

# Many-body dynamics in one dimension

A DISSERTATION PRESENTED  
BY  
ALEXANDER LUKIN  
TO  
THE DEPARTMENT OF PHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
IN THE SUBJECT OF  
PHYSICS

HARVARD UNIVERSITY  
CAMBRIDGE, MASSACHUSETTS  
NOVEMBER 2018

©2018 – ALEXANDER LIKIN  
ALL RIGHTS RESERVED.

# Many-body dynamics in one dimension

## ABSTRACT

Pure quantum states are at odds with the concepts of statistical mechanics as they completely isolated from the environment, therefore, preserve their purity even if the system's parameters are suddenly altered. However, in the wide class of systems coherent many-body dynamic dives local thermalization of the subsystem. In this case, any local observable appears thermal and can be described by the statistical ensemble, even though the full state remains pure. The key to understanding the phenomenon of quantum thermalization lies in the entanglement properties of the system: The presence of entanglement creates local entropy that validates the use of statistical physics.

In this thesis, we experimentally study the emergence of statistical mechanics in a quantum state and observe the fundamental role of quantum entanglement in facilitating this emergence. We perform microscopy on an evolving quantum system, and we see thermalization occur on a local scale, while we measure that the full quantum state remains pure.

The only known robust exception from the paradigm of quantum thermalization is provided by disordered systems. In this case, the particle transport ceases through the system, preventing the formation of sufficient amount of entanglement required for thermalization. However, if the many-body system is interacting, the state develops non-local quantum correlations, that are inac-

cessible to any local observable. These correlations spread exponentially slow in time throughout the entire system and are considered a hallmark feature of many-body-localized states.

We experimentally realize such many-body localized system and observe the formation of non-local quantum correlations, while the particle transport is frozen in the system. Our work experimentally establishes many-body localization as a qualitatively distinct phenomenon from localization in non-interacting, disordered systems.

In order to study the interplay between thermalization and localization in our system, we develop novel methods for high fidelity state preparation and readout. We achieved an exquisite isolation of our system from the environment, that allows us to monitor coherent many-body dynamics over multiple decades of time evolution. This makes our system the most advanced special purpose quantum simulator up to date.

# Contents

0	INTRODUCTION	1
1	BOSE-HUBBARD MODEL UNDER THE MICROSCOPE	7
1.1	Optical lattice . . . . .	7
1.2	Bose-Hubbard model . . . . .	10
1.3	Quantum gas microscope . . . . .	12
1.4	Single-site resolved imaging . . . . .	13
2	HOLOGRAPHIC APPROACH TO ARBITRARY-POTENTIAL GENERATION	15
2.1	Why do we need custom optical potentials? . . . . .	15
2.2	Spatial light modulators . . . . .	16
2.3	Enhancing the resolution through phase control . . . . .	17
2.4	Encoding phase information with a Fourier plane DMD . . . . .	18
2.5	Optical setup . . . . .	20
2.6	Calibration and aberration correction . . . . .	21
2.7	Holographic potentials and their limits . . . . .	24
3	HIGHT FIDELITY INITIALIZATION AND READOUT	29
3.1	Initial states with high fidelity . . . . .	29
3.2	Cutting procedure . . . . .	31
3.3	Atoms life time . . . . .	34
3.4	Atom number resolved state readout . . . . .	35
4	LONG COHERENT DYNAMICS IN OPTICAL LATTICES	38
4.1	Pure versus mixed states . . . . .	38
4.2	Hong-Ou-Mandel effect . . . . .	40
4.3	Hong-Ou-Mandel interference of massive particles . . . . .	41
4.4	Multi-particle case . . . . .	44
4.5	Many-body interference as a measure of indistinguishability . . . . .	45
4.6	Measuring the purity of a state . . . . .	47
4.7	Multi-mode systems . . . . .	48
4.8	Decoherence between copies . . . . .	49
4.9	Increasing the coherence of the system . . . . .	52
4.10	Coherence in the systems with single site control . . . . .	54
4.11	Coherent evolution with a single copy . . . . .	55

<b>5 THERMALIZATION IN AN ISOLATED QUANTUM SYSTEM</b>	<b>57</b>
5.1 Statistical mechanics: from classical to quantum . . . . .	57
5.2 Quench dynamics of an isolated quantum system . . . . .	60
5.3 Experimental protocol . . . . .	61
5.4 Local observables in the thermalized pure state . . . . .	62
5.5 Thermalization of subsystem density matrix . . . . .	65
5.6 Global state purity . . . . .	67
5.7 Entanglement in a quantum system . . . . .	69
5.8 Entanglement entropy dynamics after quench . . . . .	72
5.9 Correspondence between entanglement and thermal entropy . . . . .	72
5.10 Conclusions . . . . .	77
<b>6 MANY-BODY LOCALIZATION</b>	<b>79</b>
6.1 Locally boring globally exciting . . . . .	79
6.2 Disorder potential . . . . .	82
6.3 Calibration of the applied disorder potential . . . . .	84
6.4 Breakdown of thermalization . . . . .	87
6.5 Spatial localization . . . . .	89
6.6 Different types of entanglement . . . . .	91
6.7 MBL as a system of local integrals of motion . . . . .	93
6.8 Accessing configurational entanglement through correlations . . . . .	95
6.9 Entropy dynamics and scaling . . . . .	97
6.10 Conclusion . . . . .	98
<b>7 CONCLUSIONS AND OUTLOOK</b>	<b>102</b>
<b>APPENDIX A COMPUTING EXPECTATION VALUES IN THERMALIZED SYSTEMS</b>	<b>104</b>
<b>REFERENCES</b>	<b>109</b>

DEDICATED TO MY FAMILY.

# Acknowledgments

LOREM IPSUM DOLOR SIT AMET, Rb lab - the best!

# 0

## Introduction

The concept of thermalization is very familiar to every one of us from the everyday life. When two systems are brought in contact with each other and are allowed to exchange the energy and particles, they come to an equilibrium state. When the number of degrees of freedom in one system is substantially larger than in the other one, its state cannot be sufficiently altered due to the coupling between the two. In this case, we call a larger one bath. Coupling to a bath drives the system into thermal equilibrium, such that its state can be described by the thermodynamic ensemble. This presents a very powerful theoretical tool since the complicated microscopic dynamics can be described by a small number of thermodynamic variables.

Seemingly, this concept does not apply to the pure quantum states, as they are completely isolated from the environment. During the time evolution, the purity of the state does not change

even if the system's parameters are suddenly changed. However, in the wide class of systems this coherent many-body dynamic drives local thermalization of the subsystem. In this case, any local observable appears thermal even though the full state remains pure<sup>???</sup>. Strikingly, even with infinitely many copies of such quantum state, the local observables are fundamentally unable to reveal whether this is a single quantum state or a thermal ensemble. In other words, a globally-pure quantum state is apparently indistinguishable from a mixed, globally-entropic thermal ensemble<sup>????</sup>.

This leads us to an interesting thought: the world around us is very well described by the classical thermodynamics, which would suggest that the universe as a whole is one giant mixed state rather than in a pure quantum state. However, since there is no way of locally distinguishing between the two, we might never be able to find an answer to this conundrum, unless we happen to find an identical copy of our universe and perform the measurement of purity similar to those explained in this thesis.

In the classical systems, the thermalization is justified through the notion of ergodicity that  states that the system explores the entire parameter space allowed by the constraints. Which leads to the concept of entropy maximization. In the quantum setting the role of classical entropy is taken by the entanglement: Thermalizing systems develop an extensive amount of entanglement between the subsystems,  therefore ignoring the information about the rest of the system puts the subsystem in the statistical mixture of a large number of states. From this perspective, we can say that thermalizing quantum systems become an effective bath to themselves.

A more formal explanation of the quantum thermalization is given by the Eigenstate Thermalization Hypothesis, which connects it with the properties of the many-body eigenstates of the system. Although being numerically verified for in the number of systems<sup>??????</sup>, the full theoretical proof of ETH remains elusive to this day. Therefore, experimental studies of such this phenomenon are of a great interest in the field.

Recent experiments have demonstrated analogies between classical chaotic dynamics and the role of entanglement in few-qubit spin systems<sup>?</sup>, as well as the dynamics of thermalization within an ion system<sup>?</sup>. Furthermore, studies of bulk gases have shown the emergence of thermal ensembles and the effects of conserved quantities in isolated quantum systems through macroscopic observables and correlation functions<sup>???</sup>. However, those experiments were not able to directly verify the purity of the full state – key feature distinguishing quantum thermalization from the coupling to the external classical bath.

An exception to the paradigm of quantum thermalization can be found in the integrable systems<sup>??</sup>,  fine-tuning of the parameters results in the emergence of an extensive number of the conserved quantities even at the local level. However, small deviations from the fine-tuned point immediately restore the system's ergodicity<sup>??</sup>.

The only known robust example of thermalization breakdown is provided in the disordered systems, which localizes the particles<sup>?</sup>, leading to suppression of the transport through the system. This prevents the buildup of entanglement in the system, which is crucial for thermalization. Surprisingly, the localization persists even in the presence of interactions between the particles, this phenomenon is called many-body localization (MBL)<sup>????</sup>.

Since in the MBL state the particles are localized one might expect, that there would be no dynamics in such a system, which turns out to be true for local observables. However, numerical studies have shown that entanglement still builds up in the system<sup>??</sup>, although it happens over exponentially long timescales. This happens due to the presence of interactions, which give rise to slow coherent many-body dynamics that generate non-local correlations, inaccessible to local observables<sup>???</sup>. These dynamics are considered to be the hallmark of MBL that distinguishes it from its non-interacting counterpart, called Anderson localization<sup>?</sup>.

Experimental studies have identified MBL through the persistence of the initial density distribution<sup>??????</sup> and two-point correlation functions during transient dynamics<sup>?</sup>. However, the

observation of the buildup of non-local correlations across the full system has remained elusive, because it requires exquisite control over the system's coherence over exponentially long evolution time.

In this thesis, I present experimental techniques that extend the capabilities of our quantum gas microscope<sup>7</sup> to explore the coherent many-body dynamics of the one-dimensional Bose-Hubbard system over multiple orders of time evolution, exceeding the previously achieved timescales by more than an order of magnitude<sup>7</sup>. Using holographic beam shaping we achieve precise control over optical potentials at a single site level in our system. The ability to address individual sites enables a deterministic preparation of desired initial states as well as the ability to control the potential landscape of the system during the time evolution.

The advanced capabilities of our quantum gas microscope enable collective measurement on multiple copies of the system, which enables the measurement of purity of a quantum state and. It also gives us an access to the entanglement entropy between the subsystems of the quantum state. The ability to control optical potential with high accuracy on a single site level enables us to engineer custom potential landscapes in the system.

A quick summary of the experiments presented in this work is provided as follows, according to the organization of the rest of the thesis:

- **Chapter 1** gives the overview of the experimental setup, the quantum gas microscope. We highlight all parts of the setup that would be important for the future discussions in this thesis, and summarize the important concepts required for understanding the theory of cold-atom experiments in the optical lattices.
- **Chapter 2** describes how we use digital micro-mirror device (DMD) in the Fourier plane of our imaging system in order to create arbitrary potential landscapes on the atoms. We summarize the calibration procedure that allows us to achieve a diffraction-limited performance

on the atoms. We also discuss the limits of this technique related to the hologram binarization.

- **Chapter 3** shows how we achieve initial state preparation with near-perfect fidelity, using single site control described in the previous chapter. We discuss one of the decoherence effects coming from the atom number loss in our system, and describe how we can mitigate it by using full number-resolved readout of the final state, developed in this chapter.
- **Chapter 4** focuses on characterization of the coherence in our system. We introduce the concept of many-body interference using  beamsplitter technique in order to measure the purity of the quantum state. Using this technique we demonstrate the quantum coherence in  $\sim 12000$  dimensional Hilbert space over multiple orders of time evolution. We also introduce the measurements of coherent quantum evolution with a single copy, that allows us to extend our characterization techniques beyond the two copy method.
- **Chapter 5** describes the experiments, where we observe how a small subsystem of a globally pure state exhibit thermalization during the time evolution of the system. We measure the entanglement entropy of the subsystems of the thermalization state, and determine the correspondence of the quantum entanglement entropy between the subsystems of the pure state and the classical thermal entropy which explains the microscopic origins of thermalization in the system.
- **Chapter 6** presents the experiments with disordered system. We realize a one-dimensional many-body localized system and characterize its key properties: breakdown of thermalization, spatial localization, and study its entanglement properties over multiple decades of time evolution. In order to distill the quantum phenomenon in the system we introduce and separately measure two contributions to the entanglement entropy corresponding to a

particle transport and quantum phase evolutions in the system.

- **Chapter 7** concludes with the summery of the work done in this thesis.

# 1

## Bose-Hubbard model under the microscope

### 1.1 OPTICAL LATTICE

Neutral atoms provide are a great tool to study the coherent quantum phenomenon, since they are very well isolated from the environment including the interactions among themselves. However, this has one negative consequence, if we would like to study an interacting itinerant system we need to find a way to make the interaction energy between them to be comparable or even larger than their kinetic energy. The motional degree of freedom of the atoms can be controlled by going form the free space to the lattice geometry. This also allows to study the system in differ-

ent special configurations, such as square<sup>?</sup>, hexagonal<sup>??</sup>, kagome<sup>??</sup>, or tunable dimerized<sup>??</sup>.

In order to create a periodic potential for the atoms, one can utilize atom-light interactions. The electric field of a light beam induces a dipole moment in the atoms, which intern interacts with the electric field. This causes a shift in atomic energy levels called the ac-Stark shift. If the detuning of the light frequency from the atomic resonance is large compared to the transition linewidth, we can neglect the effects associated with photon scattering. As a result, the energy shift acts as a conservative potential for the atom.

For a two-level atom in a monochromatic laser field, if the laser detuning is large the rotating wave approximation can be used. The resulting conservative dipole potential is given by:

$$V_{dipole}(r) = \frac{3\pi c^2}{2\omega_0^3} \frac{\Gamma}{\Delta} I(r), \quad (1.1)$$

where  $\omega_0$  is the atomic resonance frequency,  $\Gamma$  is the linewidth,  $\Delta = \omega - \omega_0$  is the laser frequency detuning and  $I(r)$  is the laser intensity. Depending on the detuning the atoms are attracted to the regions of maximum (red detuned) or minimum (blue detuned) light intensity.

A simple way to generate an optical lattice in one dimension is to shine two counter propagating laser beams of the same wavelength  $\lambda$ . The interference between them results in a standing wave intensity pattern with a period of  $a = \lambda/2$ , therefore, the resulting potential is given by:

$$V(x) = V_0 \cos^2(k_l x), \quad (1.2)$$

where  $k_l = \frac{\pi}{a} = \frac{2\pi}{\lambda}$  is the lattice momentum. The lattice momentum gives a natural energy scale in the system called recoil energy  $E_r = \frac{\hbar^2 k_l^2}{2m} = \frac{\pi^2 \hbar^2}{2ma^2}$ . Throughout this thesis we will set  $\hbar = 1$ , unless otherwise specified.

A single-particle wave functions in the periodic potential are Bloch waves of the form:

$$\psi_q^{(n)}(x) = e^{iqx/\hbar} u_q^{(n)}(x), \quad (1.3)$$

where  $u$  is a function with the same periodicity as the lattice. The wavefunctions are labeled by the band index  $n$  and quasi momentum  $q \subset (-k_l; k_l]$ . The restrictions on the available momenta of the particles comes from the periodicity of the Hamiltonian.

When the lattice depth is sufficiently large, the dynamics of the atoms can be understood as hopping of particles between the sites. In order to describe this, the Bloch wavefunctions can be combined to form another complete orthogonal set of wavefunctions that are maximally localized on the individual sites of the lattice. This set is called Wannier functions and can be written up to normalization in terms of Bloch wavefunctions for a given site  $i$  as:

$$w_n(x - x_i) = \frac{1}{\mathcal{N}} \sum_q e^{iqx_i/\hbar} \psi_q^{(n)}(x). \quad (1.4)$$

In order to make the Wannier functions maximally localized we need to choose the phases between the Bloch waves accordingly. Since all even bands Bloch waves have maximum amplitude on the lattice sites, we can pick the phase such that for the lattice site of interest  $i$  it is real and positive. On the other hand every odd band has zero amplitude on the lattice sites, therefore, we can pick the phase such, that on site  $i$  the derivative is real and positive.

$$\psi_q^{(n)}(x) \rightarrow \begin{cases} \psi_q^{(n)}(x) \cdot \exp(-i \cdot \arg[\psi_q^{(n)}(x_i)]), & n \text{ even} \\ \psi_q^{(n)}(x) \cdot \exp(-i \cdot \arg[\frac{\partial \psi_q^{(n)}(x)}{\partial x}|_{x=x_i}]), & n \text{ odd} \end{cases} \quad (1.5)$$

## 1.2 BOSE-HUBBARD MODEL

The behaviour of atoms in the lattice can be well described using Bose-Hubbard model, given by the Hamiltonian:

$$\mathcal{H} = -J \sum_{i,j} (a_i^\dagger a_j + h.c.) + \frac{U}{2} \sum_i n_i(n_i - 1), \quad (1.6)$$

where  $a_i$  is a particle annihilation operator on the site  $i$ , and  $n_i = a_i^\dagger a_i$  is a number operator. This Hamiltonian can be easily understood from the Wannier functions perspective. The first term describes the tunnelling between neighbouring lattice sites and originates from the non-zero amplitude of the Wannier function on those sites. The second term represents the interaction energy shift when multiple particles occupy the same site. It arises from the contact interactions between the particles. The presence of the optical lattice greatly increases the intercalation strength between the atoms, due to the tight confinement of the Wannier functions compared to the free particle case.

In the experiment, the ratio between  $U$  and  $J$  can be tuned in two different ways. One way is to fix the lattice depth and control the contact interaction strength via Feshbach resonance<sup>?</sup>. However, due to the properties of our atom of choice, the implementation of this technique is rather challenging in our experiment. The other approach relies on the different scaling of the parameters as a function of the lattice depth: Since the overlap of the Wannier function with the neighbouring lattice sites decreases exponentially with the lattice depth, the tunnelling goes down as the depth increases. On the other hand, the interaction energy depends on the size of the Wannier function, which increases approximately as a fourth root of the lattice depth. Therefore by changing the lattice depth between  $2 - 45E_r$  we can tune the  $U/J$  ratio in the range between  $\sim 0.3 - 20000$ .

The interplay between the tunnelling and interactions leads to different ground states of this

model called superfluid and Mott insulator. In the regime where tunnelling dominates ( $J \gg U$ ) each particle tries to minimize its kinetic energy by demoralizing uniformly throughout the lattice. In the thermodynamic limit, it corresponds to the situation, where each particle is in the lowest quasi-momentum  $q = 0$ . In this case, the wavefunction of the system can be written as:

$$|\Psi_{SF}\rangle = \frac{1}{\sqrt{N}}(a_{q=0}^\dagger)^N |0\rangle \propto \left(\frac{1}{\sqrt{N}} \sum_i a_i^\dagger\right)^N |0\rangle, \quad (1.7)$$

where the last approximation becomes exact in the thermodynamic limit.

