ in which it is hoped that the observable $O'$ has a small variance. At the same time, it must be possible to efficiently calculate the ratio $\mathcal{P}[\mathbf{s}]/\mathcal{Q}[\mathbf{s}]$ and therefore generate an observable value $O'[\mathbf{s}]$ from that of the original $O[\mathbf{s}]$ for each path sampled. If this is achievable, the expected value $\mathrm{E}\left[O, \mathcal{P}\right]$ is then estimated exactly as in Sec. 8.1.1 but with $O'$ and $\mathcal{Q}$ instead of $O$ and $\mathcal{P}$.

The most popular approach for performing importance sampling of Markovian

dynamics is called sequential importance sampling (see Refs. [41, 42, 43, 44] for works introducing and using sequential importance sampling in various settings). It involves sampling according to a Markov process with altered Poisson rates $W'_{\mathbf{i}\mathbf{i}'}$. This results in a probability distribution $\mathcal{Q}$ for paths between $t = 0$ and $t_f = n_{\text{steps}}\delta t$ given by

$$\mathcal{Q}[\mathbf{s}] = \lim_{\delta t \to 0} \mathcal{Q}^{\delta t}[\mathbf{s}], \tag{8.15a}$$

$$\mathcal{Q}^{\delta t}[\mathbf{s}] = Q_{\mathbf{s}_0}(0) Q^{\delta t}_{\mathbf{s}_0\mathbf{s}_1}(0) Q^{\delta t}_{\mathbf{s}_1\mathbf{s}_2}(t_1) \cdots Q^{\delta t}_{\mathbf{s}_{n_{\text{steps}}-1}\mathbf{s}_{n_{\text{steps}}}}(t_{n_{\text{steps}}-1}), \tag{8.15b}$$

$$Q^{\delta t}_{\mathbf{i}\mathbf{i}'}(t) = \delta_{\mathbf{i}\mathbf{i}'}\big[1 - \delta t \sum_{\mathbf{i}'' \neq \mathbf{i}} W'_{\mathbf{i}\mathbf{i}''}(t)\big] + (1 - \delta_{\mathbf{i}\mathbf{i}'})\delta t W'_{\mathbf{i}\mathbf{i}'}(t). \tag{8.15c}$$

The equivalent equations for $\mathcal{P}$ are Eqs. (6.4) and (6.5), and accompany the definitions of the shorthand notations $\mathbf{s}_j$ and $t_j$. Though the two distributions $\mathcal{P}$ and $\mathcal{Q}$ are complicated quantities it is nevertheless possible to evaluate the ratio $\mathcal{P}[\mathbf{s}]/\mathcal{Q}[\mathbf{s}]$ for a given path $\mathbf{s}$. For instance, consider the path

$$\mathbf{s}(t) = \begin{cases} \mathbf{i}_0 & 0 \leq t < \tilde{t}_1 \\ \mathbf{i}_1 & \tilde{t}_1 \leq t < \tilde{t}_2 \\ \vdots & \vdots \\ \mathbf{i}_{n_{\text{tran}}} & \tilde{t}_{n_{\text{tran}}} \leq t \leq \tilde{t}_f \end{cases}, \tag{8.16}$$

in which there are $n_{\text{tran}}$ transitions. It begins in configuration $\mathbf{i}_0$ at time 0 then transitions to configuration $\mathbf{i}_1$ at time $\tilde{t}_1$ and so on until arriving at the final configuration $\mathbf{i}_{n_{\text{tran}}}$ by time $\tilde{t}_{n_{\text{tran}}} < t_f$. By taking the ratio of the quantities $\mathcal{P}^{\delta t}[\mathbf{s}]$ and $\mathcal{Q}^{\delta t}[\mathbf{s}]$ appearing in Eqs. (6.5b) and (8.15b), respectively, then taking the limit $\delta t \to 0$ we find that

$$\frac{\mathcal{P}[\mathbf{s}]}{\mathcal{Q}[\mathbf{s}]} = \frac{P_{\mathbf{i}_0}(0)}{Q_{\mathbf{i}_0}(0)} \frac{R_{\mathbf{i}_0}(0,\tilde{t}_1)}{R'_{\mathbf{i}_0}(0,\tilde{t}_1)} \frac{T_{\mathbf{i}_0\mathbf{i}_1}(\tilde{t}_1)}{T'_{\mathbf{i}_0\mathbf{i}_1}(\tilde{t}_1)} \frac{R_{\mathbf{i}_1}(\tilde{t}_1,\tilde{t}_2)}{R'_{\mathbf{i}_1}(\tilde{t}_1,\tilde{t}_2)} \cdots \frac{T_{\mathbf{i}_{n_{\text{tran}}-1}\mathbf{i}_{n_{\text{tran}}}}(\tilde{t}_{n_{\text{tran}}})}{T'_{\mathbf{i}_{n_{\text{tran}}-1}\mathbf{i}_{n_{\text{tran}}}}(\tilde{t}_{n_{\text{tran}}})} \frac{R_{\mathbf{i}_{n_{\text{tran}}}}(\tilde{t}_{n_{\text{tran}}},t_f)}{R'_{\mathbf{i}_{n_{\text{tran}}}}(\tilde{t}_{n_{\text{tran}}},t_f)}, \tag{8.17}$$

where $R_{\mathbf{i}}(t,t')$ and $T_{\mathbf{ii}'}(t)$ are defined in Eqs. (8.9) and (8.11), respectively. Similarly $R'_{\mathbf{i}}(t,t')$ and $T'_{\mathbf{ii}'}(t)$ are defined *mutatis mutandis* with each $W_{\mathbf{ii}'}(t)$ replaced by $W'_{\mathbf{ii}'}(t)$. From Eqs. (8.17) and (8.11) we see that there is a restriction on the rates $W'_{\mathbf{ii}'}$ that may be chosen. Specifically if $W_{\mathbf{ii}'}(t) = 0$ then it must be that $W'_{\mathbf{ii}'}(t) = 0$, otherwise the transition $\mathbf{i} \to \mathbf{i}'$ might occur in the altered dynamics and cause the ratio in Eq. (8.17) to be undefined. Other than this any choice $W'_{\mathbf{ii}'}$ is allowed.

The sequential importance sampling strategy is then as follows. An initial configuration $\mathbf{i}_0$ is randomly selected according to the distribution $Q_{\mathbf{i}_0}(0)$ and a weight $w[\mathbf{s}] = P_{\mathbf{i}_0}(0)/Q_{\mathbf{i}_0}(0)$ is assigned. Then this configuration is evolved using a basic Monte Carlo simulation of the type presented in Sec. 8.1 with rates $W'_{\mathbf{ii}'}$. For each period in which the path $\mathbf{s}$ stays in a configuration $\mathbf{i}_j$ between times $\tilde{t}_j$ and $\tilde{t}_{j+1}$ we update the weight according to $w[\mathbf{s}] \mapsto w[\mathbf{s}]R_{\mathbf{i}_j}(\tilde{t}_j, \tilde{t}_{j+1})/R'_{\mathbf{i}_j}(\tilde{t}_j, \tilde{t}_{j+1})$. For each transition between configurations $\mathbf{i}_j$ and $\mathbf{i}_{j+1}$ at time $\tilde{t}_{j+1}$ we update the weight according to $w[\mathbf{s}] \mapsto w[\mathbf{s}]T_{\mathbf{i}_j\mathbf{i}_{j+1}}(\tilde{t}_{j+1})/T'_{\mathbf{i}_j\mathbf{i}_{j+1}}(\tilde{t}_{j+1})$. Following the sampling of $N$ paths $\mathbf{s}^1, \cdots, \mathbf{s}^N$ we estimate the expected value $\mathrm{E}[O, \mathcal{P}]$ by

$$\bar{O}'_N = \frac{1}{N}\sum_{n=1}^{N} O'[\mathbf{s}^n] = \frac{1}{N}\sum_{n=1}^{N} O[\mathbf{s}^n]w[\mathbf{s}^n], \tag{8.18}$$

where $\langle \bar{O}'_N \rangle = \mathrm{E}[O, \mathcal{P}]$.

Importance sampling all comes down to the art of choosing a valid alternative set of rates $W'_{\mathbf{ii}'}$ that causes the system to preferentially sample the paths that most contribute to the desired expected value, i.e., paths for which $O[\mathbf{s}]\mathcal{P}[\mathbf{s}]$ is large (from above, for the optimal distribution $\mathcal{Q}[\mathbf{s}] \propto O[\mathbf{s}]\mathcal{P}[\mathbf{s}]$). This can be extremely effective. In principle, the set of rates $W'_{\mathbf{ii}'}$ could depend on the particulars of each path as it is sampled. Sequential importance sampling is therefore very flexible. The problem is that there is no general strategy for choosing the most optimal valid set of rates $W'_{\mathbf{ii}'}$ and the choice largely rests on physical intuition and a prior knowledge

of the specific process. Additionally, different high variance observables may require different choices of rates.

### 8.2.2 Go with the winners

The go with the winners method can be used in conjunction with sequential importance sampling and therefore can be thought of as an additional tool to increase the efficiency of DMC simulations for high variance observables. Its aim is to reduce the fraction of computational effort put into sampling paths with low weights. There are several strategies for doing this and below we list a few representative examples from the simplest to the more complicated. We use the terminology of Ref. [46].

Before doing this, we first note that the go with the winners strategies involve continuously tracking each DMC path as it is sampled and making decisions based on the path in 'real time', i.e., during its sampling. Second, the strategies usually involve randomly adjusting the path weights such that the weights assigned to two identical sampled paths might be different. Third, in order to make a decision about how much we expect a path to contribute to the expected value before we have finished sampling the path — the whole premise of the go with the winners approach — we must deal with observables that accumulate locally in time. If this were not true then it would be difficult to estimate how much a path will contribute to the expected value before reaching the end of the simulation. As a result we make two notational changes.

We write $O[\mathbf{s}^n](t)$ as the value of an observable at time $t$, the time up to which the path $\mathbf{s}^n$ has been sampled in the $n$-th Monte Carlo instance. As an example, consider the multiplicatively increasing observable

$$O[\mathbf{s}^n](t) = \exp\left(\int_0^t \mathrm{d}t' \dot{o}_{\mathbf{s}^n(t')}\right), \tag{8.19}$$

where $\dot{o}_{\mathbf{i}}$ is the rate at which the logarithm of the observable increases when the system is in configuration $\mathbf{i}$.

Our second notational change is to write $w^n[\mathbf{s}^n](t)$ as the weight assigned to the $n$-th Monte Carlo instance at time $t$. The additional label $n$ here indicates that there is a possibility that $w^n[\mathbf{s}^n](t) \neq w^{n'}[\mathbf{s}^{n'}](t)$ even if $\mathbf{s}^n = \mathbf{s}^{n'}$.

We now list go with the winner strategies:

1. As a trivial example, if at any time $t$ the contribution $O[\mathbf{s}^n](t)w^n[\mathbf{s}^n](t) = 0$ then no further computation time is spent on this path (though it still counts towards the total sample size $N$). This avoids any unnecessary use of computational resources. It is said that the path is killed.

2. If the contribution $O[\mathbf{s}^n](t)w^n[\mathbf{s}^n](t)$ of a path drops below a critical value $\Upsilon_-(t)$ then a random number $x_{\text{rand}}$ is uniformly selected between 0 and 1. If $x_{\text{rand}} > 1/2$ then the path is killed, otherwise its weight is increased $w^n[\mathbf{s}^n](t) \mapsto 2w^n[\mathbf{s}^n](t)$. This reduces the computational effort put into sampling paths that contribute little to the expected value. Only a few of them are kept while the weights of those remaining are increased to compensate for this.

3. If the contribution $O[\mathbf{s}^n](t)w^n[\mathbf{s}^n](t)$ of a path increases above a critical value $\Upsilon_+(t)$ then the path is cloned and the original and its clone are each assigned a weight $w^n[\mathbf{s}^n](t)/2$. This increases the computational resources spent on paths that contribute a lot to the expected value.

The go with the winners method is a combination of these three strategies (or similar alternatives). Note that all still result in an unbiased estimate of the desired expected value when averaged over the values taken by the random numbers used in the simulation.

The main difficulty lies in choosing the two critical values $\Upsilon_-$ and $\Upsilon_+$. Choosing $\Upsilon_-$ ($\Upsilon_+$) to be too small (large) will result in a lot of computational effort being

wasted, while choosing it to be too large (small) will result in the estimate of the expected value being determined by too few paths. Again the usefulness of the method comes down to making judicious choices. Usually the sampling of paths, i.e., each Monte Carlo instance, is performed in parallel, and $\Upsilon_-$ and $\Upsilon_+$ are chosen on-the-fly based on the weights of the paths that have not been killed.

Before moving on to the final discussion of this chapter, we note that the go with the winners method presented here has appeared under many guises in relation to a wide range of computational problems. Reference [254] points out the range of applications and, to some extent, the history of the go with the winners/PERM/population control strategies, something we consider beyond the purview of this thesis.

## 8.3 Discussion

In this chapter we have introduced DMC, which is by far the most widely used method to simulate Markovian dynamical processes. We have reviewed the different techniques available for tackling problems such as high variances. This chapter will therefore provide a benchmark for the new tools we introduce in the next two chapters.

To summarise, we found that a single instance of a Monte Carlo simulation is extremely efficient for time-independent rates that take values from a small set and connect each configuration to only a few others. However this might not be so simple for time-dependent rates, where either a potentially complicated integral needs to be performed at each time-step (variable time-step method) or a large number of time-steps need to be performed (fixed time-step).

As for the number of instances needed to obtain an accurate estimate of the observable, this increases linearly with the variance of that observable and so high

variance observables represent a problem for Monte Carlo simulation. Several strategies have been put forward to overcome this but each relies on judicious choices by the user and a prior knowledge or intuition about the system. If used haphazardly, these strategies could make things worse as well as better.

# Chapter 9

# TEBD for stochastic processes

In the previous chapter we examined the use of dynamical Monte Carlo (DMC) for simulating non-equilibrium stochastic processes. The goal of this chapter is to evaluate the effectiveness of using time-evolving block decimation (TEBD) for the same task[1]. We will see that TEBD provides a means of calculating expected values in a fundamentally different way to DMC. The latter averages over observable values calculated from many paths, each a possible evolution of the process, while in the former we evolve an approximation to the entire probability distribution of the process from which observable values are calculated. The motivation is to provide a complementary method to DMC with a contrasting source of error; rather than

[1]The results in this chapter were first presented by us in Ref. [1].

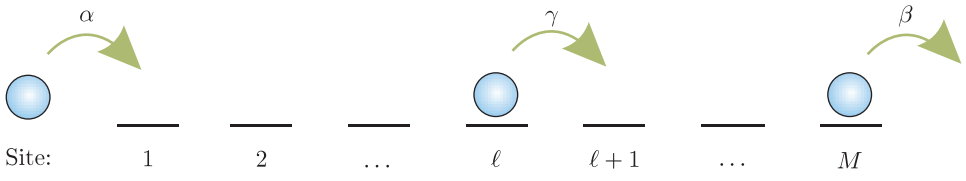


**Figure 9.1**. The TASEP is a paradigm of stochastic non-equilibrium processes. Three Poisson processes contribute to the dynamics of the process. Particles are injected into the first site with rate $\alpha$ if it is empty and ejected from the $M$-th site with rate $\beta$ if occupied. Finally, particles in the $\ell$-th site hop to the $(\ell+1)$-th site with a rate $\gamma$ if that site is empty, resulting in flow from left to right.

insufficient sampling we deal with errors arising from approximating a probability distribution by a matrix product state (MPS).

Density-matrix renormalisation group (DMRG) methods [27, 28, 29], which are related to TEBD and MPS, have previously been applied to classical non-equilibrium processes. The application of DMRG in Refs. [30, 31, 32] suggests that the stationary probability distributions of many stochastic processes may be well-approximated by MPS of small dimension, although their particular method faced numerical instabilities for large systems arising from non-Hermitian diagonalisation [31]. Using corner-transfer-matrix DMRG (CTMRG) [33, 34] and later stochastic light-cone CTMRG [35] these results have been extended to transient states in the thermodynamic limit, but the methods break down after a modest number of time-steps.

A key premise that motivates the adaption of quantum algorithms is that, as shown in Sec. 6.3, a large class of Markovian stochastic processes evolve according to a Schrödinger-like equation. We demonstrate how this similarity between quantum and stochastic evolution can be exploited by applying TEBD to the totally-asymmetric simple exclusion process (TASEP) [255, 256]. The TASEP, shown schematically in Fig. 9.1, is the quintessential driven particle hopping process. Such processes share the following features: particles occupy a one-dimensional lattice; the boundary sites are connected to particle reservoirs, so particles are injected and ejected there; and a hopping process causes particles to move through the lattice (see Refs. [257, 258] for reviews). Simple particle hopping processes have drawn much attention, like the Ising and Heisenberg models of equilibrium statistical physics [259], because they are archetypal of non-equilibrium physics.

The TASEP has an infinitesimal stochastic Hamiltonian that consists of single-site and two-site nearest-neighbour terms only [238, 260], which allows us to use the TEBD algorithm presented in Sec. 3.2.2 with almost no alteration. The only stage of the algorithm that we consider changing significantly is the matrix factorisation used

to return the description of the process to an MPS after each two-site operation. As explained in Sec. 3.2.2, the singular value decomposition (SVD) is perfectly suited to quantum systems. We will find that this is not necessarily the case for classical stochastic processes. Therefore we adapt TEBD to allow for a range of factorisations to be used as part of the algorithm, since choosing which factorisation is the key element in determining its computational cost and accuracy.

As well as the SVD the main candidate we concentrate on is non-negative matrix factorisation (NMF) [261, 262], which by restricting the matrices of the MPS to be non-negative (called an sMPS) allows for an information theoretic interpretation, as recently investigated in Ref. [263]. We compare the performance of these two factorisations and find that while SVD algorithms are well developed, stable and efficient, current NMF algorithms are unable to identify accurate factorisations due to the non-convexity of the task. As a result we focus on the SVD-based TEBD algorithm and investigate its performance in simulating the TASEP. The results obtained promise that TEBD could be applicable to and accurate in simulating a wide range of non-equilibrium processes.

The structure of the chapter is as follows. We begin by discussing the representation of probability distributions by MPS in Sec. 9.1. Next in Sec. 9.2 we describe the application of TEBD to Markovian dynamics and compare the properties of the two factorisation methods, NMF and SVD. In Sec. 9.3 we briefly explain how expected values are calculated from an MPS representing a probability distribution and discuss the scaling of the errors. The TASEP is introduced in Sec. 9.4. In Sec. 9.5 we compare the performance of TEBD with NMF and SVD before focusing on the errors of the SVD-based algorithm for small systems. We then go on to apply TEBD to study larger systems in Sec. 9.6 before summarising our results in Sec. 9.7.

## 9.1 MPS for stochastic processes

The general MPS formalism introduced in Sec. 3.2.1 can equally be applied to stochastic classical processes. To summarise earlier chapters, the state of a stochastic process at a given time (cf. Sec. 6.1) is represented by the probability distribution

$$|\,P\rangle = \sum_{\mathbf{i}} P_{\mathbf{i}}\,|\,\mathbf{i}\rangle\,, \tag{9.1}$$

where $\mathbf{i} = (i_1, \ldots, i_M)$ is a configuration of the whole system and $i_\ell = 0, \cdots, d-1$ labels the configuration of the $\ell$-th site. Classical states $|\,P\rangle$ are contained in the positive orthant of the vector space $(\mathbb{R}^d)^{\otimes M}$ and are normalised according to the $L_1$-norm as $\|\,|\,P\rangle\,\|_1 = 1$.

An MPS representation (cf. Sec. 3.2.1) of this distribution is

$$P_{\mathbf{i}} = A^{[1]i_1} A^{[2]i_2} \cdots A^{[M]i_M}, \tag{9.2}$$

where the dimension $\chi$ of the MPS gives the size of the matrices $A^{[\ell]i_\ell}$. An immediate restriction to real matrices $A^{[\ell]i_\ell}$ within the MPS expansion in Eq. (3.18) can be made in this case. A major question which we will address in this chapter is whether it is necessary or indeed desirable to further restrict the matrices to be non-negative. To do so manifestly ensures that the sMPS approximation to $|\,P\rangle$ never leaves the positive orthant. Early exact analytical calculations based on an MPS formalism used sMPS [238].