Conceptually a superfluid state represents a situation where each lattice site is occupied by a coherent state

$$|\Psi_{SF}\rangle \approx \prod_i |\alpha\rangle_i = \prod_i e^{-\frac{|\alpha|}{2} - \alpha a_i^\dagger} |0\rangle, \quad (1.8)$$

where average density  $\langle n \rangle = |\alpha|$  and the phase of each coherent state  $\arg(\alpha)$  is the same for every site of the lattice.

On the other side of the transition when interactions dominate ( $U \gg J$ ). In the extreme case, when  $J = 0$  every particle is localized to a single site. In order to minimize the interaction energy, it becomes favourable for the system to distribute atoms uniformly over all lattice sites. In the case of commensurate filling, when the number of particles is equal to the number of lattice sites times an integer functor  $n$ , the Mott insulator wave function becomes:

$$|\Psi_{MI}\rangle = \prod_i (a_i^\dagger)^n |0\rangle. \quad (1.9)$$

The transition between these two phases can be understood qualitatively from the Mott side in the case of unity filing ( $n = 1$ ) as following: Starting from the ground state for  $J = 0$  we can perturbatively include the tunnelling, this results in the coherent admixture of excitations in the form of particle-hole pairs on top of the perfect Mott ground state. As  $J$  increases the density of

excitations raises until all the particles completely delocalize.

In our experiments, we use high fidelity Mott insulating stats in order to deterministically initialize our system with the exact same state for every run of the experiment.

### 1.3 QUANTUM GAS MICROSCOPE

The experiments presented in this thesis were carried out on the quantum gas microscope. We use  $^{87}\text{Rb}$  atoms, that are confined to a single pancake of the optical lattice in  $z$  direction and loaded into a  $2D$  optical lattice in the  $x - y$  plane. The atoms are positioned in the focus of a high-resolution imaging system with a numerical aperture (NA) of 0.8, which allows us to image the gas with a single site resolution. Below, we will summarise some important aspects of the setup. Further details about the apparatus can be found in the previous theses from our group<sup>???</sup>.

In the heart of the experiment we have a long working distance microscope objective with NA = 0.55 outside the vacuum glass cell. Inside the glass cell, there is a hemispherical lens made out of the material with an index of refraction  $n = 1.45$ , mounted on top of the super polished substrate (see fig. ??). The atoms are confined to the 6th minimum of  $1.5\mu\text{m}$  spacing optical lattice,  $10\mu\text{m}$  below the super-polished substrate. Refraction on the bottom of the hemisphere leads to an effect similar to 'solid immersion', which increase the NA of the combined imaging system to 0.8. The resulting diffraction-limited resolution at the  $D_2$  transition line of  $^{87}\text{Rb}$  ( $\lambda = 780\text{nm}$ ) is  $600\text{nm}$ .

Unlike the typical lattice experiments, where the lattice potential is created by retroreflecting the laser beam onto itself. We utilize the high numerical aperture of our imaging system, to create the  $2D$  lattice by 4f imaging a pair of holographic masks with additional Fourier filtering (see fig. ??). The resulting potential has a lattice spacing of  $\sim 68\text{nm}$ , which gives the recoil energy in both directions of  $E_r = 2\pi \cdot 1240\text{Hz}$ .

This technique has one major advantage and one drawback. The advantage is that the lattice

spacing in the image plane does not depend on the wavelength of the lattice light. This effect is due to the property of 4f imaging system. As we discuss below, it allows us to increase the depth of the lattice by several orders of magnitude without increasing the light intensity on the atoms. However, the disadvantage is that the resulting potential has slow varying random offsets due to the interference of the lattice beams with the stray light. In order to decrease the amount of disorder in the system, we rely on the wavelength insensitivity of this technique in order to produce the lattice with  $\sim 3\text{nm}$  broad light source with the central wavelength of  $\sim 760\text{nm}$  in combination with additional spatial filtering in the Fourier plane. This allows us to eliminate some of the unwanted interference effects of the light, reflected from the optical surfaces along the imaging path. Further details on how to improve the quality of projected potential can be found in the thesis of Ruichao MA<sup>7</sup>.

#### 1.4 SINGLE-SITE RESOLVED IMAGING

In order to image the atoms at the end of the experiment, the depth of the 2D lattice is rapidly increased to  $\sim 5000E_r$ , which freezes the tunnelling in the  $x - y$  plane. This is done by switching the lattice light to  $\sim 30\text{GHz}$  blue detuned from the Rubidium  $D_1$  line ( $\lambda = 795\text{nm}$ ). Since the lattice light is close detuned the effects of spontaneous emission are not negligible anymore. In order to prevent the atoms from heating<sup>????</sup>, which would lead to their loss from the lattice, we cool the atoms using 80MHz detuned optical molasses on the  $D_2$  line ( $\lambda = 780\text{nm}$ ).

By collecting the photons, scattered from the molasses, on the EMCCD camera we can obtain single-site resolved images of our system (see fig. ??). During the typical exposure time of 500ms we collect  $\sim 2000$  photons from individual atoms, which correspond to  $\sim 10\%$  of the total number of scattered photons.

One might notice, that the image contains two different types of site: occupied and unoccupied, although there is no restriction for multiple bosons to occupy the same site. This is due

to the so-called light-assisted collisions<sup>?????</sup>, that occur during the imaging process. When two atoms are tightly confined in space (such as in the case of the deep imaging lattice), the presence of red-detuned light makes the atoms interact via attractive dipole-dipole potential. The energy release from such a process gives them enough kinetic energy to escape the optical lattice. In our parameter regime, this process happens on the  $\sim 100\text{ms}$  timescales, orders of magnitude faster than the time to collect enough photons for imaging. Therefore all images that we get are so-called parity projected: each site with an odd number of atom appears bright whereas sites with even number appear dark.

The obtained image is analyzed by fitting the amplitude of separately determined point spread function on each individual lattice site. The amplitude distribution has a clear bimodal structure<sup>?</sup>, therefore can be easily binarized. The fidelity of distinguishing between occupied and unoccupied sites is  $> 99\%$ .

# 2

## Holographic approach to arbitrary-potential generation

### 2.1 WHY DO WE NEED CUSTOM OPTICAL POTENTIALS?

One of the advantages of cold-atom experiments is the ability to prepare and probe the system on its characteristic time and length scales (for lattice models, those would be a tunneling time and lattice spacing respectively). Quantum gas microscopy enables us to observe our lattice system with single site resolution, which led to the first single-site-resolved observations of superfluid to Mott insulator<sup>??</sup> as well as paramagnet-to-antiferromagnet<sup>??</sup> quantum phase transitions. However, in order to fully enable the capabilities of such quantum simulators, one needs a tool to arbi-

trarily control the shape of optical potentials in the system. This would enable the preparation of a variety of different initial states with high fidelity<sup>??</sup>, and studies of systems with complex<sup>??</sup> as well as dynamically changing<sup>?</sup> geometries. To achieve this, one has to control the light intensity on a single lattice site scale, which can be achieved by means of spatial light modulators (SLM).

## 2.2 SPATIAL LIGHT MODULATORS

There are two main types of SLMs that are being used in such setups: liquid-crystal displays (LCD) and digital micro-mirror device (DMD). LCDs can be configured to provide amplitude modulation or phase modulation of the incident beam, using the birefringence of liquid crystals. With optimization and feedback algorithms<sup>??</sup>, it is possible to create complex, re-programmable potentials, using this technology. However, this type of SLM has one major drawback when it comes to cold-atom experiments. Due to the nature of liquid crystals, they have to be placed into a switching electric field that oscillates with frequencies in the  $k\text{Hz}$  range, causing the output light to blink at the same frequency. This blinking could potentially be a problem since the typical trap frequencies of the atoms in optical lattice experiments lie in the same frequency range, and hence the blinking light might induce undesired heating in the system.

Another approach is to use a DMD, which consists of CMOS arrays of micromirrors on torsion hinges which can be switched between two angles corresponding to the *on* and *off* directions. These devices are typically used in commercial video projectors and are available with resolutions up to  $2560 \times 1600$  pixels. DMDs can statically hold a desired pattern for an extended period of time, which mitigates the blinking issue of LCDs. However, this comes at a cost since, in order to control the phase information of the outgoing wavefront, one has to use special techniques<sup>??</sup> that result in low power efficiency.

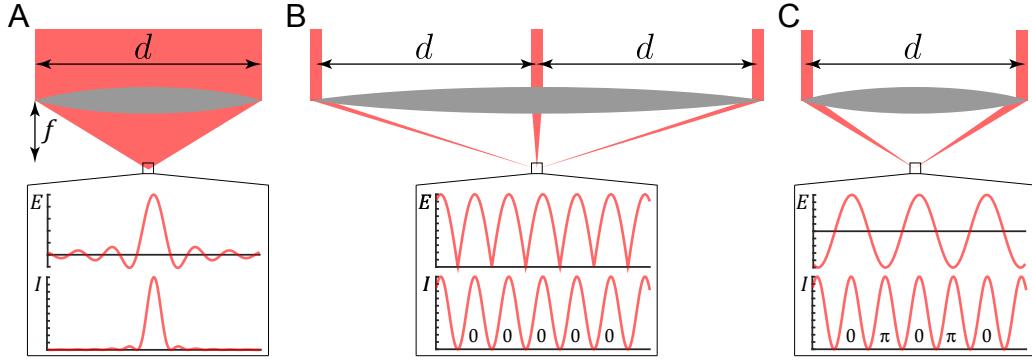
The device used in our experiment is the Keynote Photonics FlexLight X3, which uses the Texas Instruments DLP 5500 chipset. The  $1024 \times 768$  micromirror array has a pitch of  $10.8\mu\text{m}$ , mi-

cromirror tilts at  $\pm 12^\circ$  from the normal, and an update rate of 5kHz. The time required to switch between individual patterns is even shorter, on the order of a few  $\mu\text{s}$ , which can allow dynamical control over the potential.

A straightforward way to use SLMs in the microscope setup is by directly imaging it onto the atoms<sup>???</sup>. The gray scaling of the image can be achieved by setting the magnification such, that a single lattice site corresponds to an area of the SLM that contains multiple pixels. This method is very powerful for creating large-scale, smoothly-varying potentials. However, there are two major challenges, that can limit the performance of the SLM in a real setup: *First*, the wavelength of conservative light potentials, that one wants to project, is generally very different from the atom imaging wavelength, which most quantum gas microscopes are optimized for. This can lead to significant distortion of the desired wavefront due to chromatic aberrations if the imaging system is not specifically corrected for that. The first two leading orders of aberrations, in this case, are given by defocus and spherical aberration. This problem can be solved by either using an achromatic imaging system, that is wavelength insensitive in the desired range, or precompensating the imaged wavefront for the defocus caused by the imaging. *Second*, it is fundamentally impossible to create 100% intensity modulation of the potential on the single site scale without using phase control of the wavefront.

### 2.3 ENHANCING THE RESOLUTION THROUGH PHASE CONTROL

To illustrate the last point, let's consider a single lens with a focal length  $f$  and a fixed aperture of size  $d$  in the Fourier plane (see fig 2.1). To keep the math simple let's consider only two spatial dimensions  $x$  and  $z$ . The sharpest feature that one can make in the image plane with constant phase wavefront in the Fourier plane is the so-called point spread function, with the resulting light intensity in the image plane given by  $I(x) \sim \text{sinc}^2(\frac{\pi}{\Theta}x)$ , where  $\Theta = \frac{\lambda f}{d}$  (see fig 2.1 A). The question becomes: how can one make an intensity pattern that is periodic with period  $\Theta$ ? Two



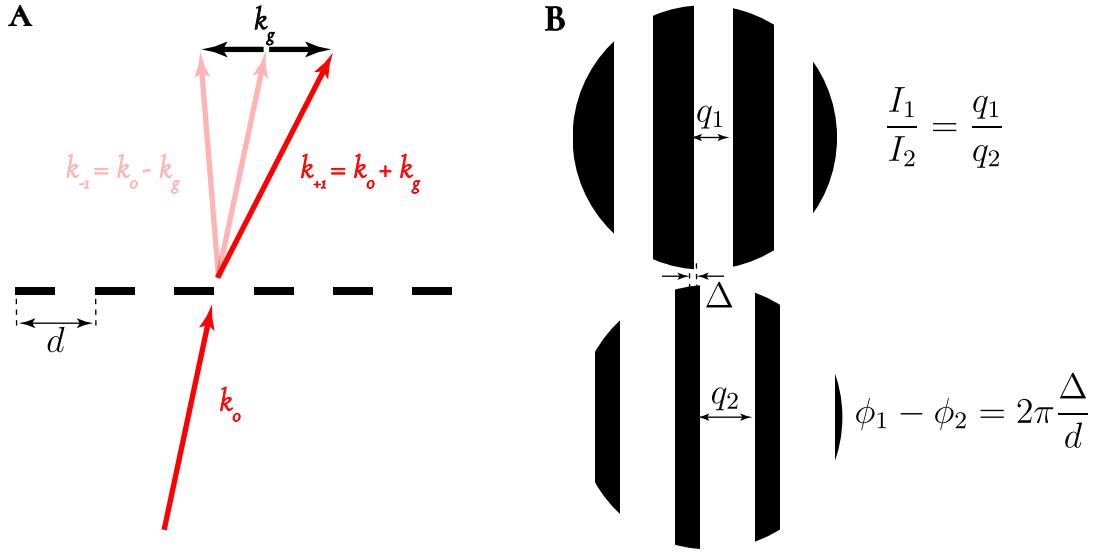
**Figure 2.1:** Enhancing image plane resolution with phase control. A Homogeneous illumination of a lens with focal distance  $f$  and a Fourier plane size  $d$  results in so-called point spread function. It determines the size of the smallest point source feature, that can be resolved through such an imaging system. The zoom in shows the image plane electric field and intensity profile for each case. B Double the size of the Fourier plane is required in order to make a periodic structure with, a period corresponding to the size of the point spread function from A, with a constant phase in the image plane. C The pattern with the same periodicity can be created using the Fourier plane of the original size, however, the phase of the adjacent intensity maxima alternate between  $0$  and  $\pi$  in this case.

different electric field profiles would give us the intensity profile that satisfies this requirement:

$E_1 \sim \sqrt{(1 + \cos(\frac{\pi}{\Theta}x))/2}$  and  $E_2 \sim \cos(\frac{\pi}{2\Theta}x)$ , both yielding  $I \sim (1 + \cos(\frac{\pi}{\Theta}x))/2$  (see fig 2.1 B,C). The first one has only positive electric field components and hence a uniform phase, however it's Fourier transform requires the Fourier plane, that is double the size of our initial point spread function. The second one has the same Fourier plane size as the one we started with at the expense of having electric field oscillate between positive and negative values, or in other words the phase of adjacent maxima alternate between  $0$  and  $\pi$ . This example shows us, that using the phase control of the potential in the image plane, allows us to increase the image plane resolution by a factor of two for a fixed aperture in the Fourier plane.

## 2.4 ENCODING PHASE INFORMATION WITH A FOURIER PLANE DMD

Another practical challenge comes from the fact that every physical imaging system has aberrations. Hence, one needs to correct for all distortions that the beam encounters on its way from



**Figure 2.2:** Encoding phase and amplitude information with a grating. (A) When a light with initial wave vector  $k_o$  passes through an amplitude grating with period  $d$ , its outgoing wave vector can be changed by any integer multiple of grating wave vector  $k_g \propto \frac{1}{d}$ . (B) Two patches (1 - top, 2 - bottom) with the grating of the same period  $d$  but different duty cycles  $q_1$  and  $q_2$  are shown. Since the duty cycle determines how much light passes through, the intensity  $I$  of the outgoing light in each case is proportional to the duty cycle  $\frac{q}{d}$ . The information about the local phase of the grating is imprinted onto the phase of the outgoing electric field  $\phi$ .

the DMD to the desired image plane (in our case, we call it the atom plane). Note, that one needs to correct for both the intensity pattern that illuminates the DMD as well as any phase deviation from the desired wavefront. The latter is particularly challenging for imaging systems with high numerical aperture, due to the breakdown of the paraxial approximation. In order to achieve the best possible quality of the optical potentials, it would be desirable to be able to measure and actively compensate for aberrations in the system. Here the use of the DMD in the Fourier plane becomes particularly useful.

the information about the local phase of the grating is imprinted onto th

Since DMDs don't have a built-in way to control the phase of the light field, one needs to find a way to encode this information using a binary structure of the DMD pixels. In our experiment

we use a holographic approach that works as follows (see fig. 2.2): Consider a plane wave incident onto an amplitude diffraction grating. The grating will create multiple diffraction orders on the output, each satisfying the condition  $k_n = k_o + n * k_g$ , where  $k_n$ ,  $k_o$  and  $k_g$  are the wave vector of the  $n$ -th diffraction order, the incident light and the grating, respectively. The grating wave vector is inversely proportional to the period of the grating  $k_g \propto \frac{1}{d}$ . We focus on the first diffraction order for the rest of the discussion. The overall spatial translation of the diffraction grating gets mapped onto the phase of the electric field in the diffracted order (see fig 2.2 B), whereas the ratio between the *on* and *off* pixels within one period of the grating (called the duty cycle) will result in the intensity modulation of the outgoing light. Therefore by controlling the phase and duty cycle of the underlying grating, we can locally control the phase and amplitude of the outgoing waveform at the DMD<sup>7</sup>.

## 2.5 OPTICAL SETUP

We chose to use an aperture of 500 pixels (which approximately equals to 5.4mm) in diameter on our DMD. In order to image it onto the objective Fourier plane, which is 18mm in diameter we use the optical set up shown in figure ?? A. In order to make the alignment easier and also to be able to control the image plane position of the potential without changing the hologram we use IP and FP mirrors (see fig. ??). The FP mirror moves the beam primarily in the Fourier plane, since it's located very close to the intermediate image plane of the imaging system, and hence does not change the position of the beam in the image plane. The IP mirror position is chosen to be close to the Fourier plane, such that it primarily displaces the beam in the image plane. In order to understand the exact location of this mirror, it is instructive to look at where the plane of the DMD (which is the Fourier plane of the overall imaging system with respect to the atoms) is imaged to (see fig ?? B). It is clear from the figure that this mirror is positioned exactly in the Fourier plane of the imaging system and hence moves the beam in the image plane.

Typically, we use the DMD to create potentials that have single-site resolved features in one direction and a slowly varying amplitude in the other direction. The Fourier transform of such a pattern occupies the entire Fourier plane along the direction of the narrow feature and only a small fraction along the other one which is inversely proportional to the size of the slowly varying amplitude. This implies that, if we would like to achieve the best power efficiency, we would like to match this profile with the intensity of the incoming beam. We accomplish this with an additional elliptical illumination channel with aspect ratio  $\frac{1}{7.5}$ . The round illumination beam is primarily used for calibration purposes.

Since the DMD and the optical lattice don't have a common interferometric reference their relative position may drift in time. This means that we need a way to track and stabilize their relative alignment. We do so by imaging the reflections off the super-polished substrate of both the lattice and the DMD on a separate camera in the intermediate image plane (see fig. ??). This method provides us with a reference that almost coincides with the atom plane. We apply feedback using the IP mirror of the DMD once every experimental cycle, corresponding to a frequency of  $\sim 17\text{mHz}$ . This allows us to stabilize the relative position to within 0.03 lattice sites RMS. The residual noise is attributed to air currents and mechanical vibrations of the breadboards.

## 2.6 CALIBRATION AND ABERRATION CORRECTION

In order to achieve a diffraction-limited resolution in the atom plane, we need a way to map out the aberrations that occur on the optical path from the DMD to the atom plane. The two main sources of the phase error in our system are the curvature of the DMD chip that is present due to the manufacturing process, and chromatic aberrations of the objective lens, which has a  $\sim 1\frac{\mu\text{m}}{\text{nm}}$  chromatic shift. We also need to compensate for the intensity profile of the beam that illuminates the DMD, in order to have spatial control over the amplitude of the electric field.

An aberration-free lens bends a beam parallel to the optical axis such that it passes through the focal point, and makes the phase of any such beam equal at that spot. If we take two beams in the Fourier plane, they will interfere to create a periodic intensity modulation in the image plane, with a period inversely proportional to their spacing in the Fourier plane (see fig. ??). The global phase of this modulation is given by the relative phase between the beams, and the amplitude is proportional to the product of electric fields. Hence, by using one beam as a reference one can map out the relative phase and amplitude of any other beam by measuring the phase and amplitude of the resulting interference pattern<sup>?</sup>.

First, we perform the calibration in the intermediate image plane. A photodiode with a pin-hole is used to detect the resulting intensity modulation (for a more detailed description of this procedure see the Master's thesis by Philip Zupancic<sup>?</sup>). The size of the pinhole (in our case  $10\mu\text{m}$ ) is chosen such that it is much smaller than the smallest period of the intensity modulation in the corresponding image plane, created by interfering two beams from the opposite ends of the Fourier plane. We use patches with diameter  $\sim \frac{1}{15}$  of the full Fourier plane to create local beams from the previous example. This procedure is relatively fast and has a high signal to noise ratio, resulting in a typical wavefront flatness of  $\sim \frac{\lambda}{40}$ . By using a series of single reference patches we are also able to create an amplitude map of the illumination beam<sup>?</sup>.

Next, we apply a similar technique to create a phase map from the intermediate plane to the atom plane, using atoms as a detector. We use weekly a weakly-interacting BEC of  $\sim 400$  atoms in a harmonic trap with a diameter of  $\sim 30$  lattice sites to read off the phase of the interference pattern. The intensity profile results in a density modulation of the atomic cloud due to the Stark shift exerted onto the atoms by the light (see fig. ??). This technique has a number of limitations: since we rely on atoms to map out the resulting intensity pattern, we require at least a few oscillations of the potential within our cloud size. This limits our ability to measure a potential with large spacing or, from the Fourier plane perspective, patches that are close together. On the other

hand, the patches at the opposite ends of the Fourier plane aperture, are also challenging to measure, since they result in an intensity modulation with a period of  $\sim 1$  lattice site. Hence, using the same technique with a fixed reference patch that was used in the intermediate plane does not seem feasible. It is also challenging to acquire enough statistics in order to properly map out the intensity of the resulting potential, so the creation of the amplitude map is not feasible either. Here, one should note that the amplitude map should stay unchanged from the intermediate plane since there is no spatial dependence of the transmission of the optical elements in between the two planes. Hence, it is sufficient to use the amplitude calibration from the intermediate image plane all the way to the atom plane.

In order to overcome the above limitations, instead of keeping one patch as a reference we keep the spacing between two patches fixed and then scan this pair across the Fourier plane (see fig. ??). The spacing is chosen such that the resulting intensity modulation has a period of  $\sim 4$  lattice sites, allowing us to fit  $\sim 7$  periods of modulation across the cloud. The resulting atom density distribution is averaged over multiple experimental realizations for a given pair of patches. During this measurement, any relative drift between the DMD pattern and the lattice directly results in an error of the extracted phase. Therefore, we use our DMD tracking routine during the measurement to illuminate this source of error. We extract the phase difference between a pair of patches by fitting a sine-wave to the resulting average density distribution. This procedure leads to a 66% confidence interval of the fit for the phase estimation of  $0.27\text{rad}$  or  $\sim \frac{\lambda}{23}$ .

Since we only measure the difference between a pair of patches, we effectively measure the derivative of the phase profile with this method. We reconstruct the phase by using a polynomial fit up to 7th order. This procedure gives us a phase profile of the wavefront along one line. In order to reconstruct the full 2D map we repeat this process along 6 different cuts through the Fourier plane, and then use a smooth surface parametrized by the Zernike polynomials up to the 6th order to fit the full 2D surface. Since, in the typical experimental setting we only use a small

part of the Fourier plane for our patterns, after applying the full 2D phase correction according to the reconstructed phase front, we repeat the calibration along the line through the center of the pattern region. The above procedure results in the RMS error of the phase between two patches on opposite sides of the Fourier plane being better than  $\frac{\lambda}{14}$ .

## 2.7 HOLOGRAPHIC POTENTIALS AND THEIR LIMITS

To create a desired potential on the atoms we rely on the imaging system property that the electric field in the image plane is the Fourier transform of the field in the Fourier plane. Therefore, in order to create a desired potential, we first numerically find its Fourier transform (note that in general, it will be a complex-valued function). Then, in order to represent both amplitude and phase information of the desired potential we, use the grating method, which encodes them in the local phase and duty cycle of the grating respectively (see fig. ??). As long as we know the aberration profile in our system, we can correct for it by locally modifying the phase to  $\phi_{total} = \phi_{target} - \phi_{aberrations}$  and amplitude  $A_{total} = \frac{A_{target}}{A_{beam}}$ .

If our DMD had pixels with a gray scale, the potential would be limited by the discretization of individual pixels, however the binary nature of the device creates additional noise that limits the precision of the resulting potentials. One straightforward way to binarize the hologram is to switch the pixel with index  $(i, j)$  on if

$$|(k_x i + k_y j + \phi_{total}) \bmod 2\pi| < \arcsin(A_{total}), \quad (2.1)$$

and off otherwise. Here  $k_x$  and  $k_y$  are  $x$  and  $y$  components of an a priori chosen wave vector of the underlying grating. This method has one straightforward problem, namely when one of the components of the grating wave vector is zero. In this case the stripes of the grating exactly coincide with pixels of the hologram, and if the wave vector is not equal to an integer number of

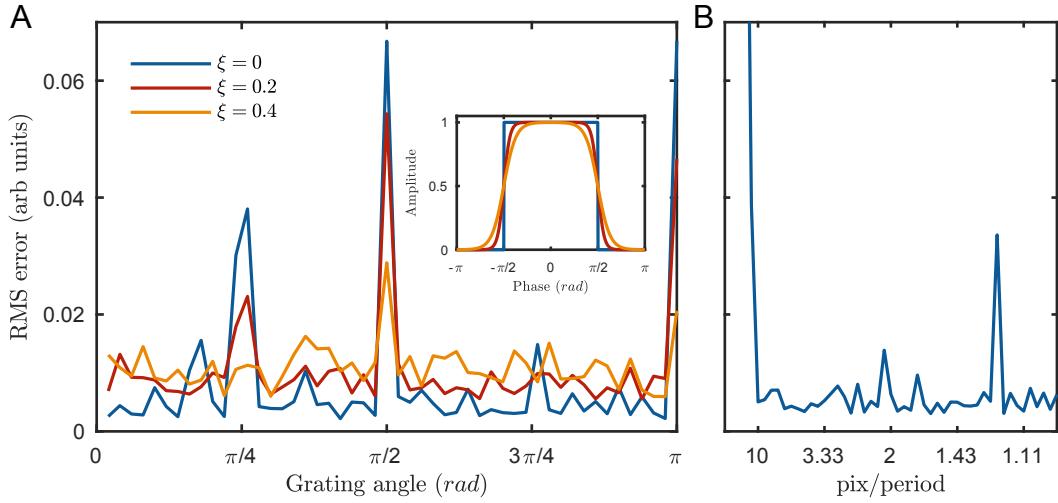
pixels, result in the distortion of the desired potential (see fig. 2.3A). This problem can be avoided by slightly randomizing the pixels at the edges of the grating slits, using the function

$$\frac{1}{2}(\tanh((p(i,j) + \arcsin(A_{total}))/\xi) - \tanh((p(i,j) - \arcsin(A_{total}))/\xi)), \quad (2.2)$$

where  $p(i,j) = |k_x i + k_y j + \phi_{total} \bmod 2\pi|$  and  $\xi$  is a parameter that controls the "smoothness" of the distribution (see fig. 2.3 A inset). Now we compare if the value of this function at any pixel position is greater than a randomly chosen number between 0 and 1, we set pixel *on* and *off* otherwise.

To study the effects of the smoothing potential, we perform numerical studies of the holograms. We use a fast Fourier transform (FFT) to determine the resulting image plane potential corresponding to a given hologram. In this study, we always compare a binarized hologram to its gray-scale counterpart, such that the binarization noise due to the finite size of the physical pixels is identical for both. From figure 2.3A it is clear that this method helps to reduce the effect of "bad" grating alignments, however, it comes at the cost of increasing the overall noise of the potential. This is easy to understand, since the more random we make the potential, the more likely it is to flip an otherwise correct pixel value. It follows from the discussion above that, in order to achieve the best performance with the holograms, it is beneficial to use a nearly deterministic method of binarization away from "bad" angles.

Another parameter that could affect the hologram quality is the absolute value of the grating wave vector, or, in simple terms, the period of the stripes. Although one might expect to find some optimal value for this parameter, our numerical simulations show that the quality of the resulting potential stays unaffected in a wide range of parameters (see fig. 2.3B). Surprisingly, even when the stripe width becomes very small (less than 2 pixels per period), the image is not affected too much. In such a situation an additional grating emerges (see fig. 2.4), resulting in additional

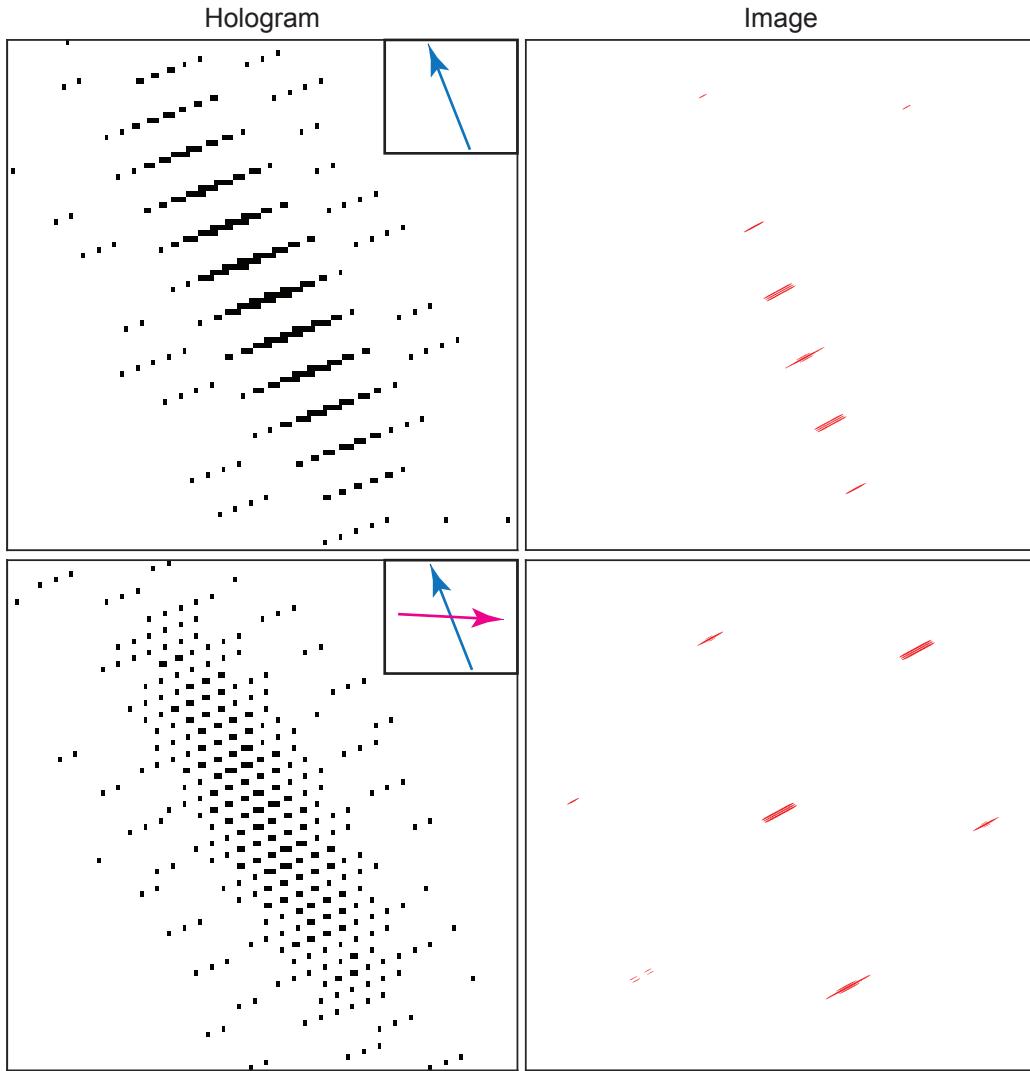


**Figure 2.3:** Potential quality as a function of grating angle. (A) RMS difference of the image plane potential created by a binary hologram compared to the gray-scaled one, for different values of smoothing parameter. Pronounced peaks at  $0$  and  $\frac{\pi}{2}$  correspond to the situation when the grating aligns to the DMD pixels that leads to aliasing. Smaller peaks at  $\frac{\pi}{4}$  and  $\frac{3\pi}{4}$  correspond to the grating that runs diagonally to the pixels, which also suffer from the same problem. The imbalance between the minor peaks ( $\frac{\pi}{4}$  and  $\frac{3\pi}{4}$ ) is due to the test pattern shape, which is elongated along the  $\sim \frac{\pi}{4}$  direction in the Fourier plane. Increasing the smoothening parameter decreases the sensitivity to the "bad" angles. However, it also decreases the overall quality of the resulting potential due to the inherent randomness of the method. The inset shows the smoothing function for the shown parameters. (B) RMS difference of the image plane potential for deterministic binarization method ( $\xi = 0$ ) as a function of grating period. Again for most of the values, the resulting pattern quality remains high. Interestingly the method works even when the period becomes smaller than 2 pixels.

images. However, those can be removed by spatial filtering.

We also try to determine the optimal value of the parameter  $\xi$ . To do that, we take the difference between two cuts that are separated by one lattice site through the resulting image plane profile. The introduction of the random parameter makes the resulting image plane profile vary slightly from realization to realization. Therefore, one can ask the question. What is the probability to obtain a pattern with the desired maximum difference between two cuts? Our numerical results suggest that the introduction of a small amount of randomness ( $\xi \sim 0.04$ ) can increase the probability of obtaining a desired potential quality compared to the strictly deterministic case.

Although more complex way to reduce the binarization artifacts like error diffusion algorithms<sup>??</sup> have been studied, their application to our system does not improve the results.



**Figure 2.4:** Binarization artifacts. On the right, examples of the binarized holograms for underlying grating period greater than two pixels (top) and lower than two pixels (bottom). The directions of the resulting wave vectors are shown in the inset. On the left, the resulting intensity pattern in the image plane. Multiple images along the direction of the wave vector correspond to the different orders of the grating. For the case of smaller grating period, the binarized hologram develops a secondary wave vector, leading to the formation of multiple additional image plane fetchers.

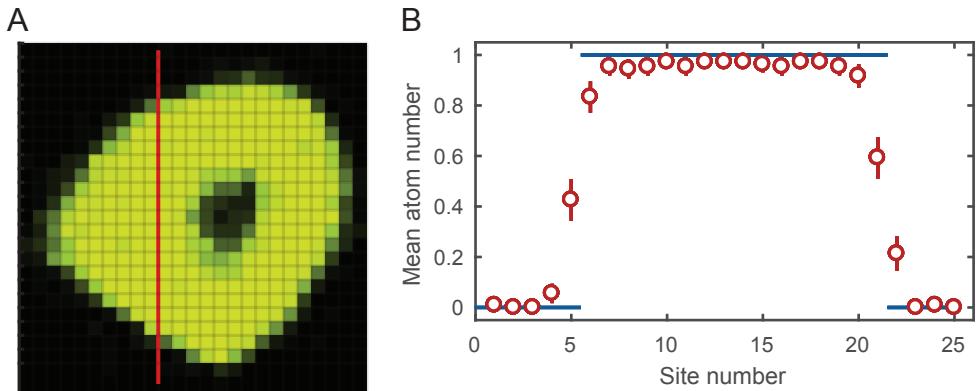
# 3

## Hight fidelity initialization and readout

### 3.1 INITIAL STATES WITH HIGH FIDELITY

Our goal is to study the dynamics of strongly-correlated many-body systems. Due to the probabilistic nature of quantum mechanics, in order to perform measurements on a quantum state, one needs to have access to multiple copies of the same state. Then, by doing repeated measurements, the probability of occupation of each particular basis state can be obtained. This means that we need a way to deterministically prepare the same state of the system with high fidelity.

In order to achieve a state of interest, one can start with some easy-to-prepare initial state, and then, by changing the parameters of the system, evolve it into the target state. For lattice systems one example of a conceptually simple initial state is a product state of the particles on indi-



**Figure 3.1:** Density distribution in the Mott insulator. A Averaged image of a Mott insulator in the harmonic trap. Green represents the region with single particle per site. The dark ring in the middle corresponds to the doubly occupied sites, where particles are lost during imaging process due to light-assisted collisions. B The density distribution of particles along the red line in A. Due to the finite temperature of the system, the sharp zero-temperature distribution (blue line) is smeared by the particle fluctuations.

vidual lattice sites. For all experiments described below, we prepare strings of the desired length with a single atom on each site. For atoms in the optical lattices there are two ways to achieve that: one is a bottom-up approach, whereby using single-site addressing and control one can trap a single atom in a tightly focused optical tweezer. Then by using Raman cooling techniques, the atom's vibrational state can be cooled to almost coincide with the perfect ground state in all three dimensions<sup>??</sup>. Finally, by subsequent rearranging an array of such tweezer traps in order to eliminate defects, one can achieve deterministic strings of up to 51 atoms<sup>?</sup> or  $?\times?$  two-dimensional arrays with unity filling<sup>?</sup>. This platform has successfully enabled a study of spin models, using Rydberg blockade techniques<sup>?</sup>. However, the current state-of-the-art cooling fidelity does not allow for efficient preparation of a single-band Hubbard model with a large number of particles, due to the residual excited-state fraction scaling exponentially with the number of atoms.