Similarly to quantum systems (see Sec. 3.2.1), for classical systems a link between the dimension $\chi$ required and correlations in $|\,P\rangle$ can be established by working exclusively with non-negative quantities. As before we bipartition the system after site $\ell$ into two contiguous blocks $\mathtt{L}^{[\ell]}$ and $\mathtt{R}^{[\ell]}$ and then expand $|\,P\rangle = \sum_{\mathbf{ij}} P_{\mathbf{ij}}\,|\,\mathbf{i}\rangle_{\mathtt{L}}\,|\,\mathbf{j}\rangle_{\mathtt{R}}$. Following Temme and Verstraete [263] we may draw a close analogy to the Schmidt

form of a quantum state via a decomposition of $P$ as

$$P_{\mathbf{ij}} = \sum_{\mu=1}^{X_\ell} p_\mu^{[\ell]} P_{\mathtt{L}}^{[\ell]}(\mathbf{i}|\mu) P_{\mathtt{R}}^{[\ell]}(\mathbf{j}|\mu). \tag{9.3}$$

This has the elegant interpretation of having $p_\mu^{[\ell]}$ as probabilities, again arranged in non-increasing order with $\mu$, which mix together the conditional marginal probability distributions $P_{\mathtt{L}}^{[\ell]}$ and $P_{\mathtt{R}}^{[\ell]}$. This could equally be expressed as the vector decomposition

$$|\, P\rangle = \sum_{\mu=1}^{X_\ell} p_\mu^{[\ell]} |\, \mathtt{L}_\mu^{[\ell]}\rangle |\, \mathtt{R}_\mu^{[\ell]}\rangle, \tag{9.4}$$

where $|\, \mathtt{L}_\mu^{[\ell]}\rangle$ and $|\, \mathtt{R}_\mu^{[\ell]}\rangle$ have elements $P_{\mathtt{L}}^{[\ell]}(\mathbf{i}|\mu)$ and $P_{\mathtt{R}}^{[\ell]}(\mathbf{j}|\mu)$, respectively. This is similar in form to the Schmidt decomposition in Eq. (3.19) except that $|\, P\rangle$ is expanded in terms of non-negative vectors rather than orthonormal ones. In this case the mutual information between the two subsystems is upper-bounded by the entropy of the mixing probability distribution [263] given by

$$S(\{p_\mu^{[\ell]}\}) = -\sum_{\mu=1}^{X_\ell} p_\mu^{[\ell]} \log\left[p_\mu^{[\ell]}\right]. \tag{9.5}$$

Akin to the truncation of a Schmidt decomposition, an (un-normalised) approximation $|\, \tilde{P}^{[\ell]}\rangle$ to $|\, P\rangle$ can be formed by truncating the decomposition in Eq. (9.3) to the first $\chi$ terms. The $L_1$-norm of the residual is $|| \,|\, P\rangle - |\, \tilde{P}^{[\ell]}\rangle||_1 = \sum_{\mu=\chi+1}^{X_\ell} p_\mu^{[\ell]}$. In Sec. 9.3.2 we will show that for classical systems the error made in computing any observable is upper-bounded by the $L_1$-norm error of the probability vector, analogous to the role of the $L_2$-distance for quantum systems.

The decomposition Eq. (9.3) is in general non-unique and the most significant decomposition at this bipartition is the one with the spectrum $p_\mu^{[\ell]}$ that has the smallest entropy $S_C$, called the entropy cost [263]. This suggests that the entropy

cost $S_C$ should be considered analogous to the entanglement entropy $S_{\mathtt{LR}}$ (introduced in Sec. 3.2.1) since it quantifies both the quality of approximation and correlations. However since $S_C$ only upper-bounds the mutual information the connection between correlations and truncation errors is substantially weaker than that seen in the quantum case.

Just like for quantum systems, repeated application of the decomposition in Eq. (9.4) at all bipartitions can be used to build an approximate sMPS $|\,\tilde{P}\rangle$. In Ref. [263] it was shown that the error is bounded by

$$|||\,P\rangle - |\,\tilde{P}\rangle||_1 \leq \sum_{\ell=1}^{M-1} \sum_{\mu=\chi+1}^{X_\ell} p_\mu^{[\ell]}. \tag{9.6}$$

Exact sMPS solutions for TASEP stationary states find that the entropy minimising spectrum decays very quickly [263] thus indicating in a similar way to the quantum case that a small $\chi$ can be chosen while maintaining an accurate description. Moreover, we saw in Sec. 7.1 that any Gibbs state with a transfer matrix solution is known to have an exact small-$\chi$ description.

Here we note that each decomposition Eq. (9.3) is in fact precisely equivalent to an exact NMF of the matrix $P$, analogous to performing an SVD in the quantum case. The exact NMF can be written as $P = UDV^T$ where the columns of $U$ and $V$ are the normalised non-negative probability vectors in the decomposition Eq. (9.4) and the diagonal elements of $D$ are the normalised probabilities $p_\mu^{[\ell]}$ [264]. As was done with the SVD, $|\,\tilde{P}^{[\ell]}\rangle$ is formed by truncating the inner-dimension of the NMF to $\chi$.

While, as shown above, the NMF may be a more natural choice for decomposing probability vectors at a formal level, it is perfectly possible to employ the SVD on $|\,P\rangle$ in precisely the same way as for a quantum state in Sec. 3.2.1. The $L_2$-norm error is given by Eq. (3.23) and the $L_2$-norm upper-bounds the more relevant $L_1$-

norm with a factor of the square root of the dimension of the vector space [132]. Hence for the approximation $|\,\tilde{P}\rangle$ built in this way

$$||\,|\,P\rangle - |\,\tilde{P}\rangle||_1 \leq \Lambda = d^{M/2}\left[2\sum_{\ell=1}^{M-1}\sum_{\mu=\chi+1}^{X_\ell}(\lambda_\mu^{[\ell]})^2\right]^{1/2}. \tag{9.7}$$

Also, several pleasing features of the SVD still apply, namely its uniqueness and the Eckart-Young theorem, expressed in Eq. (3.24), which ensures that if there exists an exact MPS of some dimension $\chi$ then there also exists an exact SVD-MPS of dimension $\chi$ or less.

A powerful feature of the SVD is the orthogonality constraint on the singular vectors (the columns of $U$ and $V$). However this typically results in the $U$ and $V$ matrices having elements of mixed signs, even for non-negative $P$, and so does not permit any obvious information theoretic interpretation of the components of the decomposition as exists for an NMF. This also means that when $|\,P\rangle$ is truncated to form $|\,\tilde{P}^{[\ell]}\rangle$ negative probabilities may enter our description. The maximum negativity of any element of $|\,\tilde{P}^{[\ell]}\rangle$ is

$$\max_{\mathbf{i},\mathbf{j}}|-\tilde{P}^{[\ell]}_{\mathbf{ij}}| \leq \|\,|\,P\rangle - |\,\tilde{P}^{[\ell]}\rangle\|_\infty \quad \leq \quad \sum_{\mu=\chi+1}^{X_\ell}\lambda_\mu^{[\ell]}. \tag{9.8}$$

Thus while non-negativity is not conserved it is controlled by a similar criterion to the accuracy of the factorisation, namely the smallness of the truncated singular values. This bound is loose. For example in the worst case limit of truncation to $\chi = 1$ non-negativity is in fact preserved due to the Perron-Frobenius theorem [265], which ensures that both the left and right singular vectors associated to the largest singular value of a non-negative matrix are themselves non-negative.

## 9.2 TEBD for stochastic processes

From the way in which we presented the TEBD algorithm in Sec. 3.2.2 it is clear that the numerical procedure does not depend on the whether the vector to be evolved is real or complex, or how it is normalised. The hermiticity of the Hamiltonian was also not important, only its decomposition into nearest-neighbour terms.

Therefore we propose the application of the TEBD algorithm to classical processes exactly as presented in Sec. 3.2.2, with the following changes. Instead of a complex vector $|\,\psi\rangle$ normalised as $|||\,\psi\rangle\,||_2 = 1$, our state to be evolved is a real vector $|\,P\rangle$ normalised as $|||\,P\rangle\,||_1 = 1$. The operator $\hat{H}$ is not Hermitian but rather infinitesimal stochastic (defined in Sec. 6.3). We assume again that the Hamiltonian decomposes as

$$\hat{H} = \sum_{\ell=1}^{M-1} \hat{h}_{\ell,\ell+1}, \tag{9.9}$$

where each nearest-neighbour two-site term $\hat{h}_{\ell,\ell+1}$ is infinitesimal stochastic. In all evolution operators $-\mathrm{i}t/\hbar$ is replaced by $t$. The full evolution operator $\mathrm{e}^{\hat{H}t}$ and similarly the two-site nearest-neighbour gates $\hat{S}_\ell = \exp(\hat{h}_{\ell,\ell+1}\delta t/2)$ are then not unitary. Rather, as we show in App. 9.A, they are stochastic. An operator $\hat{O}$ is stochastic if its elements $\langle \mathbf{i}'|\hat{O}|\mathbf{i}\rangle$ are non-negative and its columns sum to unity $\sum_{\mathbf{i}'}\langle \mathbf{i}'|\hat{O}|\mathbf{i}\rangle = 1$. When acting on a non-negative vector $|\,P\rangle$ normalised according to the $L_1$-norm, a stochastic operator will preserve the non-negativity and $L_1$-norm. This is analogous to the preservation of the $L_2$-norm by a unitary operator.

Important differences between the classical and quantum cases only emerge in the re-compression of the description of the process to an MPS following each two-site operation. This is the approximation of the $d\chi \times d\chi$ matrix $\Theta$ by the factorised matrix $\tilde{\Theta} = UDV^T$ where $D$ is a $\mathbb{R}^{\chi\times\chi}$ diagonal matrix, $U \in \mathbb{R}^{d\chi\times\chi}, V \in \mathbb{R}^{d\chi\times\chi}$ and $\chi$ is the desired MPS dimension (see Sec. 3.2.2). In Sec. 3.2.2 we saw that the truncated SVD is the optimal factorisation if it is a small $L_2$-norm error that is

desired. However this does not imply optimality for the $L_1$-norm error and so it is worth considering other types of factorisation. NMF is a natural candidate due to its analytical properties described in the previous Sec. 9.1. We now investigate its numerical properties.

### 9.2.1 Non-negative matrix factorisation

There is no known algorithm[2] for finding a non-trivial exact NMF solution for an arbitrary non-negative matrix $\Theta$, despite its existence [266], and so an NMF based approximation cannot proceed by truncation from an exact factorisation. Instead an approximation $\tilde{\Theta} = CK^T$ is formed, where $C$ and $K$ are non-negative matrices, by solving directly the reduced rank minimisation problem

$$\min_{\substack{C\in\mathbb{R}_+^{d\chi\times\chi},\\ K\in\mathbb{R}_+^{d\chi\times\chi}}} \left\{ f_{\text{cost}}(\Theta, CK^T) \right\}, \tag{9.10}$$

where $f_{\text{cost}}(\Theta, CK^T)$ is a cost function that quantifies the quality of the approximation to $\Theta$. By forming diagonal matrices $D_C$ and $D_K$ containing the column sums of $C$ and $K$ any NMF can be brought into the form $\tilde{\Theta} = UDV^T$. Here $D$ is a diagonal matrix $D = D_C D_K$ satisfying $\text{tr}(D) = \sum_{\mathbf{ij}} \Theta_{\mathbf{ij}}$, and $U = CD_C^{-1}$ and $V = KD_K^{-1}$ are stochastic [264].

[2] At the time this work was performed (2010) there was no known algorithm for finding an exact NMF. However, recently an algorithm has been found that runs in a time $\mathcal{O}((d\chi)^{2(d\chi)^2})$ [285], which is very inefficient (compare this to the $\mathcal{O}((d\chi)^3)$ scaling of the SVD). In the same work [285] they show that if an algorithm is found that scales as $\mathcal{O}((d\chi)^{2(d\chi)})$ or better then NP-hard problems can be solved in polynomial time. This is considered very unlikely. Therefore although there is now an algorithm for finding an exact NMF it is highly inefficient and further an efficient algorithm is expected to be impossible. This leaves us in the same position as in 2010; going straight to an approximate NMF is the best option.

Common choices for $f_{\text{cost}}$ are

$$f_{\text{cost}}(\Theta, CK^T) = \begin{cases} \sqrt{\sum_{\mathbf{ij}} |\Theta_{\mathbf{ij}} - (CK^T)_{\mathbf{ij}}|^2} \\ \sum_{\mathbf{ij}} \left\{ \Theta_{\mathbf{ij}} \log \left[ \frac{\Theta_{\mathbf{ij}}}{(CK^T)_{\mathbf{ij}}} \right] - \Theta_{\mathbf{ij}} + (CK^T)_{\mathbf{ij}} \right\} , \\ \sum_{\mathbf{ij}} |\Theta_{\mathbf{ij}} - (CK^T)_{\mathbf{ij}}| \end{cases} \tag{9.11}$$

which are the $L_2$-norm of the residual reshaped as a vector $\| \, |\Theta\rangle - |\tilde{\Theta}\rangle \|_2$, the generalised Kullback-Leibler (KL) divergence[3] and the $L_1$-norm $\| \, |\Theta\rangle - |\tilde{\Theta}\rangle \|_1$, respectively [267].

A crucial issue for any minimisation algorithm based on these cost functions is that although they are convex for $C$ and $K$ separately they are not convex for $C$ and $K$ simultaneously, meaning the problem can have many local minima. A common strategy for minimisation is to alternate the optimisation between $C$ and $K$ while the other is fixed, however, such algorithms can at best only guarantee convergence to a local minimum. Furthermore there is no guarantee that the global minimum, if located, is unique. The non-uniqueness of an NMF is evident given that any $\mathbb{R}_+^{\chi \times \chi}$ non-negative monomial matrix $Y$ can also form an alternative NMF as $\tilde{\Theta} = CYY^{-1}K^T$. In App. 9.B we outline some popular and simple NMF algorithms. From numerical tests (not presented in this thesis) we find, as expected, that the $L_1$-norm minimising NMF algorithms consistently deliver the smallest $L_1$-norm errors and so we concentrate on this cost function.

### 9.2.2 Optimality of NMF within TEBD

In Sec. 3.2.2 we showed that TEBD based on the SVD guarantees that the indices of $\tilde{\Theta}$ and $\Theta$ correspond to an orthonormal basis, which in turn ensures the optimality

[3]Despite not being a distance measure the KL divergence does have an information theoretic interpretation. When $\sum_{\mathbf{ij}} \Theta_{\mathbf{ij}} = \sum_{\mathbf{ij}} (CK^T)_{\mathbf{ij}} = 1$ the KL divergence reduces to the relative entropy of the two probability distributions. It then quantifies the expected number of extra bits required to code samples from $\Theta$ when using a code based on $WH^T$ rather than using a code based on $\Theta$.

of the SVD factorisation for the $L_2$-norm error made when applying each two-site operator. For TEBD based on sMPS and NMF, orthogonality is less useful since in general it cannot be imposed on an sMPS due to its non-negativity. Moreover the $L_1$-norm minimised by the NMF is not unitarily invariant and so minimising $\| \, |\,\Theta\rangle - |\,\tilde{\Theta}\rangle \|_1$ does not necessarily minimise $\| \, |\,P\rangle - |\,Q\rangle \, \|_1$, where $|\,P\rangle$ is the exact vector $|\,Q\rangle$ and the approximate one obtained by replacing $\Theta$ with $\tilde{\Theta}$. Despite this, the $L_1$-norm error of an NMF of the local $\Theta$ tensor $\tilde{\Theta}$ can be shown to upper-bound the full $L_1$-norm error of the vectors. To see this we use the fact that the full $L_1$-norm between the exact vector $|\,P\rangle$ and the approximate one $|\,Q\rangle$ simplifies in matrix product form because the sMPS of the two states differ only for the two sites $(\ell, \ell+1)$ on which an operator is applied,

$$\begin{aligned}
\| \, |\,P\rangle - |\,Q\rangle \, \|_1 &= \sum_{\mathbf{i}} \left| A^{[1]i_1} \cdots A^{[\ell-1]i_{\ell-1}} (\Theta^{i_\ell i_{\ell+1}} - \tilde{\Theta}^{i_\ell i_{\ell+1}}) A^{[\ell+2]i_{\ell+2}} \cdots A^{[M]i_M} \right| \\
&\leq \sum_{\mathbf{i}} A^{[1]i_1} \cdots A^{[\ell-1]i_{\ell-1}} \left| \Theta^{i_\ell i_{\ell+1}} - \tilde{\Theta}^{i_\ell i_{\ell+1}} \right| A^{[\ell+2]i_{\ell+2}} \cdots A^{[M]i_M} \\
&= T^{[1]} \cdots T^{[\ell-1]} \left\{ \sum_{i_\ell i_{\ell+1}} \left| \Theta^{i_\ell i_{\ell+1}} - \tilde{\Theta}^{i_\ell i_{\ell+1}} \right| \right\} T^{[\ell+2]} \cdots T^{[M]} \\
&\leq \sum_{\mu_{\ell-1}\mu_{\ell+1}} \sum_{i_\ell i_{\ell+1}} \left| \Theta^{i_\ell i_{\ell+1}}_{\mu_{\ell-1}\mu_{\ell+1}} - \tilde{\Theta}^{i_\ell i_{\ell+1}}_{\mu_{\ell-1}\mu_{\ell+1}} \right| \\
&= \| \, |\,\Theta\rangle - |\,\tilde{\Theta}\rangle \|_1 . \qquad (9.12)
\end{aligned}$$

Here we have used the non-negativity of tensors $A^{[\ell]}$ along with the fact that the left $T^{[1]} \cdots T^{[\ell-1]}$ and right $T^{[\ell+2]} \cdots T^{[M]}$ products, where $T^{[\ell]} = \sum_{i_\ell} A^{[\ell]i_\ell}$, form a row and column vector, respectively, composed of non-negative elements $\leq 1$.

This property makes the $L_1$-norm NMF very attractive as a factorisation for use with stochastic processes since it aims to minimise a quantity that upper-bounds the $L_1$-norm error $\| \, |\,P\rangle - |\,Q\rangle \, \|_1$, which in turn bounds expected value errors. The

remaining question, which we will answer in this chapter, is whether this global minimisation is achievable in practice or if local minima are found instead.

## 9.3 Expected values for stochastic processes

### 9.3.1 Calculating expected values

An expected value for an observable[4] in a classical system takes the form $\langle O \,|\, P \rangle$, where $|\, P\rangle$ is the state vector and $|\, O\rangle = \sum_{\mathbf{i}} O_{\mathbf{i}} \,|\, \mathbf{i}\rangle$ is a vector with real coefficients $O_{\mathbf{i}}$ giving the value that the observable takes for each configuration $\mathbf{i}$. In accordance with earlier chapters we denote this expected value $\mathrm{E}[O, P]$.

In this part we will be interested observables represented by product states $|\, O\rangle = \bigotimes_\ell |\, O^{[\ell]}\rangle$, where $|\, O^{[\ell]}\rangle = \sum_{i_\ell} O^{[\ell] i_\ell} \,|\, i_\ell\rangle$. The calculation of $\mathrm{E}[O, P]$, where $|\, P\rangle$ is represented by an MPS of the form of Eq. (9.2), is performed as follows. For each site $\ell$ we compute the tensor $B^{[\ell]}$ with elements $B^{[\ell]}_{\mu_{\ell-1}\mu_\ell} = \sum_{i_\ell} O^{[\ell] i_\ell} A^{[\ell] i_\ell}_{\mu_{\ell-1}\mu_\ell}$, requiring $\mathcal{O}(d\chi^2)$ computations. Then the ordered product of these matrices gives the desired expected value $B^{[1]} B^{[2]} \cdots B^{[M]} = \mathrm{E}[O, P]$, requiring $\mathcal{O}(M\chi^2)$ computations. See Sec. 3.2.3 for the equivalent procedure for expected values with pure quantum states.

### 9.3.2 Expected value errors

Here we consider how the errors in calculating an observable using an approximate probability distribution, such as that given by a truncated MPS, should scale with the $L_1$ and $L_2$-distances. Let $|\, P\rangle$ be the exact probability vector and $|\, Q\rangle$ be an approximate probability vector. The difference between the expected values of the

[4] Here we restrict ourselves to observables that depend on the configuration of the system at a given time. In Ch. 10 we will generalise this to observables that depend on the path the system takes.

observable calculated using the two probability vectors is

$$\mathrm{E}[O,P] - \mathrm{E}[O,Q] = \sum_{\mathbf{i}} O_{\mathbf{i}} \left(P_{\mathbf{i}} - Q_{\mathbf{i}}\right). \tag{9.13}$$

The worst case scenario for this error is that the observable values and the values of the probability error for each configuration are correlated such that each term in the sum has the same sign. In this case the magnitude of the error can equal its upper-bound

$$\begin{aligned} |\mathrm{E}[O,P] - \mathrm{E}[O,Q]| &\leq \sum_{\mathbf{i}} |O_{\mathbf{i}}| \, |P_{\mathbf{i}} - Q_{\mathbf{i}}| \\ &\leq \max_{\mathbf{i}}\{O_{\mathbf{i}}\} || \, |P\rangle - |Q\rangle \, ||_1. \end{aligned} \tag{9.14}$$

So the error in calculating an expected value of an observable is always bounded by the $L_1$-norm error.

However, for errors to scale as badly as this requires complete correlation between the approximation method and the observable. For many observables we expect no such correlation and it is instructive to consider how observable errors typically scale in this case. This typical scaling would be revealed by calculating the magnitude of the observable error Eq. (9.13) and averaging over all possible pairings of observable values $O_{\mathbf{i}}$ with errors $P_{\mathbf{i}} - Q_{\mathbf{i}}$. An analytically tractable approximation to this is to average over all ways of replacing each $O_{\mathbf{i}}$ by a random selection from all $O_{\mathbf{i}}$. Normalising $|Q\rangle$ such that $\sum_{\mathbf{i}} Q_{\mathbf{i}} = 1$, we find the average and root average square

error to be

$$\langle \mathrm{E}[O,P] - \mathrm{E}[O,Q] \rangle = 0, \tag{9.15a}$$

$$\sqrt{\langle (\mathrm{E}[O,P] - \mathrm{E}[O,Q])^2 \rangle} = \sqrt{\left\langle \frac{1}{2^M} \sum_{\mathbf{i}} \left( O_{\mathbf{i}} - \frac{1}{2^M} \sum_{\mathbf{i}'} O_{\mathbf{i}'} \right)^2 \right\rangle} || \, |P\rangle - |Q\rangle \, ||_2. \tag{9.15b}$$

The latter scales with the $L_2$ rather than the $L_1$-norm error, and we conclude that if there is no correlation between observable values and the approximation method then typically the observable error will scale with the $L_2$-norm error. An example of this is shown later in Fig. 9.3(a). Note that due to the possibility of negative probabilities the normalisation $\sum_{\mathbf{i}} Q_{\mathbf{i}} = 1$ is not always the same as $|| \, |Q\rangle \, ||_1 = 1$.

### 9.3.3 TEBD errors revisited

It is clear from the previous subsection that our main interest will be in the build up of the $L_1$-norm error during the TEBD algorithm. Earlier in App. 3.B we considered the build-up of the $L_2$-norm error during unitary evolution. We found that the $L_2$-norm Trotter error, the error due to the breaking down the unitary operator $\mathrm{e}^{-\mathrm{i}\hat{H}\delta t/\hbar}$ into a series of two-site nearest-neighbour unitary operators $\hat{S}_\ell$, scaled as $\mathcal{E}_{\mathrm{ST}} \sim t_f \delta t^2$. The $L_2$-norm truncation error increased at most additively for the truncation error at each gate. This gave us an upper-bound $\mathcal{E}^+_{\mathrm{tr},||\cdot||_2} = \sum_{l=1}^{l_{\max}} \epsilon_l$ for the truncation error that can be calculated efficiently during the algorithm by summing the $L_2$-norm truncation error $\epsilon_l$ made in implementing each two-site operator $\hat{S}_l$. Our proof relied only on the fact that the $L_2$-norm is preserved by unitary evolution and that the second order Suzuki-Trotter expansion of Eq. (3.27) does not affect the unitarity of the evolution.

Very similar results may be obtained for stochastic evolution for the same reasons:

the Suzuki-Trotter expansion does not effect the stochastic nature of the evolution operator and the $L_1$-norm is invariant under the action of a stochastic operator. Hence the $L_1$-norm Trotter error will scale as $\mathcal{E}_{\mathrm{ST}} \sim t_f \delta t^2$ and the worst case scenario for the $L_1$-norm truncation error will be that it equals the sum of the $L_1$-norm errors made for the two-site operations. The $L_1$-norm error for each two-site operation is difficult to compute but can be loosely upper-bounded by $2^{M/2}$ multiplied by the $L_2$-norm error. Thus we obtain an upper-bound for the $L_1$-norm truncation error for the whole evolution $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1} = 2^{M/2}\mathcal{E}^+_{\mathrm{tr},||\cdot||_2}$. This upper-bound is usually a gross overestimate.

## 9.4 TASEP

Having introduced the application of MPS and TEBD to classical stochastic processes, the rest of this chapter deals with analysing its performance on a test system, the TASEP, which we describe in Fig. 9.1. For a detailed review of the TASEP we refer the reader to Refs. [255, 256]. For our purpose it is sufficient to note that the stationary state of the system has been solved analytically [238], while numerical simulations are needed to calculate arbitrary expected values away from the stationary state. Like many of the particle hopping systems the TASEP is modelled as a chain of $M$ lattice sites, each of which can be in one of $d$ configurations, and thus can be described by the formalism introduced in Sec. 9.1. For the TASEP, $d = 2$ and the local configurations $i_\ell = 0, 1$ correspond to empty and occupied sites, respectively.

The TASEP is a Markovian process and so it is characterised by an infinitesimal stochastic Hamiltonian $\hat{H}$. This is strongly connected and thus the TASEP is ergodic

(cf. Sec. 6.4). The Hamiltonian can be written as

$$\hat{H} = \hat{h}_1 + \hat{h}_M + \sum_{\ell=1}^{M-1} \hat{h}_{\ell,\ell+1}, \tag{9.16}$$

where each term is infinitesimal stochastic. The single-site terms $h_1$ and $h_M$ describe the input and output of particles and the two-site nearest-neighbour terms $h_{\ell,\ell+1}$ describe the hopping. By combining the boundary and hopping terms, the Hamiltonian can be brought into the desired form of Eq. (9.9). The Hamiltonian terms can be calculated from the Poisson rates $W_{\mathbf{ii}'}$ using Eqs. (6.10) and are given in Ref. [260].

## 9.5 Small system analysis

### 9.5.1 Comparing NMF and SVD for the TASEP

A physically relevant comparison of the NMF and SVD can be made by considering the stationary probability distribution of the TASEP. The $L_1$-norm errors for the SVD and NMF approximations to this probability distribution bipartitioned at centre of the system are presented in Fig. 9.2(a) for $\alpha/\gamma = \beta/\gamma = 1$ and $M = 10$. This shows that the SVD is accurate to machine precision once $\chi > 5$ indicating that the stationary state has weak enough correlations to permit a considerable amount of compression. The NMF fails to identify an accurate solution as $\chi$ is increased, due to issues with local minima, and gives an error which levels off above $10^{-3}$. Note also that negativity of the approximation found from the SVD only occurs for $\chi = 2$ and is several orders of magnitude smaller than the overall $L_1$-norm error. Finally, the upper-bound to the $L_1$-norm error for the SVD truncation, computed from the $L_2$-norm error, is also plotted and seen to provide a reasonably tight upper-bound.

To move beyond a single matrix factorisation we have also tested how the NMF

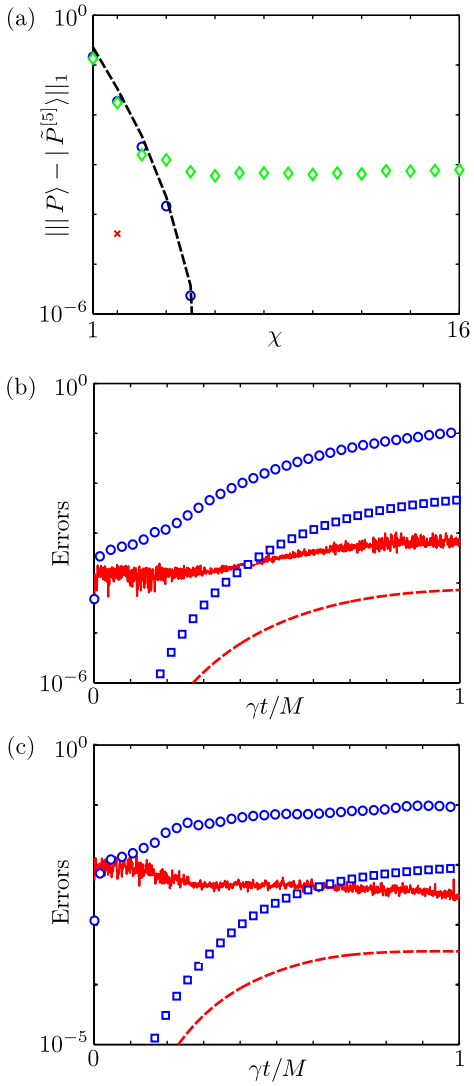


**Figure 9.2**. (a) The accuracy of approximations to the stationary state $|P\rangle$ of the 10-site TASEP with $\alpha = \beta = 1$. The $L_1$-norm error is plotted against $\chi$ for the SVD ($\circ$) and $L_1$-norm minimising NMF ($\diamondsuit$) performed at the central bipartition. For each SVD the $L_1$-norm error due to the sum of all negative elements is shown ($\times$) along with the upper-bound to the $L_1$-norm error (dashed line) derived from the exact $L_2$-norm error. (b) Errors during the evolution of the 10-site TASEP with $\alpha/\gamma = 0.25$ and $\beta/\gamma = 0.75$. The NMF and SVD algorithms are compared for $\chi = 3$ and $\gamma\delta t = 10^{-2}$, where for each NMF four random restarts have been performed. Initially the system was in the state describing a uniform distribution over all configurations. With decimation occurring at the central bipartition only, we plot the upper-bound to the $L_1$-norm error incurred during each time-step, discussed in Sec. 9.3.3, using the SVD (smooth dashed line) and NMF (jagged solid line), and the actual $L_1$-norm error for the SVD ($\square$) and NMF ($\circ$) at time $t$. (c) The same as (b) but with decimation occurring at all contiguous bipartitions.

and SVD behave when used repeatedly to simulate time-evolution. We applied TEBD to the TASEP model in which the two sets of 5 sites on either side of the central bipartition of a 10-site system were merged so only this single bipartition was considered. A calculation of the time-evolution was performed up to a time $\gamma t_f = M$ using both factorisations with $\chi = 3$. This time is how long it takes on average for an unblocked particle to hop across the whole system and so it is the typical time-scale over which the dynamics of the system occurs. At this level of truncation the SVD performed roughly the same as the NMF in the stationary state, as seen in Fig. 9.2(a).

The results for this time-evolution are shown in Fig. 9.2(b). Initially the system is in a $\chi = 1$ state but the NMF fails to find accurate decompositions at small times while the SVD succeeds. At later times the difference in errors at each time-step reduces to an order of magnitude. To examine the optimality issues explained in Secs. 3.2.2 and 9.2.2 we have repeated the calculation, this time using a full 10-site matrix product decimation. We found that while both NMF and SVD errors are slightly worse when decimation occurs at each bipartition, this is not more serious for the NMF than the SVD algorithm, as demonstrated in Fig. 9.2(c). Rather than optimality, the problem faced using current NMF algorithms is the plateau in accuracy seen for increasing $\chi$. This not only prevents any improvements from increasing $\chi$, in contrast to the SVD, but causes unnecessary accumulation of error. Also, to support the use of the SVD, it is found that even with $\chi = 3$ the evolution never produces any negative probabilities over this time. These results strongly suggest that the potential for negativity when using the SVD within the TEBD framework does not present a serious obstacle. For these reasons we shall now focus only on SVD-based decimation.

### 9.5.2 Behaviour of TEBD errors

We next examine in detail the behaviour of the different types of errors for SVD-based TEBD with small systems, in preparation for studying the behaviour of the same errors for larger systems in the next subsection. As a demonstration of the relationship between different error measures, we show in Fig. 9.3(a) the $L_1$ and $L_2$-norm errors of a simulation of the initially empty 15-site TASEP over a time $\gamma t_f = 20M$. Also in the figure we have plotted the expected value errors of three observables: the density at the middle site $\rho_8$ (which takes the value $i_8$ for each configuration $\mathbf{i}$); the current leaving the middle site $J_8$ (which for each configuration $\mathbf{i}$ is equal to 1 if $i_8 = 1, i_9 = 0$, and 0 otherwise); and an observable $O_{\text{rand}}$ that takes a different number between 0 and 1, generated randomly and uniformly for each configuration $\mathbf{i}$, and is expected to exhibit the properties of a generic observable. The random observable scales with the $L_2$-norm error because it is uncorrelated with the MPS approximation method, as expected from the discussion of observable error scaling in Sec. 9.3.2. The current scales slightly above this and the density lies even closer to the $L_1$-norm error, indicating that there is some correlation between TEBD's errors and these two observables. The $L_1$-distance lies close to the maximum factor of $2^{M/2}$ above the $L_2$-distance and, as must be the case, all errors are bounded by the $L_1$-distance.

We may isolate the Trotter error by using a $\chi$ large enough such that the truncation error is negligible. Also it is always possible to efficiently calculate $\mathcal{E}^{+}_{\text{tr},||\cdot||_1}$, introduced in Sec. 9.3.3, which upper-bounds the cumulative $L_1$-norm error due to truncation. The interplay between truncation and Trotter errors can be clearly seen in Fig. 9.3(b), which shows that the $L_1$-distance follows the Trotter error until a time at which the truncation error becomes more significant, as indicated by the increasing bound. $\mathcal{E}^{+}_{\text{tr},||\cdot||_1}$ is an overestimate of the actual truncation error and after

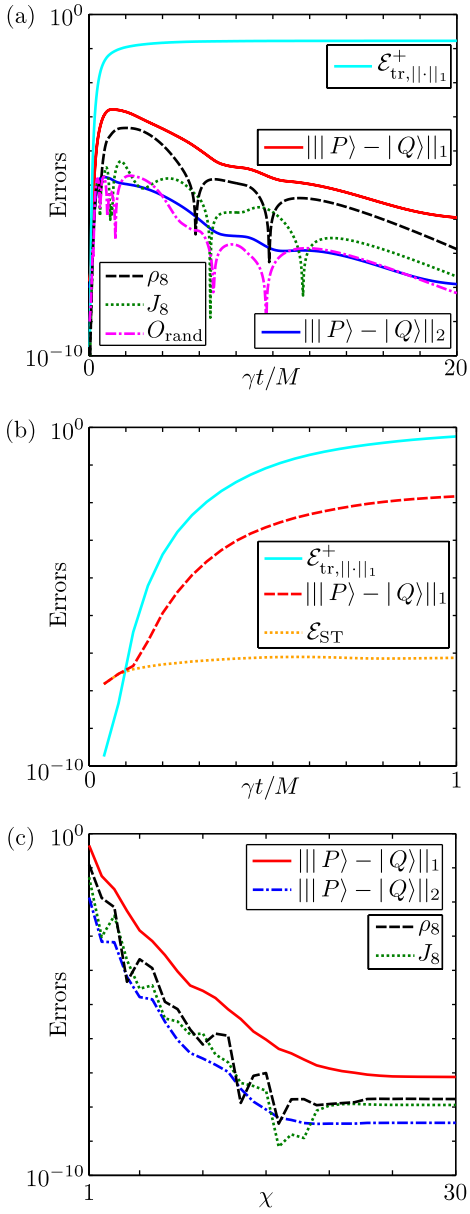


**Figure 9.3**. The errors in simulating the 15-site TASEP with $\alpha/\gamma = 0.3$ and $\beta/\gamma = 0.6$, using TEBD with $\gamma\delta t = 10^{-3}$, as compared to an exact simulation. (a) Using $\chi = 5$ we plot the distance between the approximate and exact probability vectors $||\,|\,P\rangle - |\,Q\rangle\,||_1$ and $||\,|\,P\rangle - |\,Q\rangle\,||_2$ along with the error $|\mathrm{E}[O,P] - \mathrm{E}[O,Q]|$ for each observable $O = \rho_8, J_8, O_{\mathrm{rand}}$. The upper-bound $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$ is also plotted and of the solid lines this takes the largest values, with the $L_1$ and then $L_2$-norm errors below. (b) A zoom in on the small time part of (a) along with the $L_1$-norm error due to the Trotter expansion $\mathcal{E}_{\mathrm{ST}}$ calculated using a $\chi = 100$ simulation. (c) The behaviour of some of these error measures with increasing $\chi$ at $\gamma t = M$.

times of a few $M/\gamma$ it plateaus, as can be seen in Fig. 9.3(a). Since $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_1}$ is cumulative its final value can be used to bound the $L_1$-distance between the TEBD and exact probability distributions, and so all errors, up to this time.

Another consequence of $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_1}$ being cumulative is that it has less relevance for much longer times, when the system approaches the stationary state, especially since the errors do not increase monotonically in time. Consider the evolution of the TASEP shown in Fig. 9.3(a). Initially the system is in the configuration state corresponding to an empty lattice and can be exactly represented by an MPS with $\chi = 1$. It follows there is zero error. In time, correlations are introduced into the system and the transient errors grow to a peak located around the time $\gamma t \approx M$ before decreasing as the stationary state is approached, suggesting that the stationary state is particularly compressible. This shows that even though a low $\chi$ may not be sufficient to describe the transient behaviour of the system accurately, it could still give accurate results in the stationary state. This is because the TASEP is ergodic [247] and the Trotter approximation to the evolution operator (see Eq. (3.27)) appears to preserve this property. So through whichever states the low $\chi$ approximation to the transient behaviour takes the system, whether this be erroneous or not, the approximation to the stationary state will be the same.

For small systems we can use TEBD to calculate any observable of the TASEP precisely, because by increasing $\chi$ while still severely compressing the system we can restrict the $L_1$-norm error to small values. This is shown in Fig. 9.3(c) for a time $\gamma t = M$ near the peak in errors. As $\chi$ is increased, truncation errors shrink until the errors are dominated by the Trotter error. Even though both the $L_1$ and $L_2$-norm errors seem to decrease monotonically with $\chi$, from Fig. 9.3(c) it is clear this is not necessarily the case for observable errors. This non-monotonic behaviour arises because the correlation between errors in the probability vector and an observable's values may vary with $\chi$.

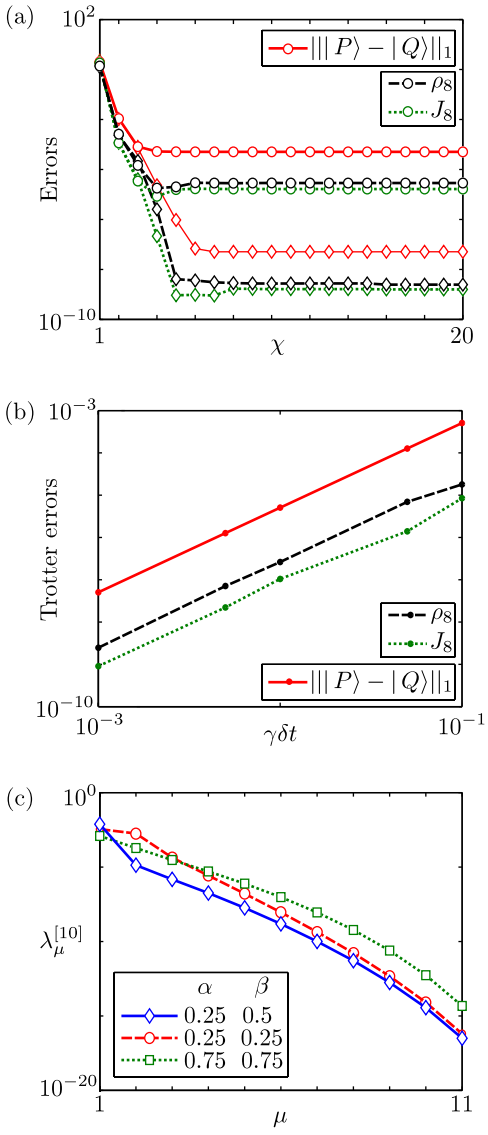


**Figure 9.4**. The stationary state of the TASEP. (a) For the 15-site TASEP with $\alpha/\gamma = 0.3$ and $\beta/\gamma = 0.6$ we have plotted the convergence of error measures to the Trotter error as $\chi$ is increased. This is shown for $\gamma\delta t = 10^{-1}$ ($\circ$) and $\gamma\delta t = 10^{-3}$ ($\diamondsuit$). For both observables $O = \rho_8, J_8$ we have plotted $|\mathrm{E}[O,P] - \mathrm{E}[O,Q]|$. (b) For the same system as (a) this shows the Trotter errors, calculated from the $\chi = 20$ simulations, plotted against $\delta t$ on a log-log axis. To guide the eye, we have connected the points with lines. They show a $\delta t^2$ dependence. (c) The stationary state singular value spectra of the 20-site TASEP, corresponding to bipartite splittings through the middle of the system. Above $\mu = 11$ the singular values are exactly zero.