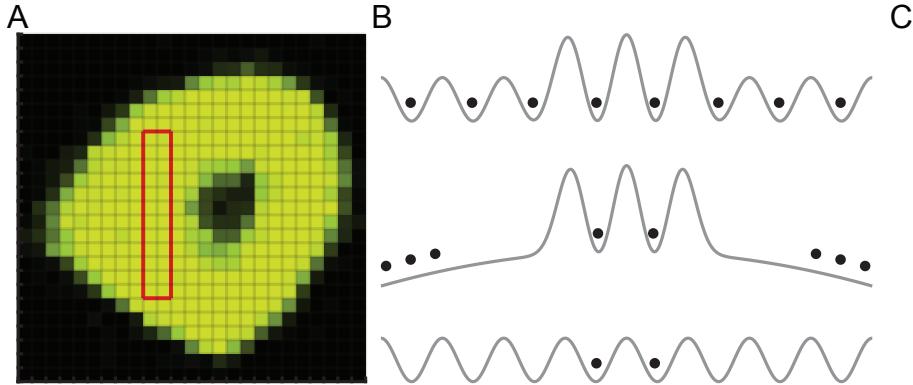
An alternative route is a top-down approach, where one uses an ensemble of atoms to first create a macroscopic number of atoms in the absolute ground state<sup>??</sup>, then selectively remove all excited state atoms<sup>?</sup>, and finally adiabatically load them into an optical lattice<sup>?</sup>. By ramping up the

lattice depth, the ratio between interaction and kinetic energies of the atoms can be increased. And at the critical ratio between those energy scales the system enters a Mott-insulating phase (see fig. 3.1A), where each lattice site is occupied by an integer number of atoms according to a local chemical potential, with a vanishing atom number variance<sup>??</sup>. However, since at the phase transition the gap between the ground and the excited state closes, it is impossible to cross the transition completely adiabatically in a finite amount of time<sup>?</sup>. Therefore at the end of the ramp the system has a small admixture of low-lying excited states, resulting in deviations of the atom number distribution from the ground-state one by "smearing" the boundary between the shells with different occupation numbers (see fig ??B). However, if the Mott insulator is sufficiently large, we can find a region with nearly unity filling up to 12 sites long. Hence, to achieve our desired target state, we just need to isolate such a region and ensure that it doesn't get disturbed by the adjacent atoms during subsequent evolution.

### 3.2 CUTTING PROCEDURE

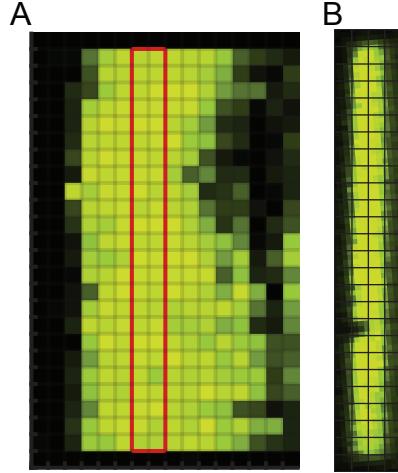
In order to isolate a desired region of the lattice, we perform a so-called "cutting" procedure. Using the DMD, we project a finite-size lattice potential on top of the initial 2D lattice, in which the Mott insulator was originally created. Then, by applying a slowly-varying repulsive beam and switching off the 2D lattice, we make the atoms leave, except for those that were kept by the DMD potential (see fig. 3.2B). Finally, we can reload the atoms into the 2D lattice by ramping it back up and switching off the DMD potential. By repeating this procedure along both directions, we isolate plaquettes up to  $12 \times 2$  in size (see fig. 3.2C). This technique allows us to achieve 99% single atom transfer fidelity from the initial Mott insulator. The preparation of larger systems could be achieved by increasing the optical power of the DMD beam. Similar results have been achieved using state-dependent light shifts and microwave pulses<sup>?</sup>.

By increasing the DMD beam power we could in principal initialize even longer chains. In



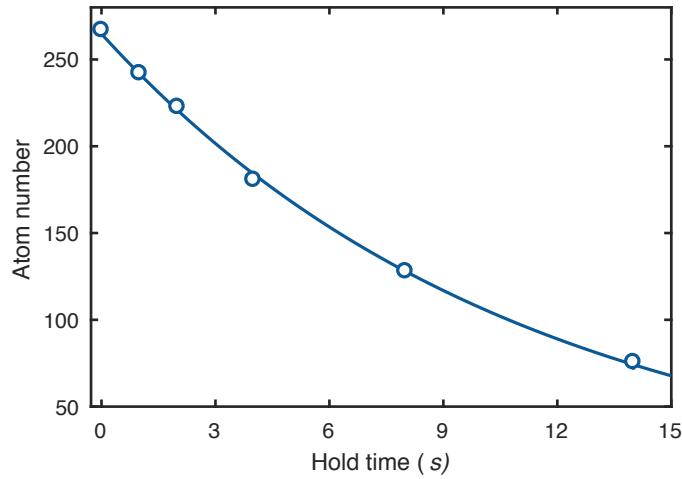
**Figure 3.2:** Cutting procedure. A Averaged image of a Mott insulator in the harmonic trap. B Using the DMD we project a potential that confines the desired number of lattice sites (in this case two) on top of the lattice, where the Mott insulator was originally created. By dropping the lattice and applying a slow-varying repulsive potential, centered on the DMD pattern, we expel all untrapped atoms. Finally the remaining atoms can be loaded back into the optical lattice. By applying this procedure along both direction we can prepare plaquettes of various sizes. C A single shot image of a  $2 \times 12$  plaquette cut from the region highlighted in A.

order to achieve a high fidelity of the  $n = 1$  shell of the Mott insulator, we rely on the idea of entropy redistribution in the system<sup>7</sup>. This concept can be applied to our system in the following way: crossing the transition from a Superfluid to a Mott insulator at finite time injects a certain amount of particle-hole excitations into the system. If the system were homogeneous, those excitations would spread evenly across the system. However, there is a strong inhomogeneity in the regions where shells of different occupation number touch one another. Since the density has to smoothly connect from one occupation to another, it deviates from an integer value creating regions of superfluid. A superfluid has a significantly smaller gap between low-lying states compared to a Mott shell, hence it is energetically favourable for the excitations to concentrate in the regions between the shells. From the above considerations, it is clear that this mechanism works better for a larger ratio of the superfluid region to the Mott shells. An analogous technique has also found a great success in Fermi gas microscopes, where a manual increase of the entropy-reservoir region led to significant entropy reductions in the gapped part of the system<sup>??</sup>.



**Figure 3.3: Mott Insulator in the box geometry.. A** Using DMD to provide the 24 sites-long box confinement in the  $y$  – direction and the magnetic field gradient in the  $x$  – direction we realize a Mott insulator in the box geometry. Average density distribution shows high fidelity  $n = 1$  shell. **B** A single shot image of the  $2 \times 24$  sites system with a single defect cut form the highlighted region in A.

From the analysis above, it follows that in order to create an  $n = 1$  shell with high fidelity, it is advantageous to have it surrounded by  $n = 0$  and  $n = 2$  shells. The current trapping geometry allows us to create chains of up to  $\sim 16$  sites long. For even longer chains, the single site fidelities become effected by the residual disorder of the lattice potential. One way to overcome this issue is to alter our trapping geometry. One simple solution that comes to mind is to make a rectangular box confinement instead of a harmonic one. In order to have well-defined shells, in this case, one can use a magnetic field gradient in order to create a uniform tilt. We realize such a configuration in our system in a 24 site long box (see fig. 3.3). Then, by using a cutting procedure along the other direction, we were able to achieve  $24 \times 2$  plaquettes. In this case, the length was only limited by the DMD beam power that provided the confinement during the cutting procedure. By increasing the DMD power and reducing the disorder, coming from the 2D lattice, this approach can enable the initialization of even longer chains of controlled length, which are particularly interesting for quantum simulations.



**Figure 3.4: Atoms life time.** Number of atoms in the harmonic trap as a function of time. Solid line corresponds to the exponential fit to the data with the  $\frac{1}{e}$  decay constant of  $21.4(7)$  seconds

### 3.3 ATOMS LIFE TIME

In our experiments, we are interested in studying coherent quantum many-body dynamics on sufficiently long timescales. In order to achieve that, we need to isolate our system from coupling to the environment. Working with neutral atoms gives us a head start, since they can only interact with the environment through collisions with photons or background atoms. Let's look at the later process in more detail.

In order to understand the effect of collisions with the background gas, one needs to consider the energy scales in the system. The typical depth of the optical potentials that we use to confine our atoms, is on the order of  $1\mu\text{K}$ . In contrast, since the vacuum chamber walls are at room temperature, the typical energy of the background gas is  $300\text{K}$ . Hence, a collision between an atom within our system with a background gas atom will lead to atom loss. This means that by measuring the atom loss rate in our system, we can determine the timescale corresponding to this process. To measure it, we prepare atoms in a shallow harmonic optical trap and image them after a variable amount of holding time (see fig. 3.4). We measure a single particle  $\frac{1}{e}$  decay time of

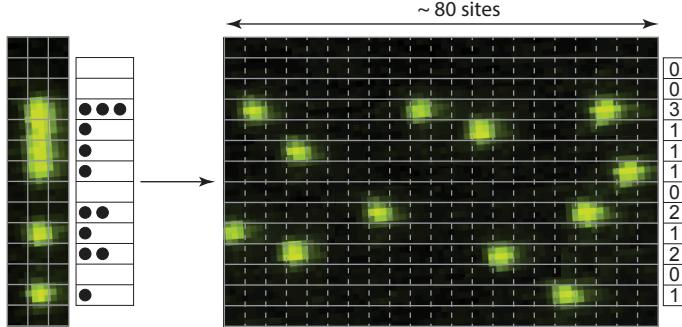
$21.4(7)$ s in our system. For 10 particles evolving for 1s this results in  $\sim 40\%$  probability of single-particle loss, which would lead to decoherence in the system.

However, if we work with systems of fixed particle number, the above source of decoherence can be overcome in our experiments. The high-fidelity initial state preparation allows us to start with a known number of atoms at the beginning of the experiment. If we could also determine the total number of atoms at the end of the sequence, we could discard the experimental runs in which the total atom number in the beginning and at the end are different. This procedure would eliminate the decoherence due to the atom loss at the expense of the number of successful runs of the experiment.

### 3.4 ATOM NUMBER RESOLVED STATE READOUT

One of the drawbacks of quantum-gas microscopes is the so-called parity projection during the imaging process. When subject to resonant light, two atoms undergo photon-assisted collisions, resulting in the formation of a molecule with a large kinetic energy. The molecules are not trapped by the optical lattice and get lost during this process. This means that in our image the sites with two particles on the same site appear dark like an empty site, and the sites with three particles appear to have one. Although this is not a problem in the Mott insulator regime, it leads to the loss of large amounts of information about the state in the superfluid regime.

We can utilize the direction transverse to the chain in order to achieve full number resolution for a one-dimensional system. The idea behind this method is simple: once the dynamics along the chain are frozen (in our case, we do this by rapidly increasing the lattice depth along the chain to  $45E_r$ ), we can lower the confinement transverse to the chain (which is provided by the other optical lattice in this case) to spread the atoms from every site of the chain into tubes of  $\sim 80$  lattice sites in the transverse direction. In this case, the probability of doubly occupying any



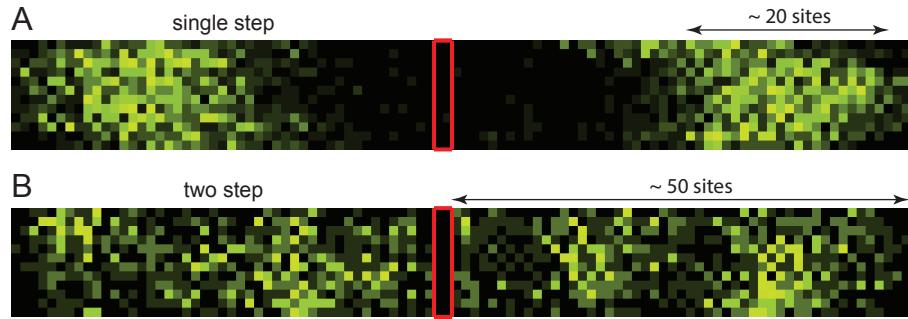
**Figure 3.5:** Atom number resolved readout. During the imaging process pairs of atoms are lost due to the light-assisted collisions, resulting in the parity projection of the single site occupations (left). For a 1D system we can use the lattice transverse to the chain to expand individual lattice sites into the region of  $\sim 80$  sites, such that the probability of multiple atoms occupying the same site is negligible. By counting the number of atoms in each tube we can reconstruct the atom number distribution along the chain.

site of the tube is given by

$$p(n, L) = \sum_{i=1}^{n-1} \frac{i}{L}, \quad (3.1)$$

where  $n$  is the number of particles that start on the site in the beginning and  $L$  is the length of the expansion tube in lattice sites. An example of such a process is shown in figure 3.5.

Although this technique is very powerful for isolated one-dimensional systems, we would like to develop a method that would allow us to perform counting of  $2 \times N$  plaquettes. At first, the solution seems very simple: one just needs a way to isolate the two sides of the plaquette and repeat the above procedure. Using the DMD we can project a single-site-wide Gaussian beam to disable the tunnelling between the two sides of the plaquette. However, that leads to an issue: the potential required to completely suppress the hopping between the two sides has a sufficient height at the position of the atoms, serving as an effective hill the atoms roll down from. This process imparts a momentum onto the atoms in the direction away from the barrier potential, which is large compared to the diffusion rate of the atoms. The result is that the atoms can only spread across  $\sim 20$  lattice sites before they leave the field of view of the microscope (see fig. 3.6 A). Such limited spread results in a significant probability of double occupancies.



**Figure 3.6:** Expansion of two adjacent lines. Resulting average density distribution of atoms after a single step (top) and a two step (bottom) expansion process starting from the highlighted region. For a single step process the atom's diffusion is slower than their center of mass momentum, hence the atoms expand much less as compared to the two step process, before they leave the field of view of the imaging system.

To overcome this issue, we add another step to this process. After a very short expansion time, when the atom distribution has moved by a few sites away from the barrier, and its effect can be neglected, we briefly capture the atoms back into the lattice and then release them again. This process randomizes the momentum of each atom and allows them to expand into the corresponding half tube with zero centres of mass momentum. This allows us to spread the occupation of a single site of the plaquette into  $\sim 50$  lattice sites, resulting in a 94% probability of successfully detecting 3 atoms or less.

# 4

## Long coherent dynamics in optical lattices

### 4.1 PURE VERSUS MIXED STATES

In order to understand what the purity of a state means and why it is important to be able to measure it, let's take a closer look at the difference between the quantum superposition and classically mixed states. While for both of them the measurements in most basis will lead to multiple different outcomes, for a pure quantum state there exists a basis in which a measurement will result in only one possible outcome. As an example, imagine that we have many copies of a spin  $\frac{1}{2}$  particle, and we can measure each of them along any of the axes  $x$ ,  $y$ , or  $z$ . If we had prepared half of the particles in a state  $|\uparrow\rangle$  and the other half in a state  $|\downarrow\rangle$  along the  $z$  direction, the measurement in any of the bases will result in the 50/50 split between the outcomes. Hence we will con-

clude that the state was mixed. Now, if the particles were initially prepared in the superposition state  $(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ , the measurement along  $y$  or  $z$  would yield the same result. However, if we measure this state in the  $x$  direction, we find that it produces only one of two possible outcomes.

In terms of formal mathematical notations, it is useful to introduce a concept of a density matrix. Given a state  $|\psi\rangle$ , the density matrix is defined as an outer product of the state with itself and can be formally written as  $\rho_{\text{pure}} = |\psi\rangle \langle \psi|$ . It is now easy to generalize this concept to include mixed states; they are just an incoherent sum of different pure states that occur with certain probabilities and can formally be written as  $\rho_{\text{mixed}} = \sum_i c_i |\psi_i\rangle \langle \psi_i|$ , where  $c_i$  is the probability of state  $|\psi_i\rangle$  so that  $\sum_i c_i = 1$ . This gives us a very intuitive way of quantifying the purity of a given state by looking at the trace of the square of the density matrix. In the case of a pure state

$$\text{Tr}(\rho_{\text{pure}}^2) = \text{Tr}(|\psi\rangle \langle \psi| |\psi\rangle \langle \psi|) = \text{Tr}(|\psi\rangle \langle \psi|) = \text{Tr}(\rho_{\text{pure}}) = 1. \quad (4.1)$$

On the other hand, if the state is mixed

$$\text{Tr}(\rho_{\text{mixed}}^2) = \text{Tr}\left(\sum_i c_i |\psi_i\rangle \langle \psi_i| \sum_j c_j |\psi_j\rangle \langle \psi_j|\right) = \text{Tr}\left(\sum_i c_i^2 |\psi_i\rangle \langle \psi_i|\right) = \sum_i c_i^2, \quad (4.2)$$

and since there should be at least two non zero terms in this sum  $\text{Tr}(\rho_{\text{mixed}}^2) < 1$ . This means that by measuring the trace of the density matrix squared, we can quantify the purity of the state.

Now let's consider what happens to the purity of a state, described by a density matrix  $\rho$ , during coherent time evolution. The time evolved density matrix can be written as  $\rho(t) = U\rho(0)U^\dagger$ , where  $U$  is the unitary evolution operator, and hence

$$\begin{aligned} \text{Tr}(\rho^2(t)) &= \text{Tr}(U\rho(0)U^\dagger U\rho(0)U^\dagger) = \text{Tr}(U\rho(0)\rho(0)U^\dagger) = \text{Tr}(U^\dagger U\rho(0)\rho(0)) = \text{Tr}(\rho(0)^2). \end{aligned} \quad (4.3)$$

This means that unitary evolution preserves the purity of the initial state. Since any coupling to

the environment will result in a loss of purity, this also gives us a way to quantify how coherence of the system's dynamics is.

## 4.2 HONG-OU-MANDEL EFFECT

In the next few sections, we will outline the procedure, that allows us to experimentally measure the purity of a quantum state. A more in-depth explanation can be found here<sup>7</sup>. In order to understand how this method works, it is instructive to consider the physics underlying the Hong-Ou-Mandel effect<sup>7</sup>. The set up is the following: Consider two photons that are incident onto two different input ports of a perfect 50/50 beamsplitter. Furthermore, let's suppose that they are identical to one another in every aspect, except for some parameter  $\lambda$ , which can be anything from photon frequency to the exact arrival time onto the beamsplitter (for simplicity let's assume that they have different wavelengths  $\lambda_1$  and  $\lambda_2$ ).

Mathematically, the beamsplitter transformation can be written in terms of creation and annihilation operators of the input modes  $i_1, i_2$  and the output modes  $o_1, o_2$  as

$$\begin{aligned} a_{i_1}^\dagger &\rightarrow (a_{o_1}^\dagger + ia_{o_2}^\dagger)/\sqrt{2} \\ a_{i_2}^\dagger &\rightarrow (ia_{o_1}^\dagger + a_{o_2}^\dagger)/\sqrt{2}. \end{aligned} \tag{4.4}$$

When two photons are incident onto two different ports of the beamsplitter, the resulting state can be found as following:

$$\begin{aligned} a_{i_1}^\dagger(\lambda_1)a_{i_2}^\dagger(\lambda_2)|o\rangle &\rightarrow \frac{1}{2}(a_{o_1}^\dagger(\lambda_1) + ia_{o_2}^\dagger(\lambda_1))(ia_{o_1}^\dagger(\lambda_2) + a_{o_2}^\dagger(\lambda_2))|o\rangle = \\ &= \frac{1}{2}(ia_{o_1}^\dagger(\lambda_1)a_{o_1}^\dagger(\lambda_2) + ia_{o_2}^\dagger(\lambda_1)a_{o_2}^\dagger(\lambda_1) + a_{o_1}^\dagger(\lambda_1)a_{o_2}^\dagger(\lambda_2) - a_{o_2}^\dagger(\lambda_1)a_{o_1}^\dagger(\lambda_2))|o\rangle, \end{aligned} \tag{4.5}$$

where  $|o\rangle$  corresponds to a vacuum state. We can see that, in general, the beamsplitter results in 3 possible states:  $|2, 0\rangle, |0, 2\rangle$  and  $|1, 1\rangle$ , where  $|m, n\rangle$  represents a state with  $m$  photons in the

output port one and  $n$  photons in the output port two of the beamsplitter. However, if the photon frequencies are identical, there is no way to determine the correspondence between the photons on the output port with those at the input and they become fundamentally indistinguishable. This is where the quantum nature of these particles comes into play. In this case the two last terms in the second line of equation 4.5 cancel each other, and the state  $|1, 1\rangle$  does not occur. This effect can be summarized as following: if two identical photons are incident onto a perfect 50/50 beamsplitter the probability of simultaneously detecting a photon at each of the output ports equals zero. However, if the photons can be distinguished in some way, the probability of simultaneous detection becomes non-zero.