The behaviour of the errors in time, shown in Fig. 9.3(a), suggests that the stationary state is more compressible than the transient states. To confirm this we evolved the TASEP up to $\gamma t_f = 100M$, by which time the system was effectively in the stationary state. For this state, we show in Fig. 9.4(a) that errors can be driven down to the Trotter error for $\chi$ even as small as 4 when using a time-step $\gamma\delta t = 10^{-1}$. To obtain more accurate results, $\delta t$ can be decreased. The results presented in Fig. 9.4(b) confirm that the scaling of $\mathcal{E}_{\mathrm{ST}}$ is with $\delta t^2$.

The success of the TEBD algorithm in describing the TASEP stationary states accurately with an MPS of small $\chi$ stems from the rapid decay of the singular value spectra of the bipartite splittings of the probability distribution. In Ref. [263] an exact sMPS for every $M$-site TASEP stationary state was obtained with a dimension $\chi = M + 1$. From our discussion in Sec. 9.1 we know an exact SVD-MPS must exist with a dimension less than or equal to that of an exact sMPS. This can be constructed explicitly by repeatedly performing an SVD on the matrices of the sMPS given in Ref. [263]. By doing this we have plotted in Fig. 9.4(c) the spectra of singular values for the 20-site TASEP for a few values of $\alpha$ and $\beta$. The spectra decay super-exponentially before becoming exactly zero after $\mu = M/2 + 1$. Numerical tests suggest that this is the matrix dimension needed for an exact MPS. How these properties extend to larger systems is investigated in the next section.

## 9.6 Scalability

The calculation of the exact $L_1$-distance is intractable for large systems. To overcome this we use three approaches to determine the accuracy of the calculated expected values. First, in the stationary state we compare expected values calculated using TEBD directly with exact analytical results. Second, away from stationary state, we use $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_1}$ to upper-bound the $L_1$-distance. Last, we show for specific observables

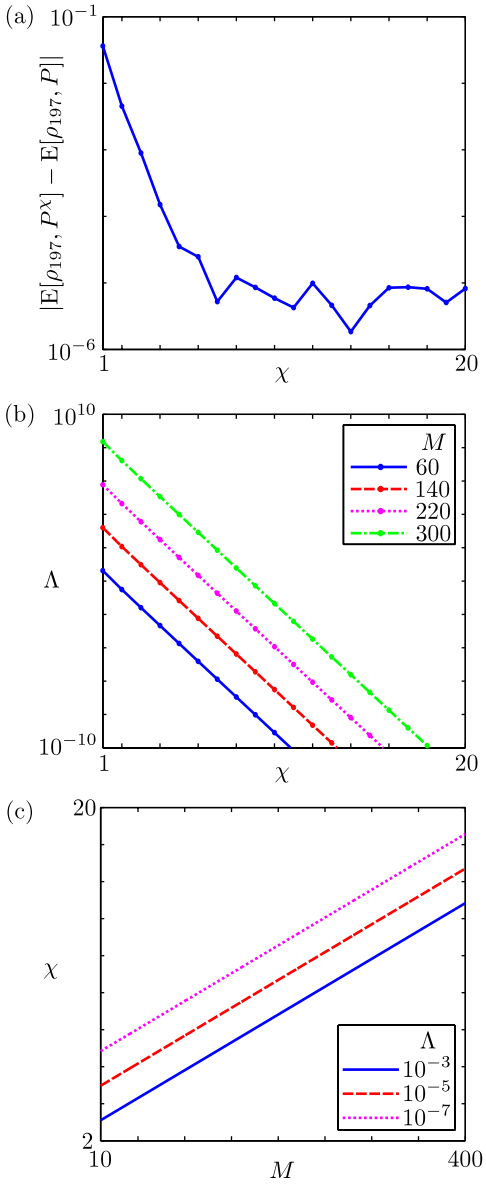


**Figure 9.5.** (a) The convergence of the TEBD density expected value $\mathrm{E}[\rho_{197}, P^\chi]$ to the analytical result $\mathrm{E}[\rho_{197}, P]$ for $M = 200$. The system is in the stationary state for $\alpha/\gamma = 0.349$ and $\beta/\gamma = 0.537$, having used $\gamma\delta t = 5 \times 10^{-3}$. (b) For fixed $\alpha/\gamma = 0.3$ and $\beta/\gamma = 0.6$ but varying $\chi$ and $M$, we plot $\Lambda$ using the exact results from Ref. [263]. (c) For the same system as in (b) this shows the MPS dimension $\chi$ needed to ensure an SVD-MPS exists with $L_1$-distance less than $\Lambda$, as a function of $M$.

that as $\chi$ is increased each expected value converges.

With the first approach in mind, we simulated the TASEP on a 200-site lattice for a total time of $100M/\gamma$. This is long enough to reach the stationary state, which from our analysis of small systems is when we expect the most accurate results to be obtained. In Fig. 9.5(a) we have plotted a comparison between TEBD and exact analytical expected values for the density at a site close to the exiting boundary of the system. Even for such a large system the error is Trotter limited for a $\chi$ of 7. This observable is actually unusual. For the vast majority of expected values considered, e.g., current, density and arbitrary two-site correlations, the Trotter error was reached for $\chi = 1$. This signifies the extreme compressibility of these states, which is also responsible for the success of mean-field theory on these systems. It is also worth noting that the non-monotonic behaviour of errors with $\chi$, observed for smaller systems, is present for larger systems.

To understand the compressibility of the stationary states it is instructive to calculate the bound on the $L_1$-norm $\Lambda$, introduced in Eq. (9.7), using the method for calculating singular spectra from the exact stationary sMPS [263], discussed in Sec. 9.5.2. This gives us an upper-bound to the $L_1$-distance that may be achieved by an SVD-MPS constructed in the way described in App. 3.A. Although this construction is not exactly what is performed by TEBD, the accuracy of both depend on the same properties. We have calculated $\Lambda$ analytically and plotted the results in 9.5(b) for several system sizes. Except for some super-exponential decay when $\chi \approx M/2 + 1$ it exhibits a nearly perfect exponential decay in $\chi$ and an exponential increase in $M$ of the form $\Lambda = \kappa_1^M \kappa_2^{-\chi}$, with $\kappa_1 = 1.0770$ and $\kappa_2 = 11.7797$. Hence the MPS dimension needed to describe the stationary state with errors less than $\Lambda$ is given by $\chi = M \log_{\kappa_2}(\kappa_1) - \log_{\kappa_2}\Lambda$, as plotted in Fig. 9.5(c). We find that accurate MPS approximations to the stationary states of large systems can be obtained using a $\chi$ of order 10 due to the rapid decay of the singular value spectra of the stationary

states, in line with the TEBD results we have obtained.

A potential limitation to the scalability of the TEBD algorithm using the SVD is the exponential scaling with $M$ of the prefactor bounding the $L_1$-norm via the $L_2$-norm, which appears in Eq. (9.7). This indicates that in order to maintain a given bound on the $L_1$-norm the simulation must be performed with exponentially increasing $L_2$-norm accuracy with $M$. The extremely rapid decay of the singular spectrum for TASEP stationary states indicates that an $L_1$-norm bound can be maintained by only a very moderate increase in $\chi$. However, we note that in order to exploit this property it may be necessary to use higher precision arithmetic than the standard 64-bit double precision. Indeed for even moderately large systems, above 50 sites or so, all but the first few singular values have a ratio to the largest that is below double machine precision. TEBD will not be able to reproduce these values. This is a similar problem to what was experienced in [31] for the reaction-diffusion model, but in our case numerical instabilities do not arise from this imprecision.

When away from the stationary state, we cannot calculate the exact spectrum or expected values. However, we can calculate $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$ and we have done this for a range of $M$ and $\chi$ to produce Fig. 9.6(a). The results show that, to ensure a small error, a larger $\chi$ is needed for larger $M$. We can also deduce, for instance, that the error in calculating the expected value of any observable for $M = 50$ using $\chi = 75$ will not exceed $10^{-4}$. In light of this we have plotted in Fig. 9.6(b) the time-evolved current profile of an initially empty 50-site TASEP for a time $\gamma t_f \approx 5M$ using a $\chi$ of 100. The system fills up as particles injected into the first site hop across the system and stationary state values are almost reached when $\gamma t = 5M$. From $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$ we can be sure that all values are correct to a factor less than $10^{-5}$, despite evolving through the transient behaviour where the largest errors are expected to appear. Note that in this argument we have ignored the Trotter error, which we have found to be much smaller than the truncation error and easily controlled by reducing $\delta t$.

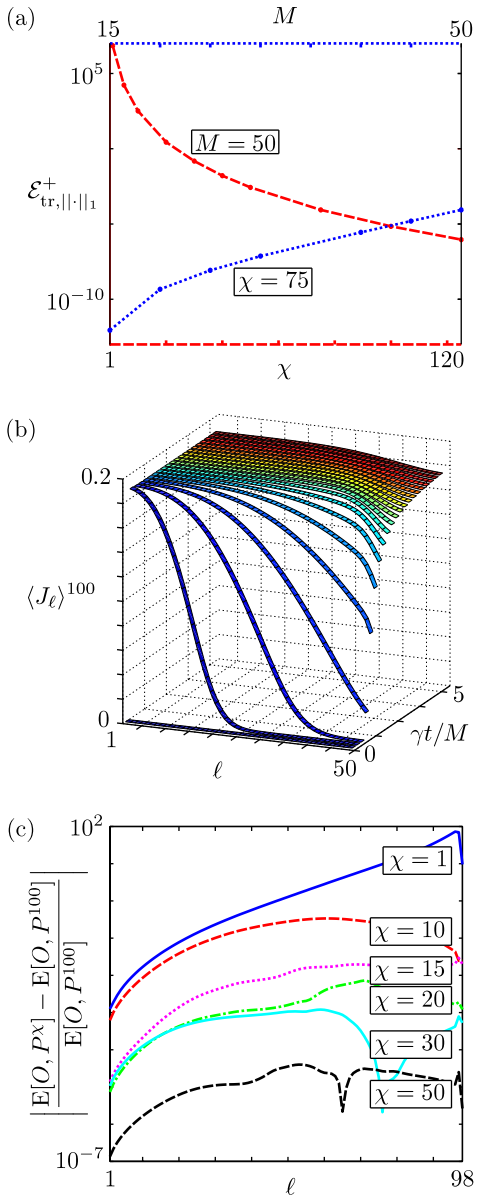


**Figure 9.6**. (a) The dependence of $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$ on $M$ and $\chi$, corresponding to the upper and lower x-axes respectively. We used $\alpha/\gamma = 0.3$ and $\beta/\gamma = 0.6$, taking the initial configuration state to be the empty lattice and setting $\gamma\delta t = 5 \times 10^{-3}$. Each simulation was ran up to $\gamma t_f = 10M$ by which time the value of $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$ had plateaued. (b) The evolution of the current expected values, calculated with $\chi = 100$ for $M = 50$, $\alpha/\gamma = 0.25$ and $\beta/\gamma = 0.25$. The errors are bounded by $\mathcal{E}^+_{\mathrm{tr},||\cdot||_1}$, which plateaus at $6 \times 10^{-6}$. (c) The relative errors between the expected values of the products of the densities of three sequential sites $O = \rho_\ell \rho_{\ell+1} \rho_{\ell+2}$ and the $\chi = 100$ result, showing a convergence to within $10^{-5}$ as $\chi$ is increased to 50. We used $\alpha/\gamma = 0.25$, $\beta/\gamma = 0.25$ and calculated expected values at time $\gamma t = M$. The time-step used was $\gamma\delta t = 5 \times 10^{-3}$ and at $t = 0$ the lattice was empty.

Importantly, $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_1}$ is an upper-bound and so it is sufficient but not necessary for it to take small values. As we saw in Fig. 9.3(a) this upper-bound is often a gross overestimation of the actual $L_1$-norm error as it is the cumulative sum of upper-bounds to the $L_1$-norm error in each two-site operation. Motivated by this, we have used TEBD to simulate the non-equilibrium dynamics of larger systems for which $\mathcal{E}^{+}_{\mathrm{tr},||\cdot||_1}$ alone could not guarantee accurate results. We considered how quickly expected values converge with $\chi$, and calculated this convergence for systems at a time $\gamma t = M$, which in the investigation of small systems was found to be a time close to that with the largest error. As shown in Fig. 9.6(c), a numerically calculated expected value for $M = 100$ converges such that the $\chi = 50$ and 100 results differ only by a fraction of $10^{-5}$. Hence only a moderate $\chi$ is needed to accurately calculate the expected value of this observable, even for transient states. This observable $\rho_\ell \rho_{\ell+1} \rho_{\ell+2}$ could, for example, be used to indicate the presence of a traffic jam with the TASEP used to model a single carriageway. Similar results were obtained for a range of observables, indicating that the TEBD algorithm is able to accurately simulate the transient dynamics of the TASEP for these system sizes.

## 9.7 Discussion

In this chapter we have investigated in detail the application of the TEBD algorithm to classical stochastic processes undergoing Markovian dynamics. A very reasonable approach to describe probability distributions is to use an sMPS [263], which is manifestly non-negative. We have shown that TEBD can produce approximations to the dynamics of stochastic processes within the sMPS class by modifying only the factorisation method, namely switching it to NMF. Unfortunately, current NMF algorithms find it difficult to locate accurate low rank solutions due to the non-convexity of the factorisation. However issues of non-optimality due to the non-

orthogonality of the MPS do not seem to cause problems, so if future algorithms overcome the former problem NMF-based TEBD could become a practical option. The main approach explored here was to relax the constraint to be explicitly non-negative by using instead a general real MPS and applying the SVD method for the evolution in an essentially identical fashion to quantum problems. We showed that potential issues involving the proliferation of negativity and that the $L_2$ rather than $L_1$-distance is minimised turn out not to present serious obstacles to the SVD-based algorithm for the system sizes considered.

With the TASEP as our test system, we demonstrated the accuracy and applicability of TEBD for non-equilibrium systems. We focused our investigation on the behaviour of errors with truncation $\chi$ and system size $M$. An interesting feature was the lack of a monotonic dependence of expected value errors on $\chi$, though in all cases considered, with system sizes up to the low hundreds, results converged. For the stationary state the comparisons with analytically calculated expected values were excellent and we obtained negligible errors for $\chi$ less than 10. Away from the stationary state we introduced an upper-bound to the $L_1$-distance, and thus all errors, which could be efficiently calculated as part of the algorithm when simulating any system, not just the TASEP. This was used to ensure that some of our 50-site simulations had negligible errors. Even for large systems, for which this bound could not guarantee small errors, we found that expected values converged for a large range of observables.

This work suggests that TEBD is a viable candidate for studying the time-evolution of a large class of non-equilibrium stochastic processes whose Markovian dynamics are governed by an infinitesimal stochastic Hamiltonian consisting of at most two-site nearest-neighbour terms. Since TEBD is easily adjusted to include many rates with time and space dependence as well as being extendable to network geometries with a bounded tree-width [139, 140], it could prove to be a powerful

simulation option. The accuracy of TEBD shown here also indicates that there is great scope to adapt other more sophisticated MPS/tensor network methods to classical stochastic processes and extend simulations to higher dimensions. These include the use of matrix product operators to express the evolution for longer ranged processes [268, 269, 270, 271] and renormalisation inspired variational techniques, which have already been used with great success for both stationary and dynamical calculations of quantum systems [141, 142].

## 9.A Stochastic and infinitesimal stochastic operators

Let us first show that the Hamiltonian $\hat{H}$ appearing in the Schrödinger equation $\partial \,|\, P(t)\rangle \,/\partial t = \hat{H} \,|\, P(t)\rangle$ must be infinitesimal stochastic in order for the $L_1$-norm and non-negativity to be conserved. First, note that unless $\langle \mathbf{i}'|\hat{H}|\mathbf{i}\rangle \geq 0$ then setting $|\, P(t)\rangle = |\, \mathbf{i}\rangle$ will result in the probability $P_{\mathbf{i}'}$ becoming negative. Second, applying $\langle \mathbf{i}\,|$ from the left, summing the Schrödinger equation over $\mathbf{i}$, and inserting the resolution of the identity results in

$$\frac{\partial}{\partial t}|| \,|\, P\rangle \,||_1 = \sum_{\mathbf{i}} \frac{\partial}{\partial t} P_{\mathbf{i}}(t) = \sum_{\mathbf{i}'} \sum_{\mathbf{i}} \langle \mathbf{i} \,|\, \hat{H} \,|\, \mathbf{i}'\rangle \, P_{\mathbf{i}'}(t). \tag{9.17}$$

For this to equal zero for arbitrary non-negative $P_{\mathbf{i}'}(t)$ we must have $\sum_{\mathbf{i}} \langle \mathbf{i} \,|\, \hat{H} \,|\, \mathbf{i}'\rangle = 0$. Together these are the requirements for $\hat{H}$ to be infinitesimal stochastic.

Similarly, it is easy to show that in order to ensure the non-negativity and $L_1$-norm of a probability vector $|\, P\rangle$, an operator $\hat{O}$ acting on it as $|\, P'\rangle = \hat{O} \,|\, P\rangle$ must be stochastic. First, if $\langle \mathbf{i}' \,|\, \hat{O} \,|\, \mathbf{i}\rangle < 0$ then $|\, P(t)\rangle = |\, \mathbf{i}\rangle$ would lead to $P'_{\mathbf{i}'} < 0$ and therefore violate non-negativity. Second, it is simple to show that $\sum_{\mathbf{i}} P'_{\mathbf{i}} = \sum_{\mathbf{i}} \sum_{\mathbf{i}'} \langle \mathbf{i} \,|\, \hat{O} \,|\, \mathbf{i}'\rangle \, P_{\mathbf{i}'}$. Setting $P_{\mathbf{i}} = \delta_{\mathbf{i},\mathbf{i}''}$ we must have $\sum_{\mathbf{i}} \langle \mathbf{i} \,|\, \hat{O} \,|\, \mathbf{i}'\rangle = 1$ in order for

$\sum_{\mathbf{i}} P'_{\mathbf{i}} = 1$. These two conditions define a stochastic operator.

Finally let us show that if $\hat{H}$ is infinitesimally stochastic then $\mathrm{e}^{\hat{H}t}$ is stochastic. Using a well-known expansion, the evolution operator may be written as

$$\mathrm{e}^{\hat{H}t} = \lim_{m\to\infty} \left( \mathbb{1} + \frac{\hat{H}t}{m} \right)^m . \tag{9.18}$$

We will show that in the limit $m \to \infty$ the operator $\mathbb{1} + \hat{H}t/m$ is stochastic and therefore so is its product $(\mathbb{1} + \hat{H}t/m)^m$. To see this we use the infinitesimally stochastic property that the columns of $\hat{H}$ add to zero to show that

$$\sum_{\mathbf{i}} \langle \mathbf{i} | \left( \mathbb{1} + \frac{\hat{H}t}{m} \right) | \mathbf{i}' \rangle = 1, \tag{9.19}$$

which is one of the conditions for the operator $\mathbb{1} + \hat{H}t/m$ to be stochastic. The other is the non-negativity of $\langle \mathbf{i} | (\mathbb{1} + \frac{\hat{H}t}{m}) | \mathbf{i}' \rangle$. For $t > 0$ the off-diagonal elements will be non-negative because the off-diagonal elements of $\hat{H}$ are non-negative. This will also be true for the diagonal elements in the limit $m \to \infty$ if the elements of $\hat{H}$ are finite. This completes the proof.

## 9.B Algorithms for non-negative matrix factorisation

The lack of convexity and non-uniqueness of the NMF problem for the three most common cost functions in Eq. (9.11) has prompted a variety of algorithms to be proposed [272]. These algorithms tackle the minimisation problem directly for the required rank $\chi$ using alternating least-squares, conjugate gradients, multiplicative updates or projected gradient descent [272, 262, 273]. Here we shall briefly mention some details for the latter two approaches. The projected gradient descent method

additively updates the elements of the $C$ and $K$ matrices as

$$C_{\mathbf{i}\mu} \mapsto C_{\mathbf{i}\mu} + \iota_C \frac{\partial f_{\text{cost}}}{\partial C_{\mathbf{i}\mu}}, \tag{9.20a}$$

$$(K^T)_{\mu\mathbf{j}} \mapsto (K^T)_{\mu\mathbf{j}} + \iota_K \frac{\partial f_{\text{cost}}}{\partial (K^T)_{\mu\mathbf{j}}}, \tag{9.20b}$$

where $\iota_C$ and $\iota_K$ are appropriately chosen descent step-sizes. For the $L_1$-norm cost function the relevant derivatives are

$$\frac{\partial f_{\text{cost}}}{\partial C_{\mathbf{i}\mu}} = -\sum_{\mathbf{j}} \frac{(\Theta - CK^T)_{\mathbf{ij}}}{|(\Theta - CK^T)_{\mathbf{ij}}|} (K^T)_{\mu\mathbf{j}}, \tag{9.21a}$$

$$\frac{\partial f_{\text{cost}}}{\partial (K^T)_{\mu\mathbf{j}}} = -\sum_{\mathbf{i}} C_{\mathbf{i}\mu} \frac{(\Theta - CK^T)_{\mathbf{ij}}}{|(\Theta - CK^T)_{\mathbf{ij}}|}. \tag{9.21b}$$

Under these update rules there is no explicit preservation of the non-negativity of $C$ and $K$ so typically a projection of the updated matrices is made to the positive orthant at each descent step by setting any negative elements to zero. Furthermore there is no preservation of the unit $L_1$-norm of $\Theta$ within the approximation $CK^T$ so this too has to be explicitly enforced [272].

Gradient descent can also be formulated for the $L_2$-norm and the KL divergence. However, by making step sizes $\iota_C$ and $\iota_K$ depend on the matrix element being minimised via specially chosen functions of the current matrices $C$ and $K$, so-called multiplicative algorithms can be derived [262, 267]. For the KL divergence cost function the updates to be applied are

$$C_{\mathbf{i}\mu} \mapsto \frac{C_{\mathbf{i}\mu}}{\sum_{\mathbf{j}} (K^T)_{\mu\mathbf{j}}} \sum_{\mathbf{j}} \left\{ (K^T)_{\mu\mathbf{j}} \frac{\Theta_{\mathbf{ij}}}{(CK^T)_{\mathbf{ij}}} \right\}, \tag{9.22a}$$

$$(K^T)_{\mu\mathbf{j}} \mapsto \frac{(K^T)_{\mu\mathbf{j}}}{\sum_{\mathbf{i}} C_{\mathbf{i}\mu}} \sum_{\mathbf{i}} \left\{ C_{\mathbf{i}\mu} \frac{\Theta_{\mathbf{ij}}}{(CK^T)_{\mathbf{ij}}} \right\}. \tag{9.22b}$$

Under these transformations the KL divergence is non-increasing and is invariant if

and only if $C$ and $K$ are at a stationary point [267].

Overall, we find that the above NMF algorithms are highly sensitive to the initial conditions. This means that to get reasonable results many random restarts are required and considerable fluctuations in the final cost function value are observed. They also often show slow convergence (if at all) and are generally much slower than the SVD. For the most relevant case of the $L_1$-norm cost function we have found that using a random initial matrix driven to a KL divergence solution often provides a much higher quality result than simply applying the gradient descent to random initial matrices. Despite this we have found that, for our task, the SVD is superior to current NMF algorithms not only at moderate truncations but also at identifying when a near-exact low-rank solution exists[5]. This is an essential ingredient of an effective decimation algorithm.

[5]We note that for data-mining applications the SVD is typically used to benchmark what rank reduction would be acceptable for a subsequent NMF, precisely because it identifies a suitable rank which contains most of the data.

# Chapter 10

# Path-dependence and high variance

Having demonstrated the applicability of time-evolving block decimation (TEBD) for simulating Markovian stochastic processes in the previous chapter, we now find an example for which TEBD outperforms other simulation methods[1]. As discussed in Ch. 8, most techniques for estimating the expected values of physical observables appearing in stochastic processes are sampling based, e.g., dynamical Monte Carlo (DMC). For high variance observables they require the use of judiciously designed variance reduction techniques, often requiring prior intuition about the process. However our TEBD approach requires no prior intuition and at first glance is not expected to perform any worse for high variance observables. Therefore in this chapter we will investigate the efficacy of TEBD with high variance observables. In addition we adapt TEBD for simulating path-dependent observables. This is a new use of matrix product states (MPS) in classical stochastic systems.

Our test observable and the focus of this chapter is the expected value of the exponential of the work done whilst quenching the magnetic field acting on a thermalising one-dimensional Ising model. Such an observable is chosen firstly because its expected value in some cases is simply expressed in terms of equilibrium free energies using Jarzynski's equality (cf. Sec. 7.2), which can be computed exactly using

[1]The results in this chapter are being prepared for publication [5].

the transfer matrix method (cf. Sec. 7.1). This allows us to ascertain the accuracy of our numerics using exact results. Secondly the variance of the observable scales exponentially with system size, making it an excellent test of how our numerics cope with high variance.

We find that the resources required by TEBD to calculate the expected value scale at most linearly in system size. In contrast, it follows from the scaling of the variance that basic DMC would require resources scaling exponentially in system size. Our work can then be interpreted as an efficient numerical simulation of the process described by Jarzynski's equality as well as a demonstration of exponential speed-up beyond a basic DMC simulation.

We begin in Sec. 10.1 by introducing the physical model that we simulate. Then in Sec. 10.2 we adapt TEBD for use with path-dependent observables before discussing the behaviour of errors in Sec. 10.3. Finally we show how our physical model can be brought into the form appropriate for simulation with the modified TEBD algorithm in Sec. 10.4 and present our simulation results in Sec. 10.5. To close, we summarise our findings in Sec. 10.6.

## 10.1 Work in a thermalising Ising model

In this section we begin by introducing the Markovian process that we will simulate. It is based on the Ising model and Glauber dynamics. We also introduce the observable, the exponential of the work done during a change in the magnetic field.

### 10.1.1 Ising model

The Ising model [274, 275] has received a remarkable amount of attention since it was introduced nearly nine decades ago to describe ferromagnetism. In its simplest one-dimensional form the configuration space of the model is a collection of con-

figurations $\mathbf{i} = (i_1, \ldots, i_M)$ describing configurations of $M$ spins, each of which is labelled by one of two possible values $i_\ell = 0, 1$ corresponding to spin down and spin up, respectively. This multi-binary configuration space has since been used to describe, for example, neural networks [276], cooperation [277] and segregation [278]. Here we will also use the interpretation in which $\ell$ labels $M$ traders in a stock market with $i_\ell = 0, 1$ corresponding to a sell or buy order, respectively [279, 280].

The defining feature of the Ising model is the assigning of an energy

$$E_{\mathbf{i}}(\lambda) = -J(\lambda) \sum_{\langle \ell\ell' \rangle} (2i_\ell - 1)(2i_{\ell'} - 1) - B(\lambda) \sum_{\ell} (2i_\ell - 1), \tag{10.1}$$

to each configuration $\mathbf{i}$, where again $\langle \ell\ell' \rangle$ denotes a sum over neighbours. For spins, the exchange $J$ represents half the difference in energy between aligned and anti-aligned neighbouring spins, while the bias $B$ represents half the difference in energy between a spin that is aligned or anti-aligned with an external magnetic field.

It is interesting that the energies $E_{\mathbf{i}}(\lambda)$ and the Gibbs distribution $P_{\mathbf{i}}(\lambda) = \exp[-\beta E_{\mathbf{i}}(\lambda)]/Z(\lambda)$ have interpretations beyond statistical physics. For the example of traders in a stock market, $\beta J$ and $\beta B$ quantify a known correlation between traders and the influence of an external bias, respectively. Specifically, given no other information, we expect that each trader is $\tanh(\beta B)$ more likely to buy than sell and $\tanh(\beta J)$ more likely to make the same decision as their neighbour. This knowledge is equivalent to knowing the expected energy $\langle E(\lambda) \rangle = \sum_{\mathbf{i}} E_{\mathbf{i}}(\lambda) P_{\mathbf{i}} = -\frac{\partial}{\partial \beta} \ln Z(\lambda)$ (cf. Eq. (6.3)) and so the Gibbs distribution represents the optimal (according to maximum entropy arguments) distribution that reproduces the known correlation and bias (cf. Sec. 6.2).

### 10.1.2 Glauber dynamics

As for the dynamics taking place on our configuration space, we choose to focus on a Markovian process for which the Gibbs distribution $P_{\mathbf{i}}(\lambda)$ is stationary, i.e., $\hat{H}(\lambda)\,|\,P(\lambda)\rangle = 0$ (cf. Sec. 6.4). We ensure this condition is met by choosing a process satisfying detailed balance (cf. Eq. (6.14))

$$W_{\mathbf{ii'}}(\lambda) = \frac{w_{\mathbf{ii'}}}{1 + \mathrm{e}^{-\beta(E_{\mathbf{i}}(\lambda) - E_{\mathbf{i'}}(\lambda))}}, \tag{10.2}$$

for $\mathbf{i} \neq \mathbf{i'}$, where $w_{\mathbf{ii'}} = w_{\mathbf{i'i}}$ are some symmetric rates. Explicitly we consider the type of process often called Glauber dynamics [244] in which non-zero values $w_{\mathbf{ii'}} = w$ are assigned for $\mathbf{i}$ and $\mathbf{i'}$ that differ by the value of a single spin, and the rate $w$ is some constant. All other rates are zero. For spins this represents their thermalisation by a reservoir, the interaction with which induces single spin flips. For traders this represents a process by which they each individually update their decisions based both on the decisions of their neighbours and the external bias.

### 10.1.3 Work done by a varying magnetic field

The observable of which we are interested in calculating the expected value is the exponential of the work done as $\lambda$ is varied. For simplicity and explicitness we focus on the case where $J$ is independent of $\lambda$ and $B = \lambda$, so that to vary $\lambda$ is to vary the magnetic field (in the spin interpretation) or the external bias (in the trader interpretation). If we are interested in the work done between times $\tau_i$ and $\tau_f$ then the observable has the form

$$\mathrm{e}^{-\beta W[\mathbf{s}]} = \exp\left(-\beta \int_0^{t_f} \mathrm{d}t \dot{w}_{\mathbf{s}(t)} \theta(t - \tau_i)[1 - \theta(t - \tau_f)]\right), \tag{10.3a}$$

$$\dot{w}_{\mathbf{i}} = -\dot{\lambda} M_{\mathbf{i}}, \tag{10.3b}$$

where $\dot{w}_{\mathbf{i}}$ is the rate of work done in configuration $\mathbf{i}$, $\theta$ is the Heaviside step function and $M_{\mathbf{i}} = \sum_\ell (2i_\ell - 1)$ is the magnetisation of configuration $\mathbf{i}$.

This observable is an appealing choice for several reasons. First, it is known to have a high variance when the work done along paths differ by much more than $1/\beta$. It is thus difficult to calculate using sampling methods and provides a good test of the TEBD algorithm in the presence of a high variance. Second, in the special case that the system is initially, at time $\tau_i$, in the Gibbs state $|\, P(\lambda_i)\rangle$ corresponding to the initial value of the parameter $\lambda$ then the expected value may be expressed in terms of equilibrium free energies using Jarzynski's equality (see Sec. 7.2). For a one-dimensional Ising model these free energies may be calculated exactly using the transfer matrix method (see Sec. 7.1). We are then able to test the predictions of TEBD in this limit. Third, in the limit that the parameter $\lambda$ is varied instantaneously by $\Delta\lambda$ at some time $\tau_*$ and we are interested in the instantaneous work done at this time, i.e., $\tau_i, \tau_f \to \tau_*$, the observable $\mathrm{e}^{-\beta W[\mathbf{s}]} = \mathrm{e}^{\beta \Delta\lambda M_{\mathbf{s}(\tau_*)}}$ has an interesting interpretation as the trading price of stock [279, 280]. We now describe this interpretation.

In the stock market model the $M$ spins are interpreted as $M$ reactionist traders. Spin up $i_\ell = 1$ represents a trader who would like to buy the stock while spin down $i_\ell = 0$ represents a trader who wants to sell. As a group the reactionist traders then buy at a rate proportional to the magnetisation $M_{\mathbf{i}}$. As we have described above, Glauber dynamics describes a process by which these traders update their orders based on an external bias, such as the news, and the orders of other traders (hence the name reactionists). Let us allow the reactionists to trade with another set of contrasting traders, called fundamentalists, who buy at a rate proportional to $\ln[p_0(t)/p(t)]$, where $p(t)$ is the trading price and $p_0(t)$ the true price of the stock [279, 280]. Matching buyers to sellers from both groups of traders, the stock price is then $p(t) = p_0(t)\mathrm{e}^{\beta\Delta\lambda M_{\mathbf{i}}}$, with the precise value of $\Delta\lambda$ depending on the

proportionality constants in the buying and selling rates of the traders. Therefore the expected value of $\mathrm{e}^{-\beta W[\mathbf{s}]}$ in the case of a sudden change is the expected value of the ratio of the stock price to its true price at the time of the sudden change.

We now go on to describe how we adapt the TEBD algorithm to calculate expected values of path-dependent observables and how we expect it to perform if the observables have a high variance.

## 10.2 TEBD for path-dependent observables

The TEBD algorithm has previously been used to evolve a probability vector through a series of time-steps, calculating the expected value of an observable after each. It has therefore been restricted to so-called configuration-dependent observables, i.e., observables that depend on the configuration at a single time. Here we examine how to extend TEBD to calculate path-dependent observables, i.e., observables that depend on the configuration of the system at many times.

### 10.2.1 Multiplicatively increasing observables

In this chapter we focus on expected values of the form (cf. Eq. (8.1))

$$\mathrm{E}\left[O, \mathcal{P}\right] = \int [\mathrm{D}\mathbf{s}] O[\mathbf{s}] \mathcal{P}[\mathbf{s}], \tag{10.4}$$

where the path-dependent observable $O[\mathbf{s}]$ is multiplicatively increasing. We may write these observables as (cf. Eq. (8.19))

$$O[\mathbf{s}] = \exp\left(\int_0^{t_f} \mathrm{d}t \dot{o}_{\mathbf{s}(t)}\right), \tag{10.5}$$

where $\dot{o}_{\mathbf{i}}$ is the rate at which the logarithm of the observable increases when the system is in configuration $\mathbf{i}$.

The reason for restricting ourselves to multiplicatively increasing observables is that other path-dependent observables are typically either calculable from many configuration-dependent observables or are non-local in time. For example, consider the additively increasing observable

$$O[\mathbf{s}] = \int_0^{t_f} \mathrm{d}t \dot{O}_{\mathbf{s}(t)}, \tag{10.6}$$

where $\dot{O}_{\mathbf{i}}$ is the rate at which the observable increases when the system is in configuration $\mathbf{i}$. The integral in Eq. (10.6) may always be approximated by a sum of configuration-dependent observables at different times, e.g., using the trapezium rule or some other higher order approximation to the integral. Consider another type of observable in which contributions to the observable value are time-non-local. Such observables would be incompatible with TEBD unless it was significantly adapted, since TEBD evolves all paths locally in time.

Our focus will therefore be on multiplicatively increasing observables. We will make use of the notation, introduced in Eq. (8.19), for defining the value

$$O[\mathbf{s}](t) = \exp\left(\int_0^t \mathrm{d}t' \dot{o}_{\mathbf{s}(t')}\right), \tag{10.7}$$

of the observable up to time $t < t_f$.

### 10.2.2 Adjusted master equation

To estimate an expected value of an observable of the type in Eq. (10.5), we work with the distribution

$$P'_{\mathbf{i}}(t) = P_{\mathbf{i}}(t)\mathtt{O}_{\mathbf{i}}(t), \tag{10.8}$$

instead of the probability distribution $P_{\mathbf{i}}$. Here $\mathsf{O}_{\mathbf{i}}(t)$ is the expected value of the observable $O[\mathbf{s}](t)$ (cf. Eq. (10.7)) up to time $t$ for all paths reaching configuration $\mathbf{i}$ at time $t$, i.e.,

$$\mathsf{O}_{\mathbf{i}}(t) = \int_{\mathbf{s}(t)=\mathbf{i}} [\mathrm{D}\mathbf{s}] O[\mathbf{s}](t) \mathcal{P}[\mathbf{s}]. \tag{10.9}$$

By this definition, it is clear that the sum of this distribution at time $t$ satisfies $\sum_{\mathbf{i}} P'_{\mathbf{i}}(t) = \mathrm{E}\left[O(t), \mathcal{P}\right]$ and, in particular, at time $t = t_f$ will equal the desired expected value of Eq. (10.4).

We would now like to derive an adjusted master equation for the time-evolution of the distribution $P'_{\mathbf{i}}$. To do this we return to the discussion of Sec. 6.3 in which we first derived the master equation for $P_{\mathbf{i}}$. There we showed that to first order in $\delta t$ we have

$$P_{\mathbf{i}}(t) = \sum_{\mathbf{i}'} P^{\delta t}_{\mathbf{i}'\mathbf{i}}(t - \delta t) P_{\mathbf{i}'}(t - \delta t), \tag{10.10}$$

where

$$P^{\delta t}_{\mathbf{i}\mathbf{i}'}(t) = \delta_{\mathbf{i}\mathbf{i}'}\big[1 - \delta t \sum_{\mathbf{i}' \neq \mathbf{i}} W_{\mathbf{i}\mathbf{i}'}(t)\big] + (1 - \delta_{\mathbf{i}\mathbf{i}'}) \delta t W_{\mathbf{i}\mathbf{i}'}(t). \tag{10.11}$$

Similarly, the first (and some higher) order changes in $\mathsf{O}_{\mathbf{i}}$ may be written as

$$\mathsf{O}_{\mathbf{i}}(t) = \frac{\begin{array}{r}\Big[P_{\mathbf{i}}(t - \delta t) P^{\delta t}_{\mathbf{i}\mathbf{i}}(t - \delta t) \mathsf{O}_{\mathbf{i}}(t - \delta t)[1 + \dot{o}_{\mathbf{i}}(t - \delta t)\delta t] \\ + \sum_{\mathbf{i}' \neq \mathbf{i}} P_{\mathbf{i}'}(t - \delta t) P^{\delta t}_{\mathbf{i}'\mathbf{i}}(t - \delta t) \mathsf{O}_{\mathbf{i}'}(t - \delta t)\Big]\end{array}}{\sum_{\mathbf{i}'} P_{\mathbf{i}'}(t - \delta t) P^{\delta t}_{\mathbf{i}'\mathbf{i}}(t - \delta t)}, \tag{10.12}$$

which we have obtained by applying Bayes' law. The first term in the numerator corresponds to the contribution to expected value $\mathsf{O}_{\mathbf{i}}(t)$ from paths where $\mathbf{s}(t) = \mathbf{s}(t - \delta t) = \mathbf{i}$ (see Eq. (10.9)). The second term in the numerator corresponds to contributions to the average from paths that have transitioned, i.e., $\mathbf{s}(t) = \mathbf{i} \neq \mathbf{s}(t - \delta t)$. The denominator is the probability of $\mathbf{s}(t) = \mathbf{i}$. Combining Eqs. (10.10)

and (10.12) we have

$$\begin{aligned} P_{\mathbf{i}}(t)\mathsf{O}_{\mathbf{i}}(t) =& P^{\delta t}_{\mathbf{ii}}(t-\delta t)[1+\dot{o}_{\mathbf{i}}(t-\delta t)\delta t]P_{\mathbf{i}}(t-\delta t)\mathsf{O}_{\mathbf{i}}(t-\delta t) \\ &+ \sum_{\mathbf{i}'\neq\mathbf{i}} P^{\delta t}_{\mathbf{i}'\mathbf{i}}(t-\delta t)P_{\mathbf{i}'}(t-\delta t)\mathsf{O}_{\mathbf{i}'}(t-\delta t), \end{aligned} \tag{10.13}$$

up to at least first order in $\delta t$. Finally, dividing by $\delta t$, taking the limit $\delta t \to 0$ and rearranging we obtain

$$\frac{\partial}{\partial t}\,|\,P'(t)\rangle = \hat{H}'\,|\,P'(t)\rangle\,, \tag{10.14}$$

$$\langle\mathbf{i}|\hat{H}'|\mathbf{i}'\rangle = \langle\mathbf{i}|\hat{H}|\mathbf{i}'\rangle + \dot{o}_{\mathbf{i}}(t)\delta_{\mathbf{i}\mathbf{i}'}. \tag{10.15}$$

Here $\hat{H}$ is the infinitesimal stochastic Hamiltonian governing the Markovian dynamics of $|\,P(t)\rangle$, $\hat{H}'$ is the adjusted non-infinitesimal stochastic Hamiltonian and $|\,P'(t)\rangle = \sum P'_{\mathbf{i}}(t)\,|\,\mathbf{i}\rangle$.

Therefore as long as the rate $\dot{o}_{\mathbf{i}}$ depends locally on the configuration $\mathbf{i}$, i.e., is composed of on-site or nearest-neighbour dependencies, then the TEBD algorithm can be used to evolve $|\,P'(t)\rangle$ according to $\hat{H}'$ in the same way it does $|\,P(t)\rangle$ according to $\hat{H}$.

## 10.3 TEBD errors revisited (again)

It is interesting to look at how errors scale for high variance and path-dependent observables. We begin by discussing high variance but configuration-dependent observables and then generalise to path-dependent observables.