### 4.3 HONG-OU-MANDEL INTERFERENCE OF MASSIVE PARTICLES

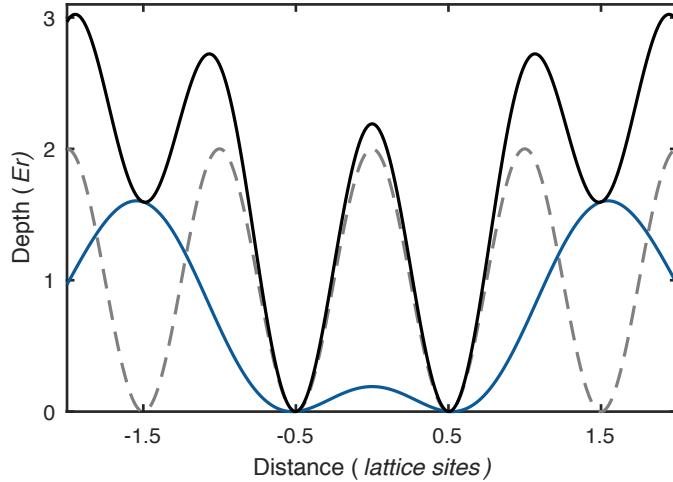
This effect can be also realized with massive bosonic particles and has first been observed in optical tweezers<sup>7</sup> and momentum states of helium BECs<sup>7</sup>. One can understand it in the following way: Consider a double-well with a coupling strength  $J$ . The corresponding Hamiltonian reads

$$H_{DW} = -J(a_L^\dagger a_R + a_R^\dagger a_L), \quad (4.6)$$

where  $a_{L(R)}^\dagger$  is bosonic creation operator for the left (right) site of the double-well. In the Heisenberg picture, this operators evolve under Hamiltonian dynamics, such that after an evolution time  $t = \frac{2\pi}{8J}$  the creation operators for the corresponding wells become

$$\begin{aligned} a_L^\dagger &\rightarrow (a_L^\dagger + ia_R^\dagger)/\sqrt{2} \\ a_R^\dagger &\rightarrow (ia_L^\dagger + a_R^\dagger)/\sqrt{2}. \end{aligned} \quad (4.7)$$

Notice that if we define the left (right) well before the time evolution as input  $i_1$  ( $i_2$ ) and the same wells after the evolution as  $o_1$  ( $o_2$ ), this transformation becomes exactly identical to the beam-



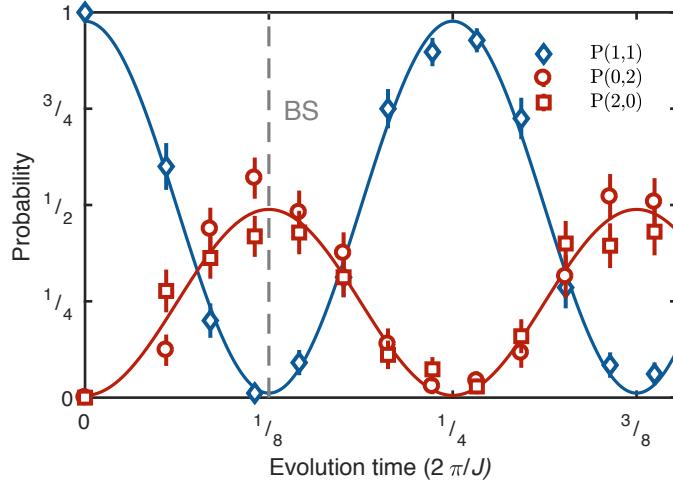
**Figure 4.1:** Double-well potential for the beamsplitter operation. The intensity profile of the DMD pattern (blue line) is combined with a shallow  $2E_r$  lattice (gray dashed line) to create the double-well potential. Such configuration allows for isolated tunneling between two adjacent sites. Coupling to the neighboring site is strongly suppressed due to the potential offset of  $1.7E_r$ .

splitter one in equation 4.5. Hence, if we start with a single particle in each well, after the time evolution corresponding to the beamsplitter duration the probability of finding one particle in each well should be equal to zero.

To observe this effect in our experiment, we prepare two atoms in the neighboring wells of our optical lattice. Using the DMD we isolate those wells by applying a repulsive potential, that detunes sites next to the double-well off-resonance for tunneling (see fig 4.1). The exact shape of the DMD pattern is given by

$$V_{DW}(x) = \left( e^{-\frac{(x-1.5)^2}{\sigma_1^2}} - ae^{-\frac{(x)^2}{\sigma_2^2}} + e^{-\frac{(x+1.5)^2}{\sigma_1^2}} \right)^2, \quad (4.8)$$

where  $x$  is measured in lattice sites. The values  $\sigma_1 = 0.95$ ,  $\sigma_2 = 0.9$ , and a relative amplitude of the middle bump  $a = 0.52$  are chosen such that the offset between two minima of the resulting potential is linearly insensitive to a relative alignment of the DMD pattern with respect to the lattice, by having quadratic minimas at the location of the lattice sites. For our experiments, we



**Figure 4.2:** Two particle osculations in the double-well potential. The probability of the system to be in each of the 3 available states as a function of time starting from initial state of one particle in each of the wells of the double-well. At the beamsplitter point (dashed line) the atoms undergo Hong-Ou-Mandel interference, where the probability of being in the  $|1, 1\rangle$  state goes to zero, while the other two states are populated equally. The solid lines show the fit to the data with residual probability  $P(1, 1) = 0.02$  at the beamsplitter point.

superimpose a  $2E_r$  deep optical lattice with a  $1.7E_r$  deep DMD potential, resulting in the tunneling rate  $J = 2\pi \times 128\text{Hz}$ .

At the beginning of the experiment, atoms are held in a deep ( $45\text{Er}$ ) lattice that suppresses the tunneling. To enable the dynamics, we rapidly change the lattice depth to  $2\text{Er}$  in  $0.5\text{ms}$ . The state of the double-well with two particles can be described in the basis of three configurational states:  $|1, 1\rangle$ ,  $|0, 2\rangle$ , and  $|2, 0\rangle$ , each corresponding to different arrangements of atoms between the two wells. In our experiment, we can measure the probability of the system occupying each of those states  $P(1, 1)$ ,  $P(0, 2)$ , and  $P(2, 0)$  as a function of evolution time by using the previously described atom number resolved state readout procedure (see fig. 4.2). As we expect, after an evolution time of  $t = \frac{2\pi}{8J}$  the probabilities of the state  $|1, 1\rangle$  almost go to zero, whereas the probability of the states  $|0, 2\rangle$  and  $|2, 0\rangle$  are equal to each other.

In our experiment, we observe a residual  $\sim 2\%$  probability of the  $|1, 1\rangle$  state for a beamsplitter

operation. This can be explained by the residual atom-atom interactions of  $U = 68\text{Hz}$  and an offset between the well of  $\Delta = 20\text{Hz}$  in our system. We can determine these two parameters independently.

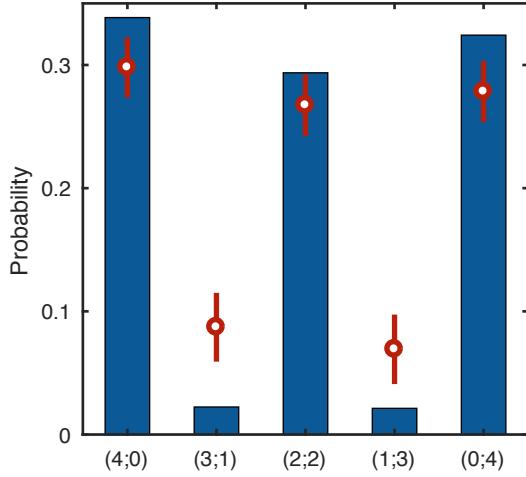
#### 4.4 MULTI-PARTICLE CASE

The Hong-Ou-Mandel effect can be also extended to the initial states with more than one particle. If we start with two identical  $n$ -particle bosonic states, each at one of the input ports of the beam-splitter, the outcome would contain only states with an even number of particles in each of the output ports (this will be explained in the following sections). In particular, if we start with two particles in each of the input ports, the beamsplitter operation should result in

$$|2, 2\rangle \rightarrow \sqrt{\frac{3}{8}}(|4, 0\rangle + |0, 4\rangle) + \frac{1}{2}|2, 2\rangle. \quad (4.9)$$

The result of such an operation in our experiment is shown in the figure 4.3. In this case, our measured probabilities of the output states deviate from the theoretical prediction, which includes finite interaction strength and the offset between two wells. However, the discrepancy can be explained by the interaction assisted coupling to the higher bands in the axial direction of the lattice, which has relatively low trapping frequency.

The interactions between particles within the same well modify the wave functions of individual particles, by making them wider compared to the non-interacting case. This, in turn, projects the single particle wave function onto the higher vibrational states of the harmonic trap in the axial direction, which makes the particles distinguishable between one another. To estimate this effect we assume that the exited band population of a single particle is proportional to  $\epsilon \approx (\frac{U(n-1)}{2\omega_z})^2$ , where  $n$  is the number of particles on one site and  $\omega_z$  is the trap frequency in the weak direction. In the case of 4 particles on the same site  $\epsilon \sim 0.017$ , which should result in



**Figure 4.3: Four particle Hong-Ou-Mandel interference.** The probability of the system being in one of the five possible states at the beamsplitter point, starting with two particles in each well of the double-well. The bars show the prediction of exact digitalization for the experimental parameter regime of  $\frac{U}{J} = 0.53$  and relative offset of  $\frac{\Delta}{J} = 0.15$ .

$\sim 6.5\%$  probability of odd outcomes. Combined with  $\sim 4.3\%$  of odd outcomes due to the finite interaction strength within the single band model leads to the total of  $\sim 10.8\%$  probability of odd outcomes, which is within the errorbar of our measurement (see fig. 4.3).

#### 4.5 MANY-BODY INTERFERENCE AS A MEASURE OF INDISTINGUISHABILITY

In order to generalize the Hong-Ou-Mandel idea to arbitrary, bosonic many-particle states, let's consider a system consisting of two states  $|\psi\rangle_1$  and  $|\varphi\rangle_2$ . The full wave function  $|\psi\rangle_1 \otimes |\varphi\rangle_2$  can be written in terms of creation operators  $a_{1(2)}^\dagger$  acting on the subspaces 1 (2) as

$$|\psi\rangle_1 \otimes |\varphi\rangle_2 = \frac{1}{(\sqrt{2})^{p+q}} (a_1^\dagger + a_2^\dagger)^p (a_1^\dagger - a_2^\dagger)^q |0\rangle, \quad (4.10)$$

where  $|0\rangle$  is a vacuum state.

If two states  $|\psi\rangle_1$  and  $|\psi\rangle_2$  are identical the full state  $|\psi\rangle_1 \otimes |\psi\rangle_2$  should be symmetric under

the exchange of particles between 1 and 2 ( $a_1^\dagger \leftrightarrow a_2^\dagger$ ), due to the symmetry of the bosonic wave function. Hence, the antisymmetric part of the wave function should be raised to an even power

$$|\psi\rangle_1 \otimes |\psi\rangle_2 = \frac{1}{(\sqrt{2})^{p+q}} (a_1^\dagger + a_2^\dagger)^p (a_1^\dagger - a_2^\dagger)^{2m} |o\rangle. \quad (4.11)$$

Notice, that by applying a unitary transformation given by

$$\begin{aligned} (a_1^\dagger + a_2^\dagger)/\sqrt{2} &\xrightarrow{\text{DFT}} a_1^\dagger \\ (a_1^\dagger - a_2^\dagger)/\sqrt{2} &\xrightarrow{\text{DFT}} a_2^\dagger, \end{aligned} \quad (4.12)$$

we will transform our state into  $|\psi\rangle_1 \otimes |\psi\rangle_2 \xrightarrow{\text{DFT}} (a_1^\dagger)^p (a_2^\dagger)^{2m} |o\rangle$ . This means that application of such transformation to two identical states results in the even number of particles in the corresponding subspace. Therefore, it can serve as a probe of indistinguishability between two quantum states. The transformation described by the equation 4.12 corresponds a two point version of the discrete Fourier transform (DFT), and the results in the following sections can be generalized to higher order correlation functions by considering multi-point DFTs ??.

The DFT transformation is very similar to the beamsplitter operation given by equation 4.5. In fact, the beamsplitter can be reduced to the DFT by applying local  $\frac{\pi}{2}$  phase shifts on one of the input ports of the beamsplitter. However, if two states have no defined phase relations, such as fixed particle number states, both operations are equivalent to one another. That's why applying the beamsplitter to the photon number states allows us to tell if the states were identical or not. In our experiment, we always work with fixed particle number states, which allows us to use the beamsplitter transformation instead of the DFT as well.

## 4.6 MEASURING THE PURITY OF A STATE

To understand how the DFT allows us to measure the purity of a given state, let's introduce one more useful concept called the *SWAP* operator. It is defined as following: Consider two states  $|\psi\rangle_1$  and  $|\varphi\rangle_2$ , *SWAP* acting on a tensor product of the two results in

$$\text{SWAP}(|\psi\rangle_1 \otimes |\varphi\rangle_2) = |\varphi\rangle_1 \otimes |\psi\rangle_2. \quad (4.13)$$

Since applying *SWAP* twice returns the state to itself, the eigenvalues of the operator are  $\pm 1$ , and the corresponding eigenstates can be written as  $(a_1^\dagger \pm a_2^\dagger)/\sqrt{2}$  in terms of creation operators in each subspace. As we have shown above, DFT will map the antisymmetric eigenstate onto a creation operator in the corresponding subspace (see eq. 4.12). Thus, by measuring the particle number in that subspace after the DFT, we can determine the eigenvalue of the *SWAP* operator (even numbers corresponding to  $+1$  and odd to  $-1$ ). This means, that if we would like to measure the expectation value of the *SWAP* operator with respect to a product state of two density matrices  $\rho^{(1)} \otimes \rho^{(2)}$ , we can do it by measuring the atom number parity of the appropriate subspace after performing the DFT

$$\text{Tr}(\text{SWAP}(\rho^{(1)} \otimes \rho^{(2)})) \xrightarrow{\text{DFT}} \langle P_2 \rangle, \quad (4.14)$$

where  $\langle P_2 \rangle$  is average parity in the second subspace.

The final step is to connect the expectation value of the *SWAP* to the quantum state overlap ??, which for two states described by the density matrices  $\rho^{(1)}$  and  $\rho^{(2)}$  is given by  $\text{Tr}(\rho^{(1)} \rho^{(2)})$ . To

see this, consider the following:

$$\begin{aligned}
Tr(SWAP(\rho^{(1)} \otimes \rho^{(2)})) &= Tr(SWAP(\rho_{ij}^{(1)} \rho_{kl}^{(2)} |i\rangle \langle j| \otimes |k\rangle \langle l|)) = Tr(\rho_{ij}^{(1)} \rho_{kl}^{(2)} |k\rangle \langle j| \otimes |i\rangle \langle l|) = \\
&= \sum_{ijkl} \rho_{ij}^{(1)} \rho_{kl}^{(2)} \delta_{kj} \delta_{il} = \sum_{ik} \rho_{ik}^{(1)} \rho_{ki}^{(2)} = Tr(\rho^{(1)} \rho^{(2)}).
\end{aligned} \tag{4.15}$$

In particular, if  $\rho^{(1)} = \rho^{(2)} = \rho$ , then the state overlapped with itself leads to the purity ??

$$Tr(\rho^{(1)} \rho^{(2)}) = Tr(\rho^2).$$

Let's summarize the conclusions above: In order to measure the purity of a state described by the density matrix  $\rho$ , we need to prepare two identical copies of the same state, perform DFT on them, and finally measure average particle parity in one of the subspaces

$$Tr(\rho^2) = Tr(\rho\rho) = Tr(SWAP(\rho \otimes \rho)) \xrightarrow{\text{DFT}} \langle P_2 \rangle. \tag{4.16}$$

Moreover, for the states without a defined phase relation (such as fixed particle number states) any unitary transform that splits the particles equally between two subspaces can play the role of DFT (for instance a beamsplitter operation).

#### 4.7 MULTI-MODE SYSTEMS

So far we have only discussed the situation where the state of interest occupies a single bosonic mode (i.e. can be written as a polynomial function of a single creation operator  $a^\dagger$  acting on the vacuum). However, in general, the states can occupy multiple modes at the same time. For example, consider a bosonic particle on a lattice that is delocalized over all sites of the lattice. Its state, up to a normalization, can be written as  $|\psi\rangle = \sum_i a_i^\dagger |0\rangle$ , where  $a_i^\dagger$  is a creation operator on the  $i$ -th site of the lattice. This means that in the scheme described above we have to perform the DFT on each mode available to the system (i.e. in this case on each every lattice site). For our experi-

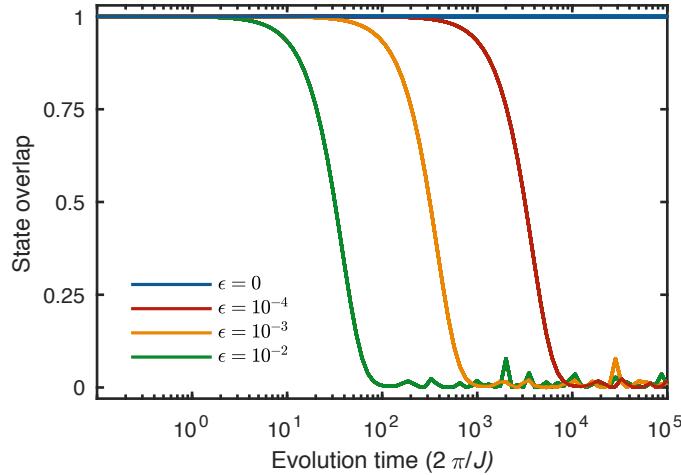
mental implementation, it means that we have to perform a beamsplitter operation between the corresponding sites of each of the copies. In order to compute the overall parity of a given realization, we can simply take the product of the parities of individual modes  $\langle P_{full} \rangle = \langle \prod_i P_i \rangle$ , where  $i$  runs through all modes of the system (see fig. ??).

The multi-mode structure of a state additionally enables us to measure the purity of the subsystem, which can be defined as a certain subset of modes within the state. By averaging the parity only among those  $\langle P_{sub} \rangle = \langle \prod_{i \in sub} P_i \rangle$  we will obtain the purity of the given subsystem. Importantly, we can obtain the purity of any subsystem through data processing, by simply blinding ourselves to the outcomes of the beamsplitters that do not belong to the subsystem of interest. This will be very important later on when we will discuss the measurement of entanglement within the system.

#### 4.8 DECOHERENCE BETWEEN COPIES

Although being a powerful tool to measure the purity of the system, many-body interference has one crucial limitation. Since the method measures the state overlap between two copies, it results in the purity only if two states are exactly identical. This task becomes even more challenging if we are interested in dynamical measurements because we have to ensure that both the initial states of the two copies as well as Hamiltonians under which each copy is evolving are identical to each other.

In order to understand these effects, we perform numerical studies in the parameter regime corresponding to our experiment. We consider the scenario where two system starts in the same initial state but evolve under slightly different Hamiltonians. For these simulations, we work with one-dimensional 8 site-long Bose-Hubbard chain with  $U/J \sim 1$ . The initial state is a charge-density wave at half filling. To make the Hamiltonians between two copies different the tunnelling strength in the second one is set to be  $J_2 = J_1(1 + \epsilon)$ , where  $\epsilon \ll 1$ . The tunnelling strength

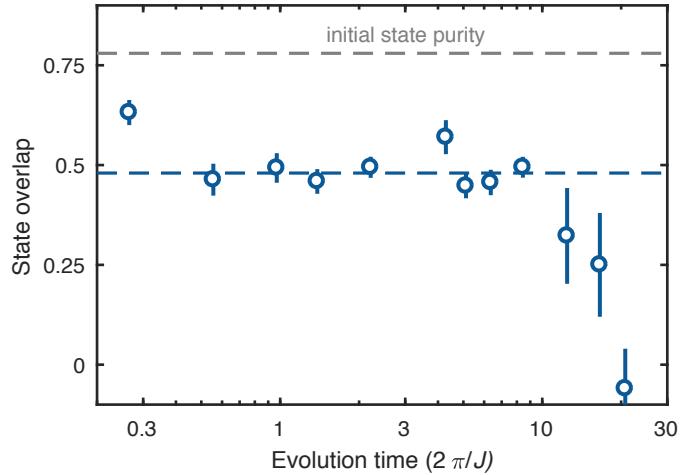


**Figure 4.4:** Dephasing of two many-body states. Quantum state overlap  $\text{Tr}(\rho_1 \rho_2)$  between two states that evolve under slightly different Hamiltonians, for which the difference in tunneling strength is parametrized by the parameter  $\epsilon = \frac{\Delta J}{J}$ . The overlap remains near unity until the time given by  $\tau = 10^{\log(\epsilon)-1}$ , when it decays to zero after  $\sim 1$  decade of time evolution.

also provides a natural time scale for the problem given by  $\frac{2\pi}{J}$ .

As we expect, if  $\epsilon = 0$  then two copies remain identical throughout the whole evolution time, resulting in the state overlap  $\text{Tr}(\rho_1 \rho_2) = 1$  (see fig. 4.4). However, if  $\epsilon > 0$  state overlap remains at 1 up to some time scale, but then decays to zero over about one decade of time evolution. From the numerical data, we can identify that the deviation from unity overlap occurs at a timescale given by  $\tau = 10^{\log(\epsilon)-1}$ . From this analysis, we can conclude that, as long as we look at the timescales shorter than  $\tau$ , many-body interference results in the purity of the state. However, it also imposes a challenge for the experiments, if we would like to do measurements on very long timescales.

In order to characterize our system, we perform the following experiment (see fig. ??): First, we prepare two copies of an unentangled product state, consisting of a chain of 6 atoms in the  $45E_r$  deep lattice in both directions. Next, we confine our system to the desired size (for this experiment either 6 or 12 sites long), by detuning the sites outside the region, in a similar way as



**Figure 4.5: Many-body coherence between two copies.** Quantum state overlap  $Tr(\rho_1, \rho_2)$  between two six sites chains in our experiment shows a steady state between  $0.5 - 10$  tunnelling times. The reduction of purity is consistent with non-perfect beamsplitter operation, therefore, we conclude that two copies remain coherent up to  $\sim 10$  tunnelling times.

for the beamsplitter experiments. Then, we rapidly drop the lattice along the chains, allowing atoms to tunnel. After a variable evolution time, we suddenly freeze the dynamics by increasing the lattice depth. Till this point, the lattice between the copies stays high such that the two states evolve independently from each other. Finally, we perform the beamsplitter operation between two copies, by dropping the lattice between them and allowing for tunnelling corresponding to the beamsplitter time. At the end of the experiment, we perform atom-number-resolved readout of the whole system.

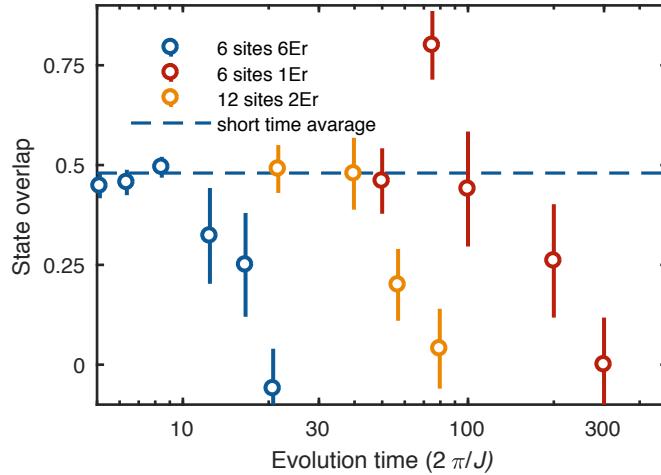
The initial state of the system is a product state of a single atom on every site of the chain in each copy. Hence, the application of a many-body beamsplitter is equivalent to the product of 6 individual beamsplitters with one particle in each well of the double-well. In this case, each beamsplitter operation leads to a HOM interference, resulting in an average purity of 78.3(1)% for the full state, limited by our beamsplitter imperfections (see fig. 4.5). Once we switch on the tunnelling along the chains, the purity of the state decreases until it reaches a stationary value

between 0.4 and 9 tunnelling times. Starting at  $\sim 10$  tunnelling times the purity drops from this stationary value all the way to zero. We attribute this decay to the decoherence between two copies.

The initial decay of purity is attributed to the fact that the system is evolving to a state with high single-site atom-number occupation. As was discussed in the previous sections, our HOM contrast is decreased for states with higher on-site atom-number, leading to an increased likelihood of non-perfect interference. We can estimate this effect in the following way: In the stationary regime, the on-site number distribution can be approximated as having 35%, 40% and 25% probability of having zero, one and two particles respectively. The beamsplitter results in the purity of 1, 0.96 and 0.7 for the respective values. Computing the result of the full beamsplitter operation on 6 sites with the above distribution results in an overall purity of 0.53, which is very close to the measured value.

#### 4.9 INCREASING THE COHERENCE OF THE SYSTEM

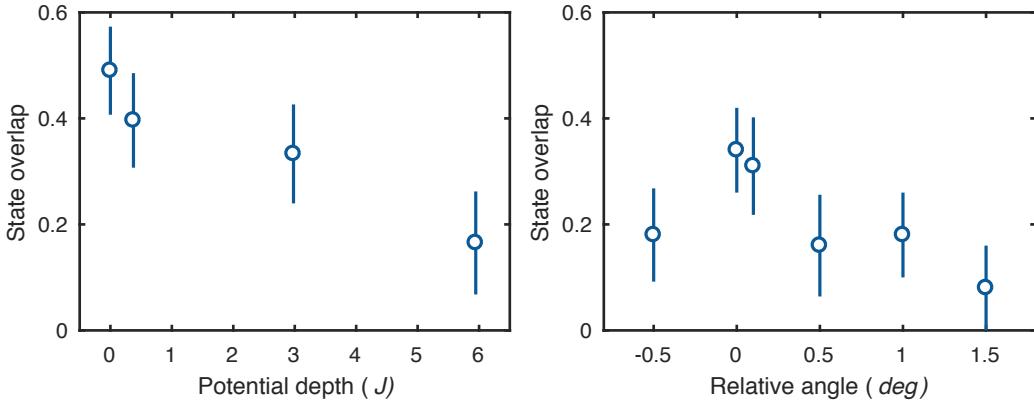
In order to improve the coherence in the systems, we can do either of two things: make the Hamiltonians in both copies more identical to one another or decrease the characteristic timescale of the experiment given by the tunnelling strength. The difference between two Hamiltonians comes from the imperfections in our optical lattice potential that occur due to the interference of the lattice beams with the stray light. Therefore, the amount of disorder in absolute units decreases with the light intensity and hence the lattice depth. On the other hand, the tunnelling strength is increased in a shallow lattice. This means that if we perform our experiments in a shallow lattice we expect to increase the coherence time in our system. We indeed see that the coherence time increases from  $\sim 10$  to  $\sim 100$  tunneling times, when we go from  $6E_r$  to  $1E_r$  deep lattice (see fig. 4.6). The decoherence timescale is consistent with the residual difference in tunnelling strength, that was separately estimated by measuring the trap frequency of individual lattice



**Figure 4.6: Long coherent dynamics in the system.** Quantum state overlap  $\text{Tr}(\rho_1, \rho_2)$  for a system of 6 atoms and various experimental parameters as a function of evolution time. Reducing the depth of the lattice along the chains increases the tunneling strength and decreases the differential disorder between two copies, resulting in increased the coherence time. The longest observed coherence of  $\sim 100$  tunnelling times shows an exquisite control over our experimental system.

sites of each copy.

Since the quench populates a large fraction of the full Hilbert space of the system, above experiments show that we can precisely control the evolution of the system in 462 dimensional Hilbert space over  $\sim 100$  tunnelling times. We can also increase the effective Hilbert space size by increasing the number of available sites to the particles. To do that, we again prepare each copy with a string of 6 atoms, but this time we confine them in the 12 site long box, such that the atoms start at its centre. This increases the Hilbert space dimension of the system to  $\sim 12000$  eigenstates. We observe, that we can maintain the coherence between such a large number of states up to  $\sim 40$  tunnelling times (see fig. 4.6). The decrease in the coherence time relative to the 6 site long system might be due to the larger accumulation of differential tilts between two copies.



**Figure 4.7:** Single site addressing with high precisions. A An additional single-site wide beam created by DMD is applied to system of 6 atoms on 12 sites, creating a potential offset on one site of each chain. Rotation of the pattern introduces a difference in the offsets experienced by each of the chains. B Quantum state overlap  $\text{Tr}(\rho_1 \rho_2)$  as a function of the applied potential depth, which is rotated by  $0.5^\circ$ . As the potential depth increases the state overlap decays due to the Hamiltonian difference between two chains. C By rotating the pattern we can restore the overlap, once the potential is properly aligned to the system.

#### 4.10 COHERENCE IN THE SYSTEMS WITH SINGLE SITE CONTROL

Purity measurements provide a very sensitive probe for the optical potentials in our system, due to its stringent requirement on how identical the Hamiltonians of the copies should be in order to maintain the coherence. In order to test the performance of our DMD, we use the following procedure: Before starting the dynamics of the system, we use DMD to create an offset on one of the middle sites of both copies. If the applied potential is different between two chains, it will lead to the loss of coherence time in the system.

We would like to determine the maximum height of the potential that can be applied before the copies start to decohere at 10 tunnelling times. Note, that this timescale sets the relative precision of two potentials at  $10^{-2}$  level. The potential we apply is a 2D Gaussian beam with a width of 1 lattice site along the chain and  $\sim 30$  sites in the transverse direction. Hence, we can intentionally make two copies slightly different by rotating the potential with respect to the copies. We start by applying a  $0.5^\circ$  rotation and increase the depth of the potential until we see the pu-

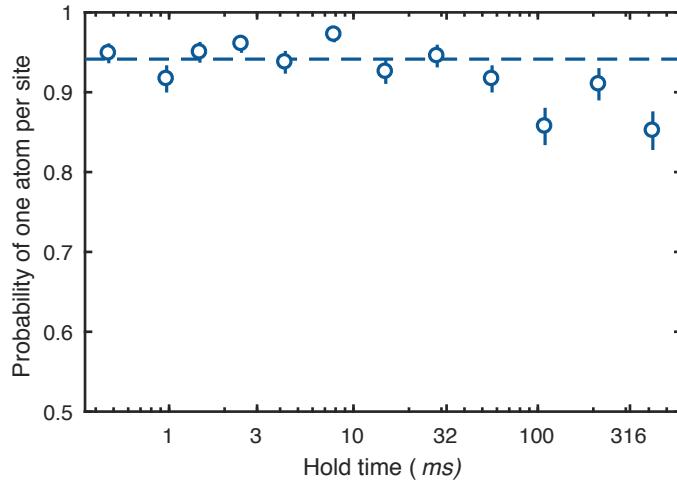
rity decrease at the depth of  $\sim 6J$  (see fig. 4.7 A). Then we fix the depth of the potential to  $6J$  and measure the state overlap as a function of the relative angle to the lattice (see fig. 4.7 B). As we expect, the state overlap has a peak when the relative angle is zero, with FWHM  $\sim 10$ , which is consistent with numerical simulations for this parameter regime.

#### 4.11 COHERENT EVOLUTION WITH A SINGLE COPY

In the previous sections, we have shown that we can prepare highly pure initial states and maintain their purity over two orders of time evolution. However, in order to achieve such outstanding performance we were restricted to work in a very shallow optical lattice, which can be an issue for certain types of experiments. The decoherence timescales observed with two copy method were consistent to be limited by the reduction in state overlap, due to the difference in the Hamiltonians between the copies. This means that our system might stay coherent over much longer timescales.

In any closed system, the decoherence of the state is always connected to some coupling of the system to the environment. Working with neutral atoms is a very good start since they only interact with the environment through collisions with photons or background atoms. As we have shown before, we can always overcome the latter by doing a post-selection on our state, and since we work with a far-detuned ( $\sim 20\text{nm}$ ) optical potentials the probability of spontaneously scattering a photon is very low. We estimate the overall scattering rate in our system to be  $\sim 0.012\text{Hz}$ , which suggest that we should be able to maintain coherent evolution of  $\sim 10$  particles over  $\sim 1$  second.

In order to probe the coherence in this case, instead of rapidly quenching the Hamiltonian parameters we adiabatically lower the lattice depth from deep in the Mott insulator phase into the superfluid phase, where atoms are allowed to tunnel. Since our Mott state is highly pure, it means that the system is occupying a single many-body state, in this case, it is a ground state of



**Figure 4.8:** Probing coherence with a superfluid. Probability of one particle on individual lattice sites averaged over the full chain of 8 atoms. The dashed line shows the early time average from 1 – 10ms. Up to the longest measured time all points lay above a single excitation threshold of 0.75.

the system. Consequently, if we do adiabatic ramps into the superfluid and back we should still stay in the ground state of the system. Hence, any decoherence should result in the formation of excitations, which deep in the Mott phase result in the formation of double-hole pairs. In our experiment, we can detect those, by doing atom number resolved readout of the final state.

In order to study the decoherence over time, we add a variable hold time in the superfluid regime. We observe that probability of one particle per lattice site stays constant over few hundred milliseconds (see fig. 4.8), which is consistent with our overall scattering rate. This timescale suggests that we have coherent many-body dynamics over  $\sim 500$  tunnelling times in a  $2E_r$  optical lattice.

# 5

## Thermalization in an isolated quantum system

Portions of this chapter have previously appeared in the following paper:

### 5.1 STATISTICAL MECHANICS: FROM CLASSICAL TO QUANTUM

Our understanding of classical statistical mechanics relies on the notion of ergodicity<sup>??</sup>, which states that during the time evolution a system explores entire phase space allowed by constraints. Time averaging plays an important role in this setup since the initial conditions of the classical non-chaotic system uniquely determine its state at any particular time during the evolution. Ul-

timately, time averaging is said to be equal to ensemble averaging<sup>??</sup>. It's important to note here, that ergodic hypothesis in such setting implies thermalization only in a *weak sense*, where weak refers to the fact that the statement is made only about *time-averaged* values of observables at long times. In order to obtain the thermalization in the *strong sense*, namely, the *instantaneous* values of observables at long times equal to the thermodynamic ensemble, one has to consider classical chaotic systems like Fermi-Ulam model<sup>?</sup>, the Kapitza pendulum<sup>?</sup>, and the kicked rotor<sup>?</sup>.

At this point, one can ask a question how to translate the classical ideas of thermalization to the quantum language, in particular, whether a quantum system can have thermalization in a strong sense? At first glance, it seems elusive, since the evolution of a quantum system is governed by the Shrödinger equation, which is linear and thus can not provide chaotic dynamics. However recent theoretical breakthroughs<sup>???</sup> have shown that quantum systems also exhibit thermalization. These ideas rely on the conjecture named eigenstate thermalization hypothesis (ETH)<sup>?</sup>, which states that in a chaotic quantum system individual eigenstates are thermal by nature. In order to understand the meaning of this statement, we first need to understand what exactly thermalizes in such systems.

In the early days of quantum mechanics von Neumann noted that, when talking about thermalization, one should focus on local observables rather than wave functions, who carry all global properties of the system (see fig. 5.1). This approach is very similar to a classical example of an isolated container with a gas, where we focus on a small volume inside the container to get a statistical ensemble for the subsystem. From this point of view, ETH can be formulated as following: in a quantum chaotic system an expectation value of a local observable is similar between the individual eigenstates and equals to statistical ensemble average. Putting it in more rigorous mathematical terms, ETH states that an expectation value of local observable  $O_{nm}$  between eigenstates

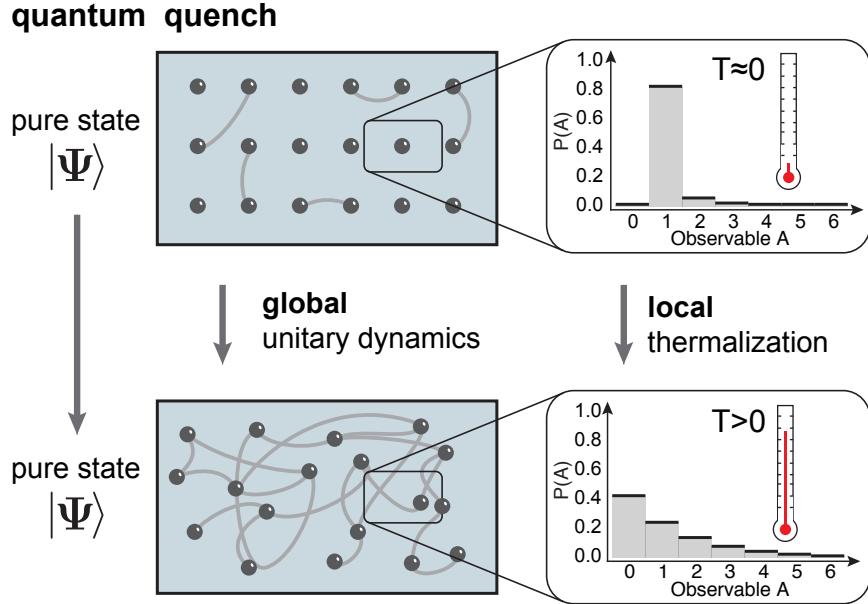


Figure 5.1: Schematic of thermalization dynamics in closed systems. An isolated quantum system at zero temperature can be described by a single pure wavefunction  $|\Psi\rangle$ . Subsystems of the full quantum state appear pure, as long as the entanglement (indicated by grey lines) between subsystems is negligible. If suddenly perturbed, the full system evolves unitarily, developing significant entanglement between all parts of the system. While the full system remains in a pure, and in this sense zero-entropy state, the entropy of entanglement causes the subsystems to equilibrate, and local, thermal mixed states appear to emerge within a globally pure quantum state.