Consider the case that $\dot{o}_{\mathbf{i}}(t) \propto \delta(t-\tau_*)$ (corresponding to a configuration-dependent observable) so the observable only increases instantaneously at time $\tau_*$. For times before this instantaneous increase the two distributions $P'_{\mathbf{i}}$ and $P_{\mathbf{i}}$ are

identical. The instantaneous evolution of $P'_{\mathbf{i}}$ at time $\tau_*$ adjusts its norm to its final value $\mathrm{E}\,[O, \mathcal{P}]$. Since the errors made in this instantaneous evolution are negligible in comparison to the errors preceding $\tau_*$ we can associate the error entirely with the probability vector at time $\tau_*$. This is the same situation as in the previous Ch. 9.

For high variance observables there are usually a few configurations $\mathbf{i} \in \mathcal{S}$ that have observable values that have magnitudes $|O_{\mathbf{i}}|$ significantly larger than the rest. To a good degree of approximation we write

$$\begin{aligned} |\mathrm{E}[O, P] - \mathrm{E}[O, Q]| &\lesssim \sum_{\mathbf{i}\in\mathcal{S}} |O_{\mathbf{i}}|\ |P_{\mathbf{i}} - Q_{\mathbf{i}}| \\ &\leq |\mathcal{S}| \max_{\mathbf{i}\in\mathcal{S}}\{|O_{\mathbf{i}}|\}||\,|\,P\rangle - |\,Q\rangle\,||_\infty. \end{aligned} \tag{10.16}$$

Here $|\mathcal{S}|$ is the size of the set $\mathcal{S}$ and the $L_\infty$-norm error is given by $||\,|\,P\rangle - |\,Q\rangle\,||_\infty = \max_{\mathbf{i}} |P_{\mathbf{i}} - Q_{\mathbf{i}}|$. To arrive at the right hand side of the first line we have assumed that the values $|O_{\mathbf{i}}|\ |P_{\mathbf{i}} - Q_{\mathbf{i}}|$ are significant only for $\mathbf{i} \in \mathcal{S}$, which is a fair assumption since we do not expect the errors of a numerical simulation to be orders of magnitude higher for some configurations than others. Importantly the $L_\infty$-norm error, which now (approximately) bounds the expected value errors, is itself bounded by the $L_2$-norm error of the probability distribution (as well as the $L_1$-norm error). Thus when only the errors for a few configurations are important the TEBD in combination with the singular value decomposition is known to be extremely effective for configuration-dependent observables, since it is well-suited for minimising the $L_2$-norm error and this (approximately) bounds the observable errors. In contrast to DMC, which is expected to perform worse for high-variance observables, MPS methods including TEBD may perform better. However, this presumes that $\max_{\mathbf{i}\in\mathcal{S}}\{|O_{\mathbf{i}}|\}$ does not grow too large.

The next step is to consider how this extends to path-dependent observables.

In this case the error is given by the difference of the sums of the correct $P'_{\mathbf{i}}$ and approximate $Q'_{\mathbf{i}}$ distributions at time $t_f$

$$\begin{aligned} |\mathrm{E}[O, P'] - \mathrm{E}[O, Q']| &= |\sum_{\mathbf{i}} (P'_{\mathbf{i}} - Q'_{\mathbf{i}})| \\ &\leq ||\,|P'\rangle - |Q'\rangle\,||_1, \end{aligned} \tag{10.17}$$

where as usual we have written $|Q'\rangle = \sum Q'_{\mathbf{i}} |\mathbf{i}\rangle$. Thus once again the $L_1$-norm is the key bounding quantity.

Things are improved using a similar argument to that in Sec. 9.3.2. The signs of the errors $P'_{\mathbf{i}} - Q'_{\mathbf{i}}$ are very unlikely to be the same. We might expect them to be random, in which case it follows from the first line of Eq. (10.17) that the typical error is given by (but not bounded by) the $L_2$-norm. Therefore we expect errors to be much smaller than the upper-bound provided by the $L_1$-norm.

It would be desirable to extend the result shown in Eq. (10.16) to path-dependent observables, i.e., to prove that the errors for high variance path-dependent observables are (approximately) bounded by the $L_2$-norm. However it is not immediately obvious whether this is possible. Previously for configuration-dependent observables the error has been entirely associated with the probability distribution, while the expected value of the observable may be calculated from this without further error. In contrast, when calculating the expected value of a path-dependent observable we approximately evolve a distribution $P'_{\mathbf{i}} = P_{\mathbf{i}}\mathtt{O}_{\mathbf{i}}$ that depends on both the path the system takes and the observable values. It would be misleading to associate a portion of the error with one or the other since it is their combined distribution that is evolved approximately. This inability to partition the error prevents us from deriving tighter bounds for the error in the case of high variance observables since it is less convincing to argue that only a few values of $|P'_{\mathbf{i}} - Q'_{\mathbf{i}}|$ dominate.

We now focus on the accumulation of errors in the TEBD algorithm, i.e., on the

scaling of the errors due to the finite time-step $\delta t$ and the truncation of the MPS representing $|\, P'(t)\rangle$ after the application of each two-site gate. The first thing to note is that the adjusted Hamiltonian is necessarily time-dependent so the dominant source of error due to the finite time-step $\delta t$ arises from the approximation of the Hamiltonian as piece-wise constant and is therefore first order in $\delta t$.

Second, in App. 3.B (Sec. 9.3.3) we were able to show that the truncation error made after each step is additive in the case that the error is quantified by the $L_2$-norm ($L_1$-norm) error and the Hamiltonian is Hermitian (infinitesimal stochastic). This was used to provide an upper-bound for the error. It relied on the fact that the $L_2$-norm ($L_1$-norm) is conserved by unitary (stochastic) operators of the form $\mathrm{e}^{-\mathrm{i}\hat{H}t/\hbar}$ ($\mathrm{e}^{\hat{H}t}$). However, for path-dependent expected values we have found no analogous measure of error that is conserved by operators of the form $\mathrm{e}^{\hat{H}'\delta t}$, where the adjusted Hamiltonian $\hat{H}'$ is infinitesimal stochastic with an additional diagonal part (see Eq. (10.15)). An upper-bound to the error can be derived by assuming the worst case scenario that the entire truncation error accumulates in the path $\mathbf{s}$ contributing the largest $O[\mathbf{s}]\mathcal{P}[\mathbf{s}]$ to the expected value. Unfortunately this upper-bound is so loose that it is effectively useless.

This does not mean that path-dependent calculations using TEBD are expected to be inaccurate. We have already seen that errors are expected to grow with the $L_2$-norm error for which TEBD is particularly optimal. It just means that we are unable, so far, to provide efficiently-calculable bounds that confirm the accuracy of the calculations from first principles. It is still possible to perform the less rigorous but more common accuracy-check of decreasing $\delta t$ and increasing $\chi$ until results converge. We now move away from error analysis, leaving it as a topic for future study.

## 10.4 Formulation of the numerical problem

We next show how the process and observable discussed in Sec. 10.1 may be brought into the form appropriate for the adjusted TEBD algorithm described in Sec. 10.2. We start by changing notation slightly. Rather than individual spins, we instead focus on pairs of neighbouring of spins (called superspins). In this viewpoint configurations are labelled by $\mathbf{i} = (i_{1,2}, i_{3,4}, \ldots, i_{M-1,M})$, where $M$ must be even. Here $i_{2\ell-1,2\ell} = (0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ for $\ell = 1, \ldots, M/2$. Each single spin-flip rate appearing in Eq. (6.14) then only depends on the values of a pair of neigbouring superspins. Thus in the superspin picture the infinitesimal stochastic Hamiltonian $\hat{H}$ is composed of two-site nearest-neighbour terms only. Similarly the rate $\dot{w}_{\mathbf{i}} = -\dot{\lambda} M_{\mathbf{i}}$ of relevant work done only depends on neigbouring superspins. Thus the adjusted Hamiltonian $\hat{H}'$ (cf. Eq. (10.15)) used to calculated the expected value of the exponential of the work done also comprises two-site nearest-neighbour terms. Therefore the TEBD algorithm can be used without alteration where the system is composed of $M/2$ sites and an on-site dimension $d = 4$.

## 10.5 Numerical results

### 10.5.1 Jarzynski processes

We begin by considering the work done over the entire process, i.e., $\tau_i = 0$, $\tau_f = t_f$, and setting the probability distribution for the initial state to the Gibbs distribution corresponding to the initial value of the parameter $\lambda$, i.e., $|\, P(0)\rangle = |\, P(\lambda_i)\rangle$ (we use the exact $\chi = d$ MPS representation of the Gibbs distribution presented in Sec. 7.1). Jarzynski's equality applies, which allows us to calculate $\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]$ exactly and therefore study the errors of the TEBD calculations.

For explicitness we set $\beta J = 1$ and vary $B = \lambda$ linearly between values $\beta\lambda_i = 0$

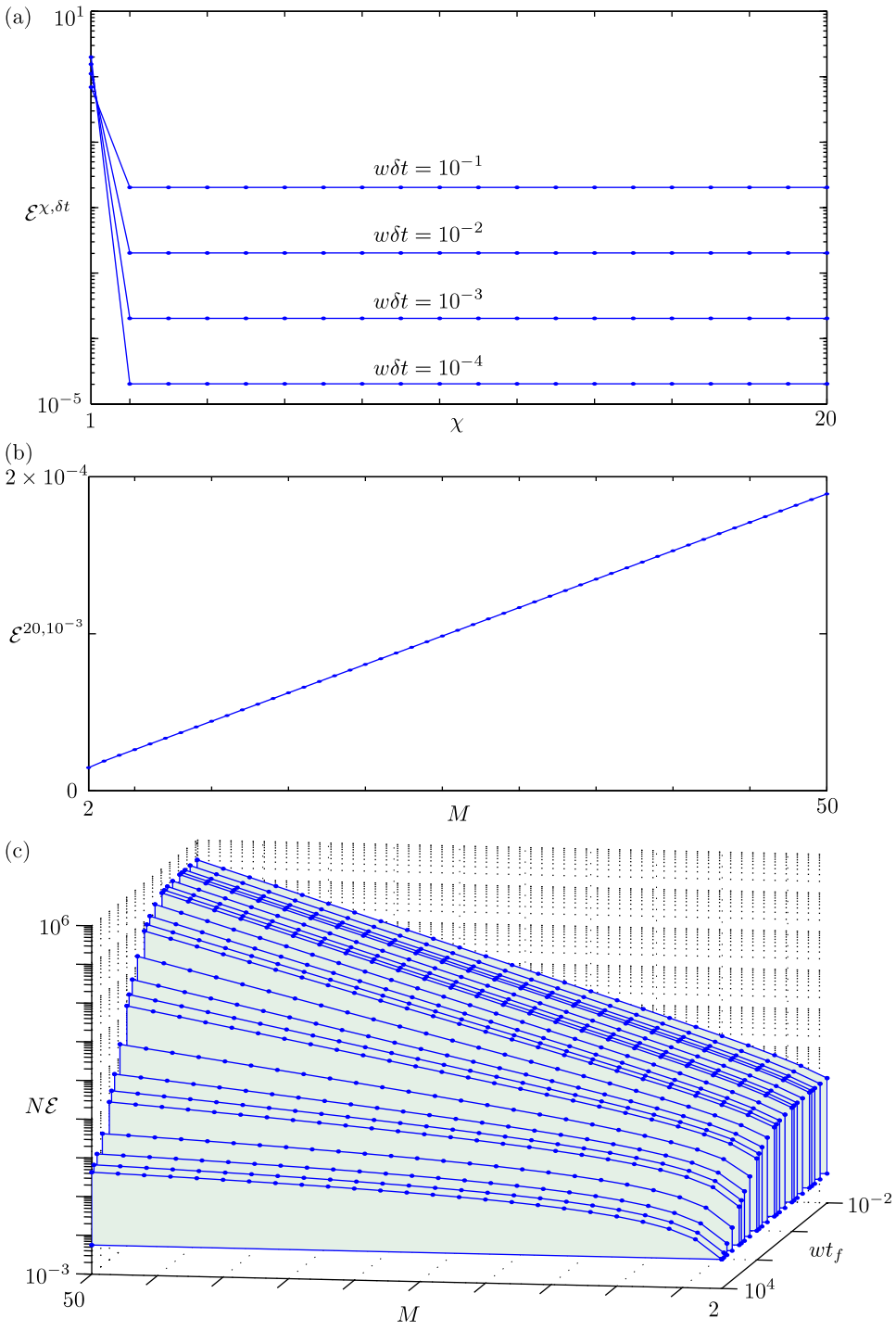


**Figure 10.1**. Realising Jarzynski's equality. (a) The fractional error $\mathcal{E}^{\chi,\delta t}$ as a function of $\chi$ for several $\delta t$. This corresponds to a linear ramp $\beta\lambda(t) = t/t_f$ over time $wt_f = 10$, correlation strength $\beta J = 1$ and system size $M = 40$. (b) The fractional error $\mathcal{E}^{20,10^{-3}}$ as a function of $M$ for a fixed time $wt_f = 100$. (c) The variance $\mathrm{Var}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right] / \mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^2$ as a function of $wt_f$ and $M$. This is equal to $N\mathcal{E}$, where $N$ is the number of sampled paths needed to ensure an average fractional error $\mathcal{E}$ in a basic DMC simulation. To calculate the variance we used $w\delta t = 10^{-2}$ and $\chi = 20$.

and $\beta\lambda_f = 1$ at times $t_i$ and $t_f$, respectively, i.e., $\beta\lambda(t) = t/t_f$. We calculate $\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^{\chi,\delta t}$ the estimate of $\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]$ using TEBD with an MPS dimension $\chi$ and timestep $\delta t$. The fractional error

$$\mathcal{E}^{\chi,\delta t} = \frac{\left|\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^{\chi,\delta t} - \mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]\right|}{\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]}, \tag{10.18}$$

of the TEBD calculations converges very quickly with increasing $\chi$, as shown in Fig. 10.1(a) for a medium simulation time $wt_f = 10$. Above $\chi > 1$ the error is purely due to the finite time-step $\delta t$ and scales as $w\delta t$. The fractional error increases only linearly in $M$, as shown in Fig. 10.1(b). This is despite the variance over expected value squared $\mathrm{Var}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right] / \mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^2$, shown in Fig. 10.1(c), increasing exponentially with $M$. We thus conclude that in this instance our method provides an exponential speed-up in $M$ with respect to naive sampling, such as would be performed by basic DMC. To arrive at this conclusion we have used that $N = \mathrm{Var}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right] / \mathcal{E}\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^2$ paths must be sampled from $\mathcal{P}$ in order to obtain an average fractional error $\mathcal{E}$ (cf. Sec. 8.1.1). As expected, we see that the variance may be reduced by adjusting the duration over which the parameter $\lambda$ is varied but this comes at the cost of having to sample long paths. It also does not avoid the exponential scaling.

We have drawn similar conclusions for other variations of $\lambda$ (though we do not show all of these results in this thesis), such as an exponential dependence

$$\beta\lambda(t) = \lambda_i \exp\left[\ln\left(\frac{\lambda_f}{\lambda_i}\right)\frac{t}{\tau_f}\right], \tag{10.19}$$

or a tanh variation

$$\beta\lambda(t) = \frac{\lambda_f + \lambda_i}{2} + \left(\frac{\lambda_f - \lambda_i}{2}\right)\frac{\tanh\left[\frac{1}{2}(t - \frac{1}{2}\tau_f)\right]}{\tanh\left[\frac{1}{4}\tau_f\right]}. \tag{10.20}$$

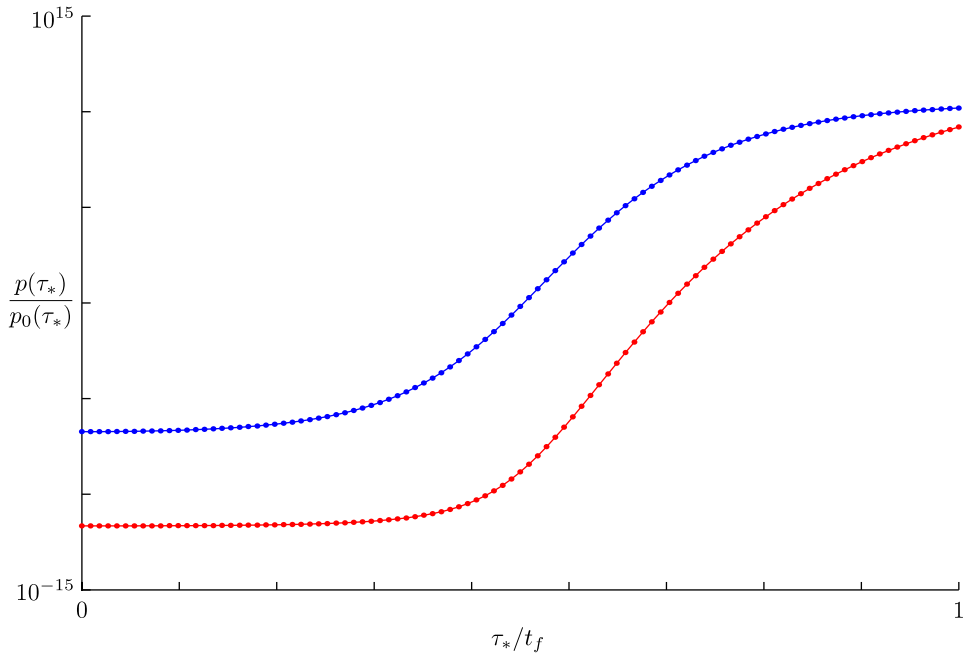


**Figure 10.2**. Expected stock price in a changing market. The expected price $p(\tau_*)$ at time $\tau_*$ for a system undergoing a tanh ramp (see Eq. (10.20)) between $\beta\lambda_i = -1$ and $\beta\lambda_f = 1$ for $M = 40$, $wt_f = 10$, $w\delta t = 10^{-3}$ and $\chi = 20$. The lower (upper) red (blue) line corresponds $\beta J = 1$ ($\beta J = 0$).

### 10.5.2 High variance stock price in a changing market

Next we go beyond what can be verified using the Jarzynski equality. We consider a process beginning in the Gibbs distribution $|\,P(0)\rangle = |\,P(\lambda_i)\rangle$. The parameter $\lambda$ is initially, for $t < \tau_*$, varied according to the tanh function in Eq. (10.20), with $\beta\lambda_i = -1$ and $\beta\lambda_f = 1$. However at time $\tau_*$ this tanh variation is interrupted by a sudden change in $\lambda$ by an amount $\beta\Delta\lambda = \ln 2$. Rather than being concerned with the work done over the whole period we are concerned only with the instantaneous work done at time $\tau_*$. As discussed in Sec. 10.1.3, this corresponds to calculating the expected stock price $p(\tau_*)/p_0(\tau_*)$ at time $\tau_*$ in a market place with a group of $M$ reactionist traders interacting with other fundamentalist traders. The tanh ramp represents an increasing confidence coming from an external bias.

Varying $\tau_*$ between 0 and $t_f$, the calculations shown in Fig. 10.2 demonstrate that

correlations between reactionist traders suppress the recovery of the stock price as the external bias becomes increasingly positive (within our simple model). From the figure we see that if there are no correlations ($\beta J = 0$) then the expected price takes small values when the bias is negative and large values when the bias is positive. Positive correlations ($\beta J > 0$) should exaggerate the small and large values, i.e., the expected price should take even smaller values when the bias is negative and even larger values when the bias is positive. However this is not quite observed in Fig. 10.2 because of a hysteresis effect. The rise in expected price lags behind the change in bias for both $\beta J = 0$ and $\beta J = 1$. Hysteresis for $\beta J = 1$ is such a strong effect that the price does not rise above that for $\beta J = 0$ before the bias has finished varying, demonstrating the importance of considering the added lag induced by correlations in a time-dependent market.

Similarly to before, we have checked (plots not shown in this thesis) that our expected value $\mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^{\chi,\delta t} = p(\tau_*)/p_0(\tau_*)$ converges quickly in $\chi$ and $\delta t$. Further our calculations of the ratio of the variance to the expected value squared $\mathrm{Var}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right] / \mathrm{E}\left[\mathrm{e}^{-\beta W}, \mathcal{P}\right]^2$ once again confirm that $N\mathcal{E}$ grows exponentially in $M$.

## 10.6 Discussion

We have calculated expected values of high variance observables in non-equilibrium stochastic processes using the TEBD method. Specifically we have calculated the exponential of the work done during the Glauber dynamics of an Ising model in a varying magnetic field and analysed the recovery of price in a simple stock market model. These observables were chosen as their variances scale exponentially with system size, thus providing an excellent test for TEBD with high variances. To calculate the expected value of the exponential of the work done over a finite period of time, we adapted the TEBD method to calculate path-dependent observables.