$|n\rangle$  and  $|m\rangle$  is equal to<sup>7</sup>:

$$O_{nm} = O(\bar{E})\delta_{nm} + e^{-S(\bar{E})/2}f_O(\bar{E}, \omega)R_{nm}, \quad (5.1)$$

where  $\bar{E} \equiv (E_n + E_m)/2$ ,  $\omega \equiv E_n - E_m$  and  $S(E)$  is thermodynamic entropy at energy  $E$ . Crucially,  $O(E)$  and  $f_O(E, \omega)$  are smooth functions of their arguments, the value  $O(E)$  is identical to the expectation value of the microcanonical ensemble at energy  $E$  and  $R_{mn}$  is a random real or complex variable with zero mean and unit variance. One important remark should be made at this point: there are obvious exceptions from this rule, namely ground state and low lying eigenstates as well as the states on top of the spectrum. Taking this into account, we say that ETH holds for

the eigenstates in the middle of the spectrum.

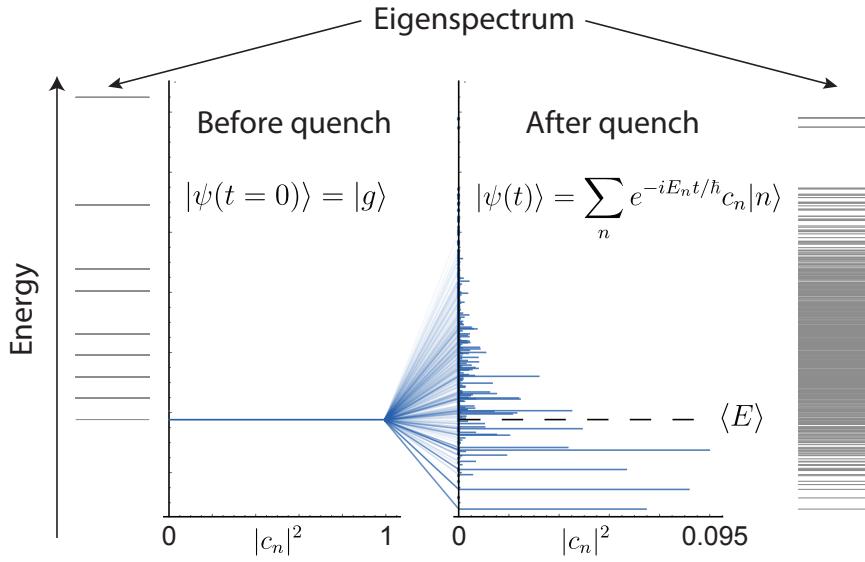
## 5.2 QUENCH DYNAMICS OF AN ISOLATED QUANTUM SYSTEM

Unfortunately, there is no known experimental technique to reliably prepare individual excited eigenstates of chaotic Hamiltonians. But since we are interested in dynamics associated with eigenstates in the middle of the spectrum we need to find a way to populate them. Here quantum quenches become handy. Quench is a sudden change in the Hamiltonian parameters, such that the wave function doesn't have time to adjust and, hence, gets reprojected onto the eigenstates of the new Hamiltonian (see fig.5.2). In our experiments, we always start in the ground state of the initial Hamiltonian. The energy of this state lies well above the ground state of the final Hamiltonian and the state itself has very small overlap with any of the final eigenstates. Hence, the initial state gets projected onto a large superposition of states primarily in the middle of the spectrum of the final Hamiltonian.

After the quench, the time evolution of the state can formally be written in the basis of final Hamiltonian eigenstates  $|n\rangle$  as

$$|\psi(t)\rangle = \sum_n e^{-iE_n t/\hbar} c_n |n\rangle, \quad (5.2)$$

where  $E_n$  and  $c_n$  are the corresponding energy and amplitude of the projection. Generally, this sum contains a large number of oscillating terms between the off-diagonal elements with large frequency, since the energy difference  $E_n - E_m$  between adjacent eigenstates can be arbitrarily small. However, ETH predicts that, for a sufficiently local observable and states with large entanglement, the contribution of the off-diagonal terms is exponentially small. Hence the expectation value of such observables is determined only by the projection amplitudes  $c_n$ , after some transient evolution time. In addition, the expectation value for thermal eigenstates is equal to the statis-



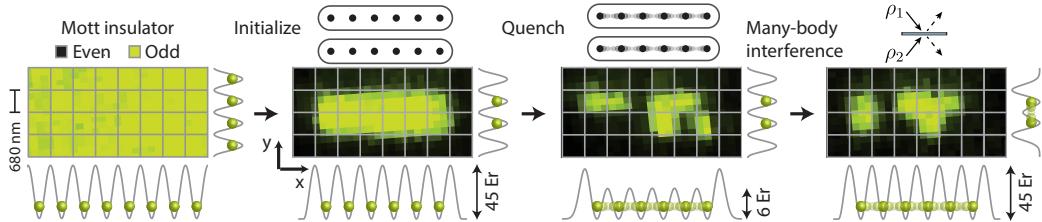
**Figure 5.2:** Local observables in a globally pure quenched state. In a quench, the ground state of the initial Hamiltonian (represented in its eigenbasis in the first panel) is projected onto many eigenstates of the new Hamiltonian and undergoes unitary evolution.

tical ensemble average. Hence, if our state populates primarily eigenstates in the middle of the spectrum, its local observables will appear thermal.

### 5.3 EXPERIMENTAL PROTOCOL

In our experiments, we study the emergence of thermalization in a six-site chain of interacting Bosons. To initiate the experiment, we isolate a  $2 \times 6$  plaquette from a larger low-entropy Mott insulator with unity filling as shown in figure 5.3. At this point, each system is in a product state of single-atom Fock states on each of the constituent sites. We then suddenly switch on tunnelling along the chains while the tunnelling between them is suppressed. Each chain is restricted to the original six sites by introducing a barrier at the ends of the chains to prevent tunnelling out of the system. These combined steps realize the setting described in the previous section. Each chain represents an identical but independent copy of a quenched system of six particles on six sites,

which evolves in the quenched Hamiltonian for a controllable duration.

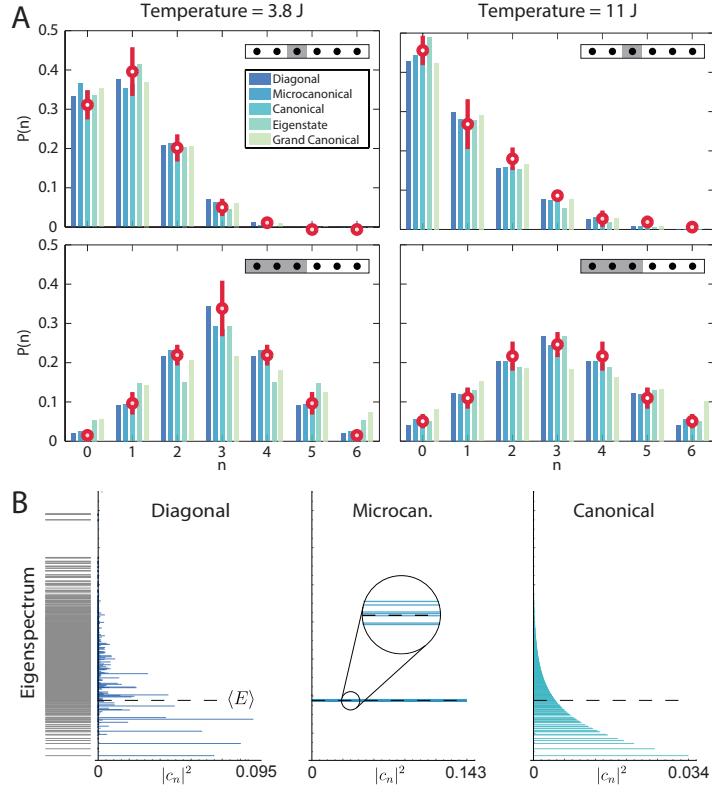


**Figure 5.3:** Experimental sequence Using tailored optical potentials superimposed on an optical lattice, we deterministically prepare two copies of a six-site Bose-Hubbard system, where each lattice site is initialized with a single atom. We enable tunnelling in the  $x$ -direction and obtain either the ground state (adiabatic melt) or a highly excited state (sudden quench) in each six-site copy. After a variable evolution time, we freeze the evolution and characterize the final quantum state by either acquiring number statistics or the local and global purity.

In the data that follow, we realize measurements of on-site number statistics and the quantum purity of the state. For measurements of the later, we append to the quench evolution a beam splitter operation that interferes the two identical copies by freezing dynamics along the chain and allowing for tunnelling in a projected double-well potential for a prescribed time<sup>7</sup>. In the last step for both measurements, a potential barrier is raised between the two copies and a full counting procedure is performed to measure the resulting occupation on each site of each copy.

#### 5.4 LOCAL OBSERVABLES IN THE THERMALIZED PURE STATE

In order to observe thermalization in our system, we first focus on measurements of atom number distributions in various subsystems of the chain. In figure 5.4B the number distribution for a single site and half of the chain is shown for two different final values of  $U/J = 1.56$  and  $U/J = 0.38$ . The data is averaged in the saturated regime over 5 times between 10 and 20 ms, and the error bars are the standard deviation in the measured probabilities. We can already make an observation that the standard deviation of measured probabilities is small compared to their value, which indicates that the system is in a quasi-stationary state with small temporal fluctua-

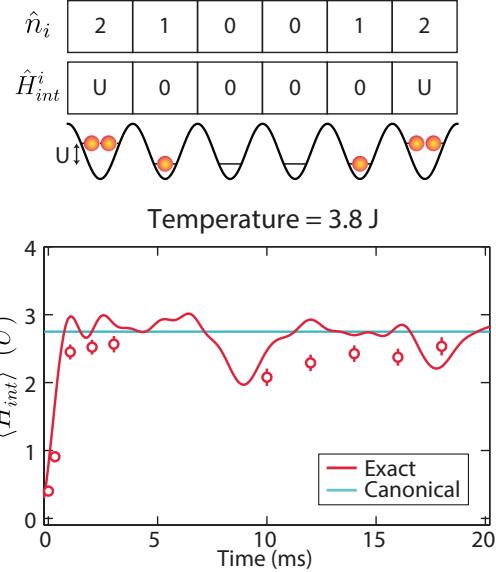


**Figure 5.4:** Local observables in a globally pure quenched state. (A) Along with the microcanonical ensemble, several other closely related ensembles are compared to the data. (B) Thermalization of local observables. For the different temperatures and subsystems shown, the measured number statistics are in excellent agreement with microcanonical and canonical thermal ensembles, verifying the thermal character of the local density matrix. A grand-canonical ensemble reproduces the data very well as long as the subsystem is small compared to the full system. The error bars are the standard deviation of our observation over times between 10 and 20 ms.

tions around the mean value- a characteristic of a thermal state.

We compare our measurements to the predictions of thermal ensembles that are illustrated in Figure 5.4A, as well as a grand-canonical ensemble truncated to our total atom number<sup>7</sup>: this ensemble perhaps most closely models how well the many-body state can act as a reservoir for its constituent subsystems. The consistency within the error bars indicates that in this temporal range our observations remain near the thermal predictions despite the presence of temporal fluctuations. For the single site subsystem, the data is in good agreement with all the ensembles

$$\hat{H}_{int} = \sum_i \hat{H}_{int}^i = \frac{U}{2} \sum_i n_i(n_i - 1)$$



**Figure 5.5:** Local observables in a globally pure quenched state. Thermalization occurs even for global quantities such as the full system interaction energy. The thermalization dynamics as calculated from our number-resolved images are in near agreement with exact numerical simulation and a canonical prediction.

considered. Despite the fact that the quenched state is in a large distribution of eigenstates, surprisingly, we find favorable agreement for the case of a single eigenstate ensemble: this illustrates a key principle of ETH, which holds that expectation value of local observables, very slowly from eigenstate to eigenstate and is therefore relatively insensitive to the width of the distribution of populated states from the quench. We perform the same comparisons to the three-site case in the bottom two panels. Here we also observe agreement with most ensembles, though, interestingly, there is relatively less agreement with the single eigenstate and grand-canonical ensembles, particularly for the lower temperature quench. This variation in the agreement may suggest that these ensembles are more sensitive to the relative size of the traced out reservoir compared to the subsystem, which indicates directions of further experiments<sup>7</sup>.

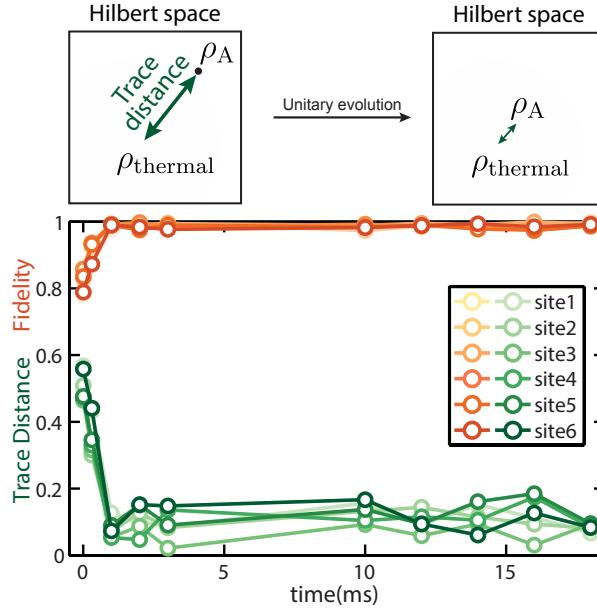
The above measurements were on specific subsystems, but our measurements also allow extraction of the average global interaction energy given by Hamiltonian:

$$\hat{H}_{int} = \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1), \quad (5.3)$$

where  $\hat{n}_i$  is the number operator on  $i$ -th site and the summation runs over all sites of the chain (see fig. ??). Since the interaction Hamiltonian is diagonal in the Fock basis, we can use our measurements of the final particle configurations to compute the expectation value  $\langle \hat{H}_{int} \rangle$ . For the  $T = 3.8J$  data, we show a time scan indicating the initial growth in this quantity, which starts at zero since the initial state is a single particle per site. These observations, at long times, are in near agreement with the canonical prediction. Interestingly, this measurement is sensitive to the entire six-site system as opposed to some subset of sites, which might suggest that it is global and unlikely to thermalize. Yet,  $\langle \hat{H}_{int} \rangle$  undergoes thermalization because it is a sum of local operators, each of which thermalizes and is insensitive to the global purity of the full system. The observed agreement is consistent with the idea that only a small set of operators, such as the global purity we measure or other specific fine-tuned state projectors, can truly distinguish the pure state we produce from a thermal state.

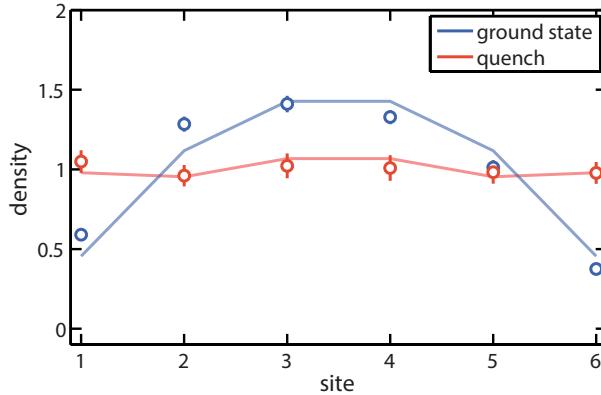
## 5.5 THERMALIZATION OF SUBSYSTEM DENSITY MATRIX

We can perform a more rigorous test of single-site thermalization by comparing the measured density matrix of each site with the reduced density matrix of a canonical thermal ensemble  $\rho_A^T$  (Figure 5.6B). Our measurements of probabilities to observe a given particle number on a site completely characterize that single-site density matrix, because there are no coherences between different number states due to super-selection rules. With this measured density matrix, we can perform a quantitative comparison to a thermal ensemble using the trace distance



**Figure 5.6: Observation of local thermalization.** In measuring the probabilities to observe a given particle number on a single site, we can obtain the local, single-site density matrix and observe the approach to thermalization. Using two different metrics, we compare the mixed state observed to the mixed state derived from the subsystem of a canonical thermal ensemble, after a quench to  $U/J = 1.56$ . The trace distance provides an effective distance between the mixed states in Hilbert space, while the fidelity is an overlap measure for mixed states. The two metrics illustrate how the pure state subsystem approaches the thermal ensemble subsystem shortly after the quench. The starting value of these quantities is given by the overlap of the initial pure state with the thermal mixed state. Solid lines connect the data points.

$(\frac{1}{2}\text{Tr}(|\rho_A^T - \rho_A|))$  and quantum fidelity  $(\text{Tr}(\sqrt{\sqrt{\rho_A^T}\rho_A\sqrt{\rho_A^T}}))$ , both of which quantify the similarity of two mixed quantum states. After a short time, we see a quantum fidelity exceeding 99% and a trace-distance that fluctuates between 0 and 0.1, indicating the similarity between the local density matrix of a verified pure state with the local density matrix of a thermal state. The correspondence between the observables of a pure state and thermal state depends on the equivalence of their reduced density matrices within the Hilbert space sampled by the observable. The measurement of Figure 5.6B therefore shows that observables for the single-site Hilbert space should agree with the predictions of thermal ensembles.

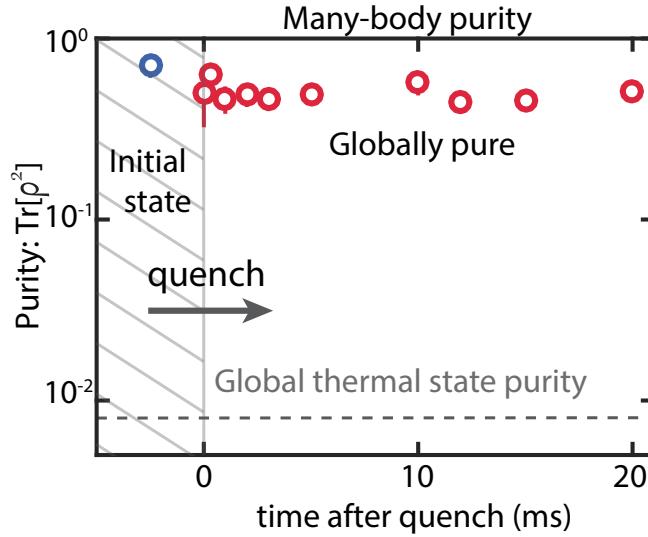


**Figure 5.7: Observation of local thermalization.** (A) After quenching to  $U/J = 0.38$ , the saturated average particle number on each site (density) is nearly equal among the sites of the system, which resembles a system at thermal equilibrium. By comparison, the ground state for the same Bose-Hubbard parameters has significant curvature.