We found that the fractional errors achieved using TEBD showed only a small (linear) dependence on system size, suggesting that high variance expected value calculations are tractable using TEBD. In contrast we showed that an exponentially increasing sample size would be needed to achieve the same fractional errors using basic DMC. Thus, in this sense, TEBD provides an exponential speed-up to basic DMC for the cases considered.

Of course any serious attempt at performing the same calculations using DMC would exploit variance reduction techniques, for example, sequential importance sampling and go with the winners, discussed in Sec. 8.2. Perhaps a good choice for sequential importance sampling would be to sample from Glauber dynamics with a stationary Gibbs distribution $P_{\mathbf{i}} \propto \exp[-\beta(\lambda_f - \lambda_i)M_{\mathbf{i}}]$ for some choice of $\beta$ and to begin the process in this distribution. This would ensure the frequent sampling of paths that spend a lot of time in configurations with highly negative (for positive $\lambda_f - \lambda_i$) magnetisation $M_{\mathbf{i}}$, which we might expect to contribute the most to the expected value. Also, since the growth of the observable could be tracked in time, a go with the winners strategy could be used to further focus computational resources on the most important paths. However a fair bit of trial and error would be needed with both of these methods and without doing this there is no way of knowing whether an exponentially scaling variance reduction is achievable, as would be required to match our TEBD results. This comparison also emphasises the convenience of our TEBD calculations. Setting up our calculations required no judicious system-dependent choices and we did not need to know anything about what to expect.

To close, we discuss the generality of our approach. While we limited ourselves to one dimension for ease of comparison with analytical results, several aspects of our approach are applicable in higher dimensions and more complicated geometries. Our derivation of an adjusted master equation to simulate path-dependent observables

would be applicable with any geometry and so could be used in conjunction with the generalisation of TEBD and MPS methods to tree-like, two-dimensional or three-dimensional geometries [139, 144, 141, 142]. The complication is that, for non-tree-like geometries, extracting information exactly from the compressed distribution, even calculating its sum, is often inefficient and another approximation must be made [141, 142]. Secondly, part of the success of our method lay in the fact that Gibbs distributions are represented by MPS with a small dimension (see Sec. 7.1). This extends naturally to tensor network representations of Gibbs distributions in higher dimensions and follows from the transfer matrix method approach. Future work will include the study of the performance of tensor network methods for two-dimensional stochastic processes. In fact, classical simulations will provide a useful test for the two-dimensional tensor network software currently in development by our research group[2] since the dimension $\chi$ needed tends to be much smaller than for quantum simulations.

[2]See the Tensor Network Theory Library project at `http://ccpforge.cse.rl.ac.uk/gf/project/tntlibrary/`.

# Chapter 11

# Conclusion

We conclude this thesis by discussing the exciting prospects for the development of the ideas contained within it. However, before we do this, let us first summarise some significant and substantial contributions we have already made.

**Chapter 4.** We found that the dynamics of an impurity moving in a weakly-interacting Bose gas at low temperatures is well-described by taking into account only the effects of self-trapping, the inhomogeneity of the Bose gas and the dissipation of the energy of the impurity into the phononic excitations of the Bose gas. However for stronger inter-boson interactions the effects of the Bose gas on the impurity effective mass must be taken into account, at least for a quasi one-dimensional system.

**Chapter 5.** Next we showed that placing both the impurity and the Bose gas in separate optical lattices, and tilting that of the impurity, realises a system that is closely analogous to theoretical models of driven electronic motion in a solid. We showed that this cold atom setup exhibits clear signatures of the Van Hove singularity in the phonon spectrum that have yet to be clearly observed in solid state systems. Inspired by this result, we showed more generally that the motion of the impurity clearly reveals properties of the excitation spectrum of its environment.

We studied this for the case of strongly-interacting bosons in an optical lattice, specifically the one-dimensional Bose-Hubbard model, and identified the effects of the different strongly-correlated phases of the Bose gas on the impurity dynamics.

**Chapter 9.** We adapted the time-evolving block decimation (TEBD) algorithm for use with non-equilibrium classical stochastic processes. We examined the effectiveness of using non-negative matrix factorisation (NMF) instead of the singular value decomposition (SVD) as the fundamental building-block of the algorithm, since it in principle directly minimises an upper-bound of the relevant errors. However we found that current NMF algorithms both demand significant computational resources and struggle to find an accurate factorisation, therefore using the SVD is currently the best approach for classical as well as quantum systems. We then used the TEBD algorithm in conjunction with the SVD and showed that it accurately reproduced the known long-time properties of the totally-asymmetric exclusion process for small and medium sized systems. For large systems we found that observables quickly converged to physically-sensible values, demonstrating the accuracy of the algorithm even for transient dynamics.

**Chapter 10.** We then adapted the TEBD algorithm to compute the expected values of observables that depend on the whole path of a system, not just its configuration at one time. We used this to estimate the expected value of an observable whose variance increases exponentially with the size of the system. Nevertheless the error of our numerical method scaled linearly with system size, much better than the expected exponential scaling for basic Monte Carlo algorithms. This suggests that TEBD may provide an excellent alternative to Monte Carlo-based algorithms when the observables of interest have an extremely high variance.

Let us now suggest three future research directions that would lead to important extensions of the work in this thesis. A first research direction relates to the

dynamics of impurities immersed in both periodic potentials and weakly-interacting Bose gases. A rich system that provides a novel twist on this idea is a rotating quasi two-dimensional impurity-Bose gas mixture. For sufficiently high rotation frequencies the rotating Bose gas forms a triangular vortex lattice [281]. This therefore provides both a lattice potential for the impurity atoms and dissipative interactions with a phononic environment. The rotation also naturally leads to Peierls phases on the hopping elements associated with the tunnelling of the impurity between lattice sites [282]. A particularly exciting aspect of this system is the softness of the lattice potential, i.e., the presence of an impurity at a lattice site deepens the lattice potential at that site, leading to both self-trapping and an attractive contribution to the on-site impurity-impurity interaction. Future investigation will extend and combine the work on impurity-Bose gas mixtures in this thesis to describe the impurity-vortex lattice system.

Another prospect that arose in this thesis is that of using an impurity as a probe. While we indeed found that by observing an impurity we could glean information about a surrounding Bose gas, the impurity Hamiltonian itself was quite complicated and this made it difficult to use the impurity as a probe. In the future we will analyse the use of a much simpler impurity system, e.g., a qubit formed by two atomic internal states [18, 21], as a probe. We will consider this qubit probe quite generally, but two specific applications spring to mind. First, we will ask to what extent the probe could distinguish between the mean density of a Bose gas and its underlying corpuscular make-up. This would necessarily involve a highly-localised impurity as might realistically be achieved using an atom in a tight harmonic trap. Second, we ask whether the probe could distinguish between a coherent superposition and an incoherent mixture of different particle numbers. This would reveal whether spontaneous symmetry breaking in Bose gases is a real phenomenon or just a useful theoretical tool [283].

Last, as mentioned in Sec. 10.6, the TEBD algorithm is only efficient for one-dimensional systems. The next step is thus to consider the related algorithms that are effective in higher dimensions and adapt these also for use with stochastic processes. If this is successful the impact of these approaches on the stochastic process community would be significant.

# Bibliography

[1] T. H. Johnson, S. R. Clark, and D. Jaksch, Phys. Rev. E **82**, 036702 (2010).

[2] M. Bruderer, T. H. Johnson, S. R. Clark, D. Jaksch, A. Posazhennikova, and W. Belzig, Phys. Rev. A **82**, 043617 (2010).

[3] T. H. Johnson, S. R. Clark, M. Bruderer, and D. Jaksch, Phys. Rev. A **84**, 023617 (2011).

[4] T. H. Johnson, M. Bruderer, Y. Cai, S. R. Clark, W. Bao, and D. Jaksch, Europhys. Lett. **98**, 26001 (2012).

[5] T. H. Johnson, T. J. Elliot, S. R. Clark, and D. Jaksch, *Calculating expected values of high-variance observables in dynamical stochastic processes using matrix products*, in preparation (2013).

[6] T. H. Johnson, J. D. Biamonte, S. R. Clark, and D. Jaksch, Sci. Rep. **3**, 1235 (2013).

[7] P. Haikka, T. H. Johnson, and S. Maniscalco, Phys. Rev. A **87**, 010103(R) (2013).

[8] R. Walters, G. Cotugno, T. H. Johnson, S. R. Clark, and D. Jaksch, Phys. Rev. A. **87**, 043613 (2013).

[9] M. Faccin, T. H. Johnson, J. D. Biamonte, S. Kais, and P. Migdał, Phys. Rev. X **3**, 041007 (2013).

[10] M. Faccin, P. Migdał, T. H. Johnson, J. D. Biamonte, and V. Bergholm, preprint arXiv:1310.6638 (2013).

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

[12] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).

[13] S. Palzer, C. Zipkes, C. Sias, and M. Köhl, Phys. Rev. Lett. **103**, 150601 (2009).

[14] J. Catani, G. Lamporesi, D. Naik, M. Gring, M. Inguscio, F. Minardi, A. Kantian, and T. Giamarchi, Phys. Rev. A **85**, 023623 (2012).

[15] S. Will, Th. Best, S. Braun, U. Schneider, and I. Bloch, Phys. Rev. Lett. **106**, 115305 (2011).

[16] C. Weber, S. John, N. Spethmann, D. Meschede, and A. Widera, Phys. Rev. A **82**, 042722 (2010).

[17] N. Spethmann, F. Kindermann, S. John, C. Weber, D. Meschede, and A. Widera, Phys. Rev. Lett. **109**, 235301 (2012).

[18] M. Bruderer and D. Jaksch, New J. Phys. **8**, 87 (2006).

[19] A. Recati, P. O. Fedichev, W. Zwerger, J. von Delft, and P. Zoller, Phys. Rev. Lett. **94**, 040404 (2005).

[20] P. Haikka, S. McEndoo, G. De Chiara, G. M. Palma, and S. Maniscalco, Phys. Rev. A **84**, 031602 (2011).

[21] J. Goold, T. Fogarty, N. Lo Gullo, M. Paternostro, and Th. Busch, Phys. Rev. A **84**, 063632 (2011).

[22] M. Knap, A. Shashi, Y. Nishida, A. Imambekov, D. A. Abanin, and E. Demler, Phys. Rev. X **2**, 041020 (2012).

[23] A. V. Ponomarev, J. Madroñero, A. R. Kolovsky, and A. Buchleitner, Phys. Rev. Lett. **96**, 050404 (2006).

[24] M. Bruderer, A. Klein, S. R. Clark, and D. Jaksch, Phys. Rev. A **76**, 011605(R) (2007).

[25] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, Phys. Rev. Lett. **81**, 3108 (1998).

[26] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Phys. Rev. B **40**, 546 (1989).

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

[28] S. R. White, Phys. Rev. B **48**, 10345 (1993).

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

[30] Y. Hieida, J. Phys. Soc. Jpn. **67**, 369 (1998).

[31] E. Carlon, M. Henkel, and U. Schollwöck, Eur. Phys. J. B **12**, 99 (1999).

[32] E. Carlon, M. Henkel, and U. Schollwöck, Phys. Rev. E **63**, 036101 (2001).

[33] A. Kemper, A. Schadschneider, and J. Zittartz, J. Phys. A **34**, L279 (2001).

[34] T. Enss and U. Schollwöck, J. Phys. A **34**, 7769 (2001).

[35] A. Kemper, A. Gendiar, T. Nishino, A. Schadschneider, and J. Zittartz, J. Phys. A **36**, 29 (2003).

[36] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes*, p. 6 (Methuen, London, 1965).

[37] W. M. Young and E. W. Elcock, Proc. Phys. Soc. **89**, 735 (1966).

[38] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comp. Phys. **17**, 10 (1975).

[39] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).

[40] A. P. J. Jansen, preprint arXiv:cond-mat/0303028 (2003).

[41] C. Dellago, P. G. Bolhus, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).

[42] O. Mazonka, C. Jarzynski, and J. Blocki, Nuc. Phys. A **641**, 335 (1998).

[43] H. K. Lee and Y. Okabe, Phys. Rev. E **75**, 015102(R) (2005).

[44] A. Kundu, S. Sabhapandit, and A. Dhar, Phys. Rev. E **83**, 031119 (2011).

[45] D. Aldous and U. Vazirani, *Go with the winners algorithms* in *Proc. 35th IEEE Sympos. on Found. of Comp. Sci.*, p. 492 (IEEE, 1994).

[46] P. Grassberger, Comp. Phys. Comm. **147**, 64 (2002).

[47] T. Dean and P. Dupuis, Stochastic Process. Appl. **119**, 562 (2009).

[48] L. Pitaevskii and S. Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford, 2003).

[49] R. M. Ziff, G. E. Uhlenbeck, and M. Kac, Phys. Rep. **32**, 169 (1977).

[50] M. Gajda and K. Rzazewski, Phys. Rev. Lett. **78**, 2686 (1997).

[51] R. Shankar, *Principles of quantum mechanics*, 2nd Ed. (Plenum Press, New York, 1994).

[52] H. D. Politzer, Phys. Rev. A **54**, 5048 (1996).

[53] N. L. Balazs and T. Bergeman, Phys. Rev. A **58**, 2359 (1998).

[54] K. Huang, *Statistical mechanics*, 2nd Ed. (John Wiley & Sons, New York, 1987).

[55] K. Huang and C. N. Yang, Phys. Rev. **105**, 767 (1957).

[56] T. D. Lee and C. N. Yang, Phys. Rev. **105**, 1119 (1957).

[57] T. D. Lee, K. Huang, and C. N. Yang, Phys. Rev. **106**, 1135 (1957).

[58] I. Bloch, J. Dalibard, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).

[59] M. A. Cazalilla, R. Citro, T. Giamarchi, E. Orignac, and M. Rigol, Rev. Mod. Phys. **83**, 1405 (2011).

[60] V. L. Ginzburg and L. P. Pitaevskii, Sov. Phys. JETP **34**, 858 (1958).

[61] L. P. Pitaevskii, Sov. Phys. JETP **13**, 451 (1961).

[62] E. P. Gross, Nuovo Cimento **20**, 454 (1961).

[63] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).

[64] Y. Castin, *Bose-Einstein condensates in atomic gases: simple theoretical results* in *Procs. of the Les Houches Summer School, Session LXXII*, Ed. R. Kaiser, p. 1 (EDP Sciences & Springer-Verlag, Berlin, 2001).

[65] E. H. Lieb, R. Seiringer, and J. Yngvason, Phys. Rev. A **61**, 043602 (2000).

[66] A. J. Leggett, Rev. Mod. Phys. **73**, 307 (2001).

[67] C. W. Gardiner, Phys. Rev. A **56**, 1414 (1997).

[68] Y. Castin and R. Dum, Phys. Rev. A **57**, 3008 (1998).

[69] P. R. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966).

[70] N. N. Bogoliubov, J. Phys. USSR **11**, 23 (1947).

[71] P. Öhberg, E. L. Surkov, I. Tittonen, S. Stenholm, M. Wilkens, and G. V. Shlyapnikov, Phys. Rev. A **56**, R3346 (1997).

[72] M. Girardeau and R. Arnowitt, Phys. Rev. **113**, 775 (1959).

[73] A. L. Fetter, Ann. Phys. **70**, 67 (1972).

[74] S. A. Gardiner and T. P. Billam, preprint arXiv:1207.6382 (2012).

[75] F. Schwabl, *Advanced quantum mechanics*, 2nd Ed. (Springer-Verlag, Berlin, 2004).

[76] D. S. Petrov, D. M. Gangardt, and G. V. Shlyapnikov, J. Phys. IV **116**, 5 (2004).

[77] D. S. Petrov, M. Holtzmann, and G. V. Shlyapnikov, Phys. Rev. Lett. **84**, 2551 (2000).

[78] M. Olshanii, Phys. Rev. Lett. **81**, 938 (1998).

[79] V. N. Popov, *Functional integrals in quantum field theory and statistical physics* (D. Reidel, Dordrecht, 1983).

[80] Y. Castin, J. Phys. IV **116**, 89 (2004).

[81] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).

[82] P. C. Hohenberg, Phys. Rev. Lett. **158**, 383 (1967).

[83] W. Kane and L. Kadanoff, Phys. Rev. **155**, 80 (1967).

[84] L. Reatto and G. V. Chester, Phys. Rev. **155**, 88 (1967).

[85] C. N. Yang and C. P. Yang, Rev. Mod. Phys. **34**, 694 (1962).

[86] D. S. Petrov, G. V. Shlyapnikov, and J. T. M. Walraven, Phys. Rev. Lett. **85**, 3745 (2000).

[87] M. Schwartz, Phys. Rev. B **15**, 1399 (1977).

[88] F. D. M. Haldane, Phys. Rev. Lett. **47**, 1840 (1981).

[89] V. Bagnato and D. Kleppner, Phys. Rev. A **44**, 7439 (1991).

[90] W. Ketterle and N. J. van Druten, Phys. Rev. A **54**, 656 (1996).

[91] S. Dettmer, D. Hellweg, P. Ryytty, J. J. Arlt, W. Ertmer, K. Sengstock, D. S. Petrov, G. V. Shlyapnikov, H. Kreutzmann, L. Santos, and M. Lewenstein, Phys. Rev. Lett. **87**, 160406 (2001).

[92] S. Richard, F. Gerbier, J. H. Thywissen, M. Hugbart, P. Bouyer, and A. Aspect, Phys. Rev. Lett. **91**, 010405 (2003).

[93] J.-B. Trebbia, J. Esteve, C. I. Westbrook, and I. Bouchoule, Phys. Rev. Lett. **97**, 250403 (2006).

[94] P. Krüger, S. Hofferberth, I. E. Mazets, I. Lesanovsky, and J. Schmiedmayer, Phys. Rev. Lett. **105**, 265302 (2010).

[95] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen, and U. Sen, Adv. Phys. **56**, 2 (2007); M. Lewenstein, A. Sanpera, and V. Ahufinger, *Ultracold Atoms in Optical Lattices: Simulating quantum many-body systems* (Oxford University Press, Oxford, 2012).

[96] N. W. Ashcroft and D. Mermin, *Solid State Physics* (Holt, Rinehardt and Winston, New York, 1976).

[97] G. Wannier, Phys. Rev. **52**, 191 (1937).

[98] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[99] T. Stöferle, H. Moritz, C. Schori, M. Köhl, and T. Esslinger, Phys. Rev. Lett. **92**, 130403 (2004).

[100] H. Feshbach, Ann. Phys. **5**, 337 (1958).

[101] P. O. Fedichev, M. W. Reynold, and G. V. Shlyapnikov, Phys. Rev. Lett. **77**, 2921 (1996).

[102] J. L. Bohn and P. S. Julienne, Phys. Rev. A **60**, 414 (1999).

[103] E. Timmermans, K. Furuya, P. W. Milonni, and A. K. Kerman, Phys. Rev. A **285**, 228 (2001).

[104] K. D. Nelson, X. Li, and D. S. Weiss, Nature Physics **3**, 556 (2007); T. Gericke, P. Würtz, D. Reitz, T. Langen, and H. Ott, Nature Physics **4**, 949 (2008); W. S. Bakr, J. I. Gillen, A. Peng, S. Fölling, and M. Greiner, Nature **462**, 74 (2009); M. Karski, L. Förster, J. M. Choi, W. Alt, A. Widera, and D. Meschede, Phys. Rev. Lett. **102**, 053001 (2009); J. F. Sherson, C. Weitenberg, M. Endres, M. Cheneau, I. Bloch, and S. Kuhr, Nature **467**, 68 (2010); A. Itah, H. Veksler, O. Lahav, A. Blumkin, C. Moreno, C. Gordon, and J. Steinhauer, Phys. Rev. Lett. **104**, 113001 (2010).

[105] P. W. Anderson, P. A. Lee, M. Randeria, T. M. Rice, N. Trivedi, and F. C. Zhang, J. Phys.: Condens. Matter **16**, R755 (2004).

[106] D. van Oosten, P. van der Straten, and H. T. C. Stoof, Phys. Rev. A **63**, 053601 (2001).

[107] K. Burnett, M. Edwards, C. W. Clark, and M. Shotter, J. Phys. B: At. Mol. Phys. **35**, 1671 (2002).

[108] A. M. Rey, K. Burnett, R. Roth, M. Edwards, C. J. Williams, and C. W. Clark, J. Phys B: At. Mol. Phys. **36**, 825 (2003).

[109] M. A. Cazalilla, Phys. Rev. A **70**, 41604 (2004).

[110] T. D. Kühner, S. R. White, and H. Monien, Phys. Rev. B **61**, 12474 (2000).

[111] C. Kollath, U. Schollwöck, J. von Delft, and W. Zwerger, Phys. Rev. A **69**, 031601(R) (2004).

[112] Y. Cai, *Mathematical theory and numerical methods for Gross-Pitaevskii equations and applications* (PhD Thesis, National University of Singapore, 2011).

[113] A. Jüngel and R.-M. Weishäupl, *Blow-up in two-component nonlinear Schrödinger systems with an external deriven field*; to appear in Math. Models Meth. Appl. Sci. (2013).

[114] M. M. Cerimele, M. L. Chiofalo, F. Pistella, S. Succi, and M. P. Tosi, Phys. Rev. E **62**, 1382 (2000).

[115] M. L. Chiofalo, S. Succi, and M. P. Tosi, Phys. Rev. E **62**, 7438 (2000).

[116] W. Bao and Q. Du, SIAM J. Sci. Comput. **25**, 1674 (2004).

[117] W. Bao, Multiscale Model Sim. **2**, 210 (2004).

[118] W. Bao and Y. Cai, East Asia J. Appl. Math. **1**, 49 (2010).

[119] G. I. Marchuk, *On the theory of the splitting-up method* in *Procs. of the Second Symposium on the Numerical Solution of Partial Differential Equations*, Ed. B. Hubbard (Academic Press, London, 1971).

[120] N. N. Yanenko, *The method of fractional steps* (Springer, New York, 1971).

[121] G. Strang, SIAM J. Numer. Anal. **5**, 506 (1968).

[122] W. Bao, D. Jaksch, and P. A. Markowich, J. Comput. Phys. **187**, 318 (2003).

[123] W. Bao, Q. Du, and Y. Z. Zhang, SIAM J. Appl. Math. **66**, 758 (2006).

[124] D. Gobert, C. Kollath, U. Schollwöck, and G. Schütz, Phys. Rev. E **71**, 036102 (2005).

[125] S. R. White and A. E. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004).

[126] S. R. Clark and D. Jaksch, Phys. Rev. A **70**, 043612 (2004).

[127] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, J. Stat. Mech. P04005 (2004).

[128] A. J. Daley, S. R. Clark, D. Jaksch, and P. Zoller, Phys. Rev. A **72**, 043618 (2005).

[129] S. R. Clark and D. Jaksch, New J. Phys. **8**, 160 (2006).

[130] M. Bruderer, A. Klein, S. R. Clark, and D. Jaksch, New J. Phys. **10**, 033015 (2008).

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

[132] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd Ed. (Johns Hopkins University Press, Baltimore, 1996).

[133] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[134] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Rev. Mod. Phys. **80**, 517 (2008).

[135] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

[136] C. Eckart and G. Young, Psychometrika **1**, 211 (1936).

[137] M. Suzuki, Phys. Lett. A **146**, 319 (1990); J. Math. Phys. **32**, 400 (1991).

[138] M. Zwolak and G. Vidal (private communication, 2004).

[139] Y. Y. Shi, L. M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).

[140] I. L. Markov and Y. Y. Shi, Algorithmica **59**, 461 (2011).

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

[142] J. I. Cirac and F. Verstraete, J. Phys. A **42**, 504004 (2009).

[143] G. Vidal, Phys. Rev. Lett. **99**, 220405 (2007).

[144] G. Vidal, Phys. Rev. Lett. **98**, 070201 (2007).

[145] F. Verstraete and J. I. Cirac, Phys. Rev. Lett. **104**, 190405 (2010).

[146] Y. Kawaguchi and M. Ueda, Phys. Rep. **520**, 253 (2012).

[147] F. Zhou and G. W. Semenoff, Phys. Rev. Lett. **97**, 180411 (2006).

[148] K. Eckert, L. Zawitkowski, M. J. Leskinen, A. Sanpera, and M. Lewenstein, New J. Phys. **9**, 133 (2007).

[149] B. Damski, H.-U. Everts, A. Honecker, H. Fehrmann, L. Santos, and M. Lewenstein, Phys. Rev. Lett. **95**, 060403 (2005).

[150] B. Damski, H. Fehrmann, H.-U. Everts, M. Baranov, L. Santos, and M. Lewenstein, Phys. Rev. A **72**, 053612 (2005).

[151] A. Honecker, D. C. Cabra, H.-U. Everts, P. Pujol, and F. Stauffer, J. Phys.: Condens. Matter **19**, 145249 (2007).

[152] B. DeMarco and D. S. Jin, Science **285**, 1703 (1999).

[153] A. Klein, M. Bruderer, S. R. Clark, and D. Jaksch, New J. Phys. **9**, 411 (2007).

[154] J. Tempere, W. Casteels, M. K. Oberthaler, S. Knoop, E. Timmermans, and J. T. Devreese, Phys. Rev. B **80**, 184504 (2009).

[155] J. Catani, L. De Sarlo, G. Barontini, F. Minardi, and M. Inguscio, Phys. Rev. A **77**, 011603(R) (2008).

[156] B. Gadway, D. Pertot, R. Reimann, and D. Schneble, Phys. Rev. Lett. **105**, 045303 (2010).

[157] H. T. Ng and S. Bose, Phys. Rev. A **78**, 023610 (2008).

[158] P. Haikka, J. Goold, S. McEndoo, F. Plastina, and S. Maniscalco, Phys. Rev. A **85**, 060101(R) (2012).

[159] M. Punk and S. Sachdev, Phys. Rev. A **87**, 033618 (2013).

[160] R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, and V. Vedral, Phys. Rev. Lett. **110**, 230601 (2013).

[161] L. Mazzola, G. De Chiara, and M. Paternostro, Phys. Rev. Lett. **110**, 230602 (2013).

[162] C. Sabín, A. White, L. Hackermuller, and I. Fuentes, preprint arXiv:1303.6208 (2013).

[163] D. K. K. Lee and J. M. F. Gunn, Phys. Rev. B **46**, 301 (1992).

[164] R. M. Kalas and D. Blume, Phys. Rev. A **73**, 043608 (2006).

[165] M. Bruderer, W. Bao, and D. Jaksch, Europhys. Lett. **82**, 30004 (2008).

[166] D.-S. Lühmann, K. Bongs, K. Sengstock, and D. Pfannkuche, Phys. Rev. Lett. **101**, 050402 (2008).

[167] *The Bible: Authorized King James Version with Apocrypha*, Sec. Exodus 14 (Oxford University Press, Oxford, 1997).

[168] J. Bonart and L. F. Cugliandolo, Phys. Rev. A **86**, 023636 (2012).

[169] J. Bonart and L. F. Cugliandolo, Europhys. Lett. **101**, 16003 (2013).

[170] S. Peotta, D. Rossini, M. Polini, F. Minardi, and R. Fazio, Phys. Rev. Lett. **110**, 015302 (2013).

[171] A. R. Kolovsky, Phys. Rev. A **77**, 063604 (2008).

[172] A. Zenesini, C. Sias, H. Lignier, Y. Singh, D. Ciampini, O. Morsch, R. Mannella, E. Arimondo, A. Tomadin, and S. Wimberger, New J. Phys. **10**, 053038 (2008).

[173] M. Ben Dahan, E. Peik, J. Reichel, Y. Castin, and C. Salomon, Phys. Rev. Lett. **76**, 4508 (1996).

[174] G. Roati, E. de Mirandes, F. Ferlaino, H. Ott, G. Modugno, and M. Inguscio, Phys. Rev. Lett. **92**, 230402 (2004).

[175] C. Sias, H. Lignier, Y. P. Singh, A. Zenesini, D. Ciampini, O. Morsch, and E. Arimondo, Phys. Rev. Lett. **100**, 040404 (2008).

[176] H. Ott, E. de Mirandes, F. Ferlaino, G. Roati, G. Modugno, and M. Inguscio, Phys. Rev. Lett. **92**, 160601 (2004).

[177] T. Best, S. Will, U. Schneider, L. Hackermüller, D. van Oosten, I. Bloch, and D.-S. Lühmann, Phys. Rev. Lett. **102**, 030408 (2009).

[178] P. T. Ernst, S. Götze, J. S. Krauser, K. Pyka, D.-S. Lühmann, D. Pfannkuche, and K. Sengstock, Nature Physics **6**, 56 (2009).

[179] K. W. Madison, C. F. Bharucha, P. R. Morrow, S. R. Wilkinson, Q. Niu, B. Sundaram, and M. G. Raizen, Appl. Phys. B **65**, 693 (1997).

[180] L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 61 (1970).

[181] G. H. Döhler, R. Tsu, and L. Esaki, Solid State Commun. **17**, 317 (1975).

[182] K. Leo, *High-Field Transport in Semiconductor Superlattices* (Springer, Berlin, 2003).

[183] V. V. Bryxin and Y. A. Firsov, Solid State Commun. **10**, 471 (1972).

[184] V. V. Bryxin and P. Kleinert, J. Phys.: Condens. Matter **9**, 7403 (1997).

[185] S. Feng, C. H. Grein, and M. E. Flatté, Phys. Rev. B **68**, 085307 (2003).

[186] L. Van Hove, Phys. Rev. **89**, 1189 (1953).

[187] D. Emin and C. F. Hart, Phys. Rev. B **36**, 2530 (1987).

[188] M. Glück, A. R. Kolovsky, and H. J. Korsch, Phys. Rep. **366**, 103 (2002).

[189] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd Ed. (Cambridge Universtiy Press, Cambridge, 1995).

[190] F. Rossi, *Bloch oscillations and Wannier-Stark localization in semiconductor superlattices* in *Theory of Transport Properties of Semiconductor Nanostructures*, Ed. E. Schöll (Chapman and Hall, London, 1998).

[191] A. Sibille, J. F. Palmier, H. Wang, and F. Mollot, Phys. Rev. Lett. **64**, 52 (1990).

[192] A. V. Ponomarev, *Dynamics of cold Fermi atoms in one-dimensional optical lattices* (PhD Thesis, University of Freiburg, 2008).

[193] A. Wacker and A.-P. Jauho, Phys. Rev. Lett. **80**, 369 (1998).

[194] A. R. Kolovsky, J. Stat. Mech. P02018 (2009).

[195] L. D. Landau, J. Phys. (Moscow) **5**, 71 (1941).

[196] A. J. Daley, P. O. Fedichev, and P. Zoller, Phys. Rev. A **69**, 022306 (2004).

[197] G. D. Mahan, *Many-Particle Physics*, 3rd ed. (Kluwer Academic, New York, 2000).

[198] M. A. Cazalilla, Phys. Rev. A **67**, 053606 (2003).

[199] J. Stenger, S. Inouye, A. P. Chikkatur, D. M. Stamper-Kurn, D. E. Pritchard, and W. Ketterle, Phys. Rev. Lett. **82**, 4569 (1999).

[200] E. Haller, R. Hart, M. J. Mark, J. G. Danzl, L. Reichsöllner, M. Gustavsson, M. Dalmonte, G. Pupillo, and H.-C. Nägerl, Nature **436**, 09259 (2010).

[201] C. Kollath, A. Iucci, T. Giamarchi, W. Hofstetter, and U. Schollwöck, Phys. Rev. Lett. **97**, 050402 (2006).

[202] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, Nature **415**, 39 (2002).

[203] B. Gadway, D. Pertot, J. Reeves, and D. Scheble, Nature Physics **8**, 544 (2012).

[204] G. E. Astrakharchik and L. P. Pitaevskii, Phys. Rev. A **70**, 013608 (2004).

[205] A. Privitera and W. Hofstetter, Phys. Rev. A **82**, 063614 (2010).

[206] L. Rutherford, J. Goold, Th. Busch, and J. F. McCann, Phys. Rev. A **83**, 055601 (2011).

[207] C. J. M. Mathy, M. B. Zvonarev, and E. Demler, Nature Physics **8**, 881 (2012).

[208] E. Timmermans and R. Côté, Phys. Rev. Lett. **80**, 3419 (1998).

[209] D. S. Rokhsar and B. G. Kotliar, Phys. Rev. B **44**, 10328 (1991).

[210] W. Krauth, M. Caffarel, and J. P. Bouchaud, Phys. Rev. B **45**, 3137 (1992).

[211] K. Sheshadri, H. R. Krishnamurthy, R. Pandit, and T. V. Ramakrishnan, Europhys. Lett. **22**, 257 (1993).

[212] A. Marte, T. Volz, J. Schuster, S. Dürr, G. Rempe, E. G. M. van Kempen, and B. J. Verhaar, Phys. Rev. Lett. **89**, 283202 (2002).

[213] H. Wang, A. N. Nikolov, J. R. Ensher, P. L. Gould, E. E. Eyler, W. C. Stwalley, J. P. Burke, J. L. Bohn, C. H. Greene, E. Tiesinga, C. J. Williams, and P. S. Julienne, Phys. Rev. A **62**, 052704 (2000).

[214] F. Ferlaino, C. D'Errico, G. Roati, M. Zaccanti, M. Inguscio, and G. Modugno, Phys. Rev. A **73**, 040702(R) (2006).

[215] F. Ferlaino, C. D'Errico, G. Roati, M. Zaccanti, M. Inguscio, G. Modugno, and A. Simoni, Phys. Rev. A **73**, 039903(E) (2006).

[216] T. P. Meyrath, F. Schreck, J. L. Hanssen, C.-S. Chuu, and M. G. Raizen, Phys. Rev. A **71**, 041604(R) (2005).

[217] J. Catani, L. D. Sarlo, G. Barontini, F. Minardi, and M. Inguscio, Phys. Rev. A **77**, 011603 (2008).

[218] A. Albus, F. Illuminati, and J. Eisert, Phys. Rev. A **68**, 023606 (2003).

[219] M. Krämer, C. Menotti, L. Pitaevskii, and S. Stringari, Eur. Phys. J. D **27**, 247 (2003).

[220] S. Bergkvist, P. Henelius, and A. Rosengren, Phys. Rev. A **70**, 053601 (2004).

[221] V. G. Rousseau, D. P. Arovas, M. Rigol, F. Hébert, G. G. Batrouni, and R. T. Scalettar, Phys. Rev. B **73**, 174516 (2006).

[222] L. Pollet, S. M. A. Rombouts, and P. J. H. Denteneer, Phys. Rev. Lett. **93**, 210401 (2004).

[223] B. Paredes, A. Widera, V. Murg, O. Mandel, S. Fölling, I. Cirac, G. V. Shlyapnikov, T. W. Hänsch, and I. Bloch, Nature **429**, 277 (2004).

[224] G. G. Batrouni, V. Rousseau, R. T. Scalettar, M. Rigol, A. Muramatsu, P. J. H. Denteneer, and M. Troyer, Phys. Rev. Lett. **89**, 117203 (2002).

[225] S. Wessel, F. Alet, M. Troyer, and G. G. Batrouni, Phys. Rev. A **70**, 053615 (2004).

[226] G. G. Batrouni, R. T. Scalettar, and G. T. Zimanyi, Phys. Rev. Lett. **65**, 1765 (1990).

[227] B. DeMarco, C. Lannert, S. Vishveshwara, and T.-C. Wei, Phys. Rev. A **71**, 063601 (2005).

[228] A. M. Rey, *Ultracold bosonic atoms in optical lattices* (PhD Thesis, University of Maryland, 2003).

[229] J. Krug, Phys. Rev. Lett. **67**, 1882 (1991).

[230] G. Grinstein, D.-H. Lee, and S. Sachdev, Phys. Rev. Lett. **64**, 1927 (1990).

[231] A. Schadschneider, Physica A **313**, 153 (2002).

[232] D. Helbing, Rev. Mod. Phys. **73**, 1067 (2001).

[233] D. Chowdhury, L. Santen, and A. Schadschneider, Phys. Rep. **329**, 199 (2000).

[234] D. Chowdhury, A. Schadschneider, and K. Nishinari, Phys. Life Rev. **2**, 318 (2005).

[235] Y. Aghababaie, G. I. Menon, and M. Plischke, Phys. Rev. E **59**, 2578 (1999).

[236] T. Chou and D. Lohse, Phys. Rev. Lett. **82**, 3552 (1999).

[237] A. Parmeggiani, T. Franosch, and E. Frey, Phys. Rev. Lett. **90**, 086601 (2003).

[238] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, J. Phys. A **26**, 1493 (1993).

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

[240] G. M. Schütz, *Exactly solvable models for many-body systems far from equilibrium* in *Phase Transitions and Critical Phenomena*, Vol. 19, Eds. C. Domb and J. L. Lebowitz (Academic Press, London, 2000).

[241] L. D. Landau and E. M. Lifshitz, *Statistical Physics Part 1*, 3rd Ed., Ch. 28 (Pergamon, Oxford, 1980).

[242] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, 1998).

[243] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).

[244] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[245] F. C. Alcaraz, M. Droz, M. Henkel, and V. Rittenberg, Ann. Phys. **230**, 250 (1994).

[246] J. C. Baez and J. Biamonte, *A Course on Quantum Techniques for Stochastic Mechanics*, preprint arXiv:1209.3632 (2012).

[247] J. Keizer, J. Stat. Phys. **6**, 67 (1972).

[248] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **31**, 526 (1959).

[249] S. Lifson and A. Roig, J. Chem. Phys. **34**, 1963 (1961).

[250] L. Onsager, Phys. Rev. **65**, 117 (1944).

[251] T. Nishino, J. Phys. Soc. Japan **64**, L3598 (1995).

[252] R. Orus and G.Vidal, Phys. Rev. B **78**, 155117 (2008).

[253] A. Slavík, *Product integration, its history and applications* (Matfyzpress, Prague, 2007).

[254] Y. Iba, Journal of Computational and Graphical Statistics **7**, 175 (2000).

[255] O. Golinelli and K. Mallick, J. Phys. A **39**, 12679 (2006).

[256] B. Derrida, Phys. Rep. **301**, 65 (1998).

[257] B. Schmittmann and R. K. P. Zia, *Statistical Mechanics of Driven Diffusive Systems* in *Phase Transitions and Critical Phenomena*, Vol. 17, Eds. C. Domb and J. L. Lebowitz (Academic Press, New York, 1995).

[258] R. Stinchcombe, Adv. Phys. **50**, 431 (2001).

[259] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2001).

[260] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg, J. Stat. Phys. **92**, 151 (2004).

[261] P. Paatero and U. Tapper, Environmetrics **5**, 111 (1994).

[262] D. D. Lee and H. S. Seung, Nature **401**, 788 (1999).

[263] K. Temme and F. Verstraete, Phys. Rev. Lett. **104**, 210502 (2010).

[264] E. Gaussier and C. Goutte, *Relation between PLSA and NMF and implications* in *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, p. 601 (Salvador, 2005).

[265] E. Seneta, *Non-negative matrices and Markov chains*, 2nd Ed. (Springer, New York, 2006).

[266] N. Vasiloglou, A. G. Gray, and D. V. Anderson, *Non-Negative Matrix Factorization, Convexity and Isometry* in *Proceedings of the Ninth SIAM International Conference on Data Mining*, p. 673 (Sparks, Nevada, 2009).

[267] D. D. Lee and H. S. Seung, Advances in Neural Information Processing Systems **13**, 556 (2000).

[268] R. Huebener, V. Nebendahl, and W. Dür, New J. Phys. **12**, 025004 (2010).

[269] F. Fröwis, V. Nebendahl, and W. Dür, Phys. Rev. A **81**, 062337 (2010).

[270] M. C. Bañuls, M. B. Hastings, F. Verstraete, and J. I. Cirac, Phys. Rev. Lett. **102**, 240603 (2009).

[271] E. M. Stoudenmire and S. R. White, New J. Phys. **12**, 055026 (2010).

[272] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, Computational Statistics and Data Analysis **52**, 155 (2007).

[273] C.-J. Lin, Neural Computation **19**, 2756 (2007).

[274] W. Lenz, Physikalische Zeitschrift **21**, 613 (1920).

[275] E. Ising, Z. Phys. **31**, 253 (1925).

[276] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, Nature **440**, 1007 (2006).

[277] E. W. Montroll, Proc. Natl Acad. Sci. USA **78**, 36 (1981).

[278] T. C. Schelling, Journal of Mathematical Sociology **1**, 143 (1971).

[279] S. Bornholdt, Int. J. Mod. Phys. C **12**, 667 (2001).

[280] T. Kaizoji, S. Bornholdt, and Y. Fujiwara, Physica A **316**, 441 (2002).

[281] R. A. Williams, S. Al-Assam, and C. J. Foot, Phys. Rev. Lett. **104**, 050404 (2010).

[282] D. Jaksch, Physics **5**, 60 (2012).

[283] P. W. Anderson, Rev. Mod. Phys. **38**, 298 (1966).

[284] Q. Thomen, J. C. Garreau, and V. Zehnlé, J. Opt. B: Quantum Semiclass. Opt. **6**, 301 (2004).

[285] S. Arora, R. Ge, R. Kannan, and A. Moitra, *Computing a Nonnegative Matrix Factorization – Provably* in *Proceedings of the 44th symposium on Theory of Computing, STOC '12*, p. 145 (ACM, New York, 2012).