One consequence of single-site thermalization is that for a thermal state individual sites look identical to one another. In a system with open boundary conditions, it appears in a striking contrast with the ground states, that are trying to minimize their energy. As an example, we look at the average particle number distribution in our system (see fig. 5.7). While the ground state exhibits significant curvature, corresponding to the ground state solution of the Shrödinger equation for a particle in the box. The quenched state exhibits a flat density distribution, manifesting the indestructibility between different sites of the system.

## 5.6 GLOBAL STATE PURITY

One straightforward consequence of the subsystem's density matrix being thermal is that the state of the subsystem is no longer pure. One potential explanation could be that during the time evolution that leads to thermalization the state as a whole got couples to the environment and hence loses its global purity. In order to exclude this scenario, we perform the measurement of purity in our system during the time evolution.



**Figure 5.8: Experimental sequence** We can measure the global many-body purity, and observe a static, high purity. This is in stark contrast to the vanishing global purity of the canonical thermal ensemble, yet this same ensemble accurately describes the local number distribution we observe.

Tomography of the full quantum state would typically be required to extract the global purity, which is particularly challenging in the full 462-dimensional Hilbert space defined by the itinerant particles in our system. Furthermore, while in spin systems global rotations can be employed for tomography<sup>7</sup>, there is no known analogous scheme for extracting the full density matrix of a many-body state of itinerant particles. The many-body interference described here, however, allows us to extract quantities that are quadratic in the density matrix, such as the purity<sup>7</sup>. After performing the beam splitter operation, we can obtain the quantum purity of the full system and any subsystem simply by counting the number of atoms on each site of one of the six-site chains (Figure ??C). Each run of the experiment yields the parity  $p^{(k)} = \prod_i p_i^{(k)}$ , where  $i$  is iterated over a set of sites of interest in copy- $k$ . The single-site parity operator  $p_i^{(k)}$  returns 1 (-1) when the atom number on site- $i$  is even (odd). It has been shown that the beam splitter operation yields  $\langle p^{(1)} \rangle = \langle p^{(2)} \rangle = \text{Tr}(\rho_1 \rho_2)$ , where  $\rho_i$  is the density matrix on the set of sites considered for each copy<sup>???</sup>. Because the preparation and quench dynamics for each copy are identical, yielding

$\rho_1 = \rho_2 \equiv \rho$ , the average parity reduces to the purity:  $\langle P^{(k)} \rangle = \text{Tr}(\rho^2)$ . When the set of sites considered comprises the full six-site chain, the expectation value of this quantity returns the global many-body purity, while for smaller sets it provides the local purity of the respective subsystem.

Surprisingly we find that the global many-body state retains its quantum purity during the entire time of our experiment, affirming the unitarity of its evolution following the quench (see fig. 5.8). This global measurement also clearly distinguishes the quantum state we produce from a canonical thermal ensemble, we used to describe the local observables in our system, as the purity between the two is different by two orders of magnitude. The deviation of the global state purity from unity can be explained by non-perfect beamsplitter fidelity in our system due to the residual interaction between the particles.

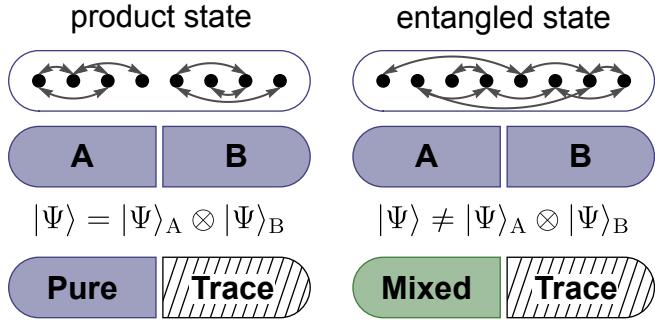
## 5.7 ENTANGLEMENT IN A QUANTUM SYSTEM

One of the defining concepts of quantum mechanics is a superposition principle, that allows the system to be in the multiple states at the same time. This can lead to the formation of correlations that go beyond classical local theories<sup>???</sup>. The most famous example is a Bell state<sup>???</sup>, which can be written in the context of two spin  $\frac{1}{2}$  particles as

$$|\psi\rangle = (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}, \quad (5.4)$$

where up and down arrows represent one of the two states of each spin. The remarkable property of this state is, that no matter which basis one would choose to measure the spin of both particles, the results of the measurement will always be anti-correlated. This basis independent correlations are called entanglement.

For a many-body system, it can be a formidable task to find all possible correlations that exist between its subsystems. To overcome this complication, we rely on the property of the entangled



**Figure 5.9: Entanglement and purity** For a globally pure state the purity of the subsystem is related to the entanglement (represented by arrows) between the subsystems. If subsystems A and B are not entangled the full state is a tensor product of pure states of the subsystems. Therefore each of them remain pure if the information about the other one is ignored. Contrary, when A and B are entangled, the subsystems purity is decreased upon tracing, as the correlations between A and B are ignored in the reduced density matrix

states, that ignoring the information about one subsystem puts the other into a classical mixture of states that contribute to entanglement. The number of states in this mixture can be quantified by the purity of the subsystem's density matrix  $\rho$ , given by  $\text{Tr}(\rho^2)$ . If the state is pure  $\text{Tr}(\rho^2) = 1$ , whereas for a mixed state  $\text{Tr}(\rho^2) < 1$ .

Consider an example of a globally pure state  $|\psi_{AB}\rangle$  that is split into two subsystems A and B (see fig 5.9). Ignoring the information of one of the subsystems is equivalent to tracing the full system density matrix  $\rho = |\psi_{AB}\rangle \langle \psi_{AB}|$  with respect to degrees of freedom of ignored subsystem, for example  $\rho_A = \text{Tr}_B(\rho)$ . In case subsystems A and B are not entangled the full state is a tensor product of pure states of the subsystems  $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ . Hence  $\rho_A = |\psi_A\rangle \langle \psi_A|$  and  $\rho_B = |\psi_B\rangle \langle \psi_B|$ , therefore  $\text{Tr}(\rho_A^2) = \text{Tr}(\rho_B^2) = 1$ . For an entangled state, the subsystems are less pure than the full system as the correlations between A and B are ignored in the reduced density matrix,  $\text{Tr}(\rho_A^2) = \text{Tr}(\rho_B^2) < 1$ .

Even if the full many-body state is mixed ( $\text{Tr}(\rho_{AB}^2) < 1$ ), it is still possible to measure entanglement between the subsystems<sup>???</sup>. It is sufficient<sup>?????</sup> to prove this entanglement by showing

that the subsystems are less pure than the full system, i.e.

$$\begin{aligned} \text{Tr}(\rho_A^2) &< \text{Tr}(\rho_{AB}^2) \\ \text{Tr}(\rho_B^2) &< \text{Tr}(\rho_{AB}^2) \end{aligned} \quad (5.5)$$

These inequalities provide a tool for detecting entanglement in the presence of experimental imperfections. Furthermore, quantitative bounds on the entanglement present in a mixed many-body state can be obtained from these state purities<sup>??</sup>.

A common measure of the entropy both in the classical and quantum case is von Neumann entropy, defined with respect to the density matrix  $\rho$  as:

$$S_{vN} = -\text{Tr}[\rho \text{Log}(\rho)], \quad (5.6)$$

However, in our experiments we quantify the entanglement using the second-order Rényi entropy  $S_A = -\text{Log}(\text{Tr}[\rho_A^2])$ , which is the logarithm of the purity of the subsystem density matrix. While the von Neumann entropy is typically used in the context of statistical mechanics, both quantities grow as a subsystem density matrix becomes mixed and increasingly entropic. In the Rényi case, the purity in the logarithm quantifies the number of states contributing to the statistical mixture described by the density matrix.

It is important to note that the thermodynamic relations in statistical mechanics, defined with the von Neumann definition, do not directly apply for the Rényi definition. However, Rényi entropy provides a lower bound to the von Neumann entropy and is directly accessible by our measurements.

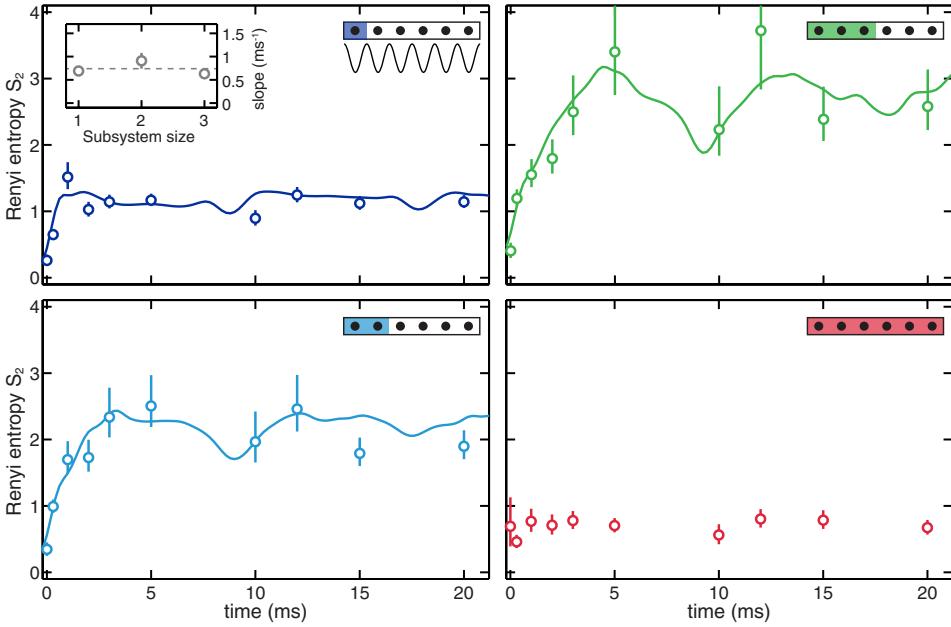
## 5.8 ENTANGLEMENT ENTROPY DYNAMICS AFTER QUENCH

The growth of entanglement following a quench is key to understanding how entropy forms within the subsystems of a pure quantum state, thereby facilitating thermalization<sup>???</sup>. We study the dynamics of the entanglement entropy, for varying subsystem sizes (Figure 5.10). Initially, we observe an approximately linear rise in the entropy, with similar slope among the subsystems considered (Figure 5.10 inset)<sup>?</sup>. After an amount of time that depends on the subsystem size, the entanglement entropy saturates to a steady-state value, about which there are small residual temporal fluctuations. The presence of residual fluctuations is, in part, attributable to the finite-size of our system. An exact numerical calculation of the dynamics with no free parameters shows excellent agreement with our experimental measurements. Crucially, the data indicate that while the subsystems acquire entropy in time (Figure 5.10,A-C), the full system entropy remains constant and is small throughout the dynamics (Figure 5.10D). The high purity of the full system allows us to conclude that the dynamical increase in entropy in the subsystems originates in the propagation of entanglement between the system's constituents.

The approximately linear rise at early times (Figure 5.10 inset) is related to the spreading of entanglement in the system within an effective light cone<sup>???</sup>. Furthermore, in analogy to the growth of thermodynamic entropy in an equilibrating classical mechanical system, such as a gas in a closed container, we observe the growth of local entropy in a closed quantum mechanical system. In the quantum mechanical case, however, the mechanism responsible for entropy is entanglement, which is absent from a system modeled by classical mechanics.

## 5.9 CORRESPONDENCE BETWEEN ENTANGLEMENT AND THERMAL ENTROPY

When a system thermalizes, we expect that the saturated values of local observables should correspond to the predictions of a statistical ensemble. By analogy, if the entanglement entropy



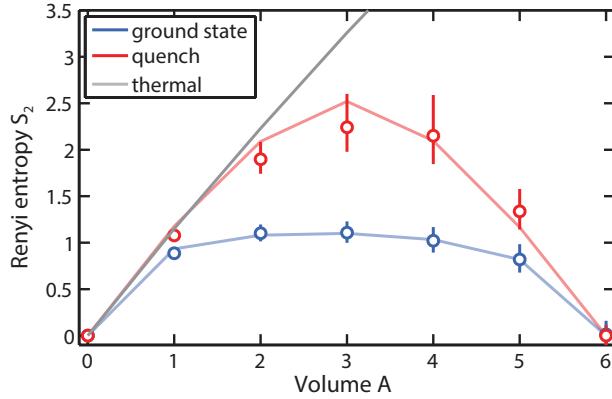
**Figure 5.10:** Dynamics of entanglement entropy. Starting from a low-entanglement ground state, a global quantum quench leads to the development of large-scale entanglement between all subsystems. We quench a six-site system from the Mott insulating product state ( $J/U \ll 1$ ) with one atom per site to the weakly interacting regime of  $J/U = 0.64$  ( $J/(2\pi) = 66\text{Hz}$ ) and measure the dynamics of the entanglement entropy. As it equilibrates, the system acquires local entropy while the full system entropy remains constant and at a value given by measurement imperfections. The dynamics agree with exact numerical simulations with no free parameters (solid lines). Error bars are the standard error of the mean (S.E.M.). For the largest entropies encountered in the three-site system, the large number of populated microstates leads to a significant statistical uncertainty in the entropy, which is reflected in the upper error bar extending to large entropies or being unbounded. Further information about statistics is discussed in the Supplementary Materials<sup>7</sup>. Inset: slope of the early time dynamics, extracted with a piecewise linear fit<sup>7</sup>. The dashed line is the mean of these measurements.

plays the role of thermal entropy, then in a thermalized pure state we expect extensive growth in the entanglement entropy with subsystem volume. When the entanglement entropy in a quantum state grows linearly with the size of the subsystem considered, it is known as a volume law. Ground-breaking theoretical work using conformal field theory has shown that indeed, at long times, a volume law is expected for a quenched, infinite, continuous system, while only an area law with a log correction is expected for the ground state<sup>???</sup>. Characterizing the large amount of

entanglement associated with a volume law is particularly challenging because it results in nearly every entry of the density matrix having small, but importantly non-zero magnitude.

Using the techniques outlined in this work, we show measurements displaying a near volume law in the entanglement entropy (Figure 5.11A). A linear growth with volume in the entanglement entropy occurs when each subsystem incoherently populates a number of states that scales with the size of the subsystem Hilbert space. This is because, for the Bose-Hubbard model, the Hilbert space is approximately exponential in the lattice size, which results in a linear growth in  $S_A = -\text{Log}(\text{Tr}[\rho_A^2])$ . Furthermore, the exact slope of the entanglement entropy versus subsystem volume depends on the average energy of the thermalized pure state<sup>7</sup>. By contrast, we can prepare the ground state of the quenched Hamiltonian by adiabatically reducing the lattice depth. Here, the superfluid ground state of the Bose-Hubbard model has suppressed entanglement, which is predicted to incur slow logarithmic growth in the entanglement entropy with subsystem volume<sup>7</sup>. Our measurements clearly distinguish the two cases. The back-bending of the entanglement entropy as the subsystem surpasses half the system size indicates that the state is globally pure. In the quenched state, the high global purity is striking in a state that locally appears completely dephased, which is behaviour often associated with environmentally-induced decoherence or other noise sources.

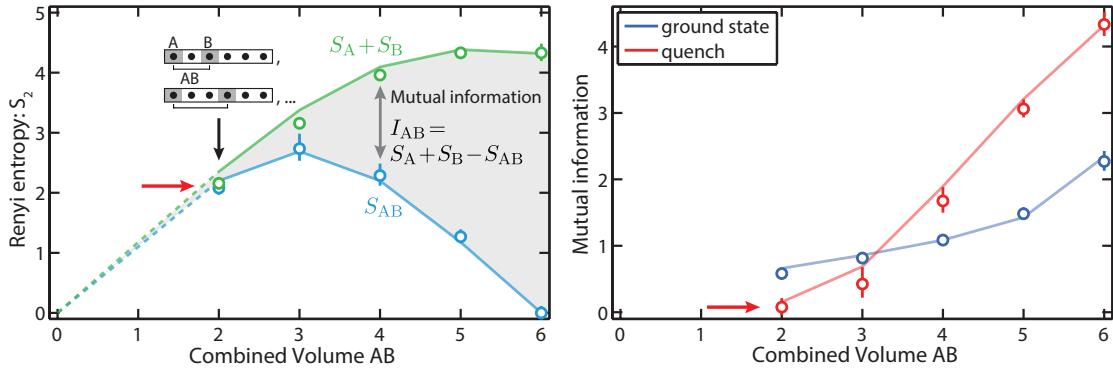
We further observe near-quantitative agreement between the exact dependence of the entanglement entropy with subsystem volume and the prediction of a thermal ensemble. We make this comparison by computing a canonical thermal ensemble  $\rho^T$  with an average energy that is the same as the quenched quantum state produced experimentally<sup>7</sup>. The grey line in Figure 5.11A is the Rényi (thermal) entropy as a function of subsystem size for this calculated thermal state. Although our limited system size prevents comparison over a large range of subsystem sizes, the initial rise of the entanglement entropy with subsystem volume mimics that of the thermal entropy. Despite their similarity, it is worth emphasizing the disparate character of the thermal and



**Figure 5.11:** Thermalized many-body systems. After the quench, the many-body state reaches a thermalized regime with saturated entanglement entropy. In contrast to the ground state, for which the Rényi entropy only weakly depends on subsystem size, the entanglement entropy of the saturated, quenched state grows almost linearly with size. As the subsystem size becomes comparable to the full system size, the subsystem entropy bends back to near zero, reflecting the globally pure zero-entropy state. For small subsystems, the Rényi entropy in the quenched state is nearly equal to the corresponding thermal entropy from the canonical thermal ensemble density matrix.

entanglement entropy. The entanglement entropy (either the Rényi or von Neumann) is instantaneously present in the pure quantum state after coherent unitary evolution, arising from the non-separability of the quantum state between the subsystem and traced out degrees of freedom. On the other hand, the von Neumann thermal entropy within a subsystem of a mixed thermal state is the thermodynamic entropy in statistical mechanics, which could be extracted from irreversible heat flow experiments on the subsystem<sup>7</sup>. Therefore, the similarity of the Rényi entropies we observe points to an experimental equivalence between the entanglement and thermodynamic entropy<sup>??</sup>.

The behaviour of the entanglement entropy provides a clear framework for understanding the entropy within thermalizing, closed quantum systems. However, one of the most famous features of entanglement, the presence of non-local correlations, appears inconsistent with what one expects of thermalized systems. In particular, the massive amount of entanglement implied by a volume law suggests a large amount of correlation between disparate parts of the system, while a



**Figure 5.12:** Thermalized many-body systems. The mutual information  $I_{AB} = S_A + S_B - S_{AB}$  quantifies the amount of classical (statistical) and quantum correlations between subsystems A and B. For small subsystems, the thermalized quantum state has  $S_A + S_B \approx S_{AB}$  due to the near volume law scaling (red arrow), leading to vanishing mutual information. When the volume of AB approaches the system size, the mutual information will grow because  $S_A + S_B$  exceeds  $S_{AB}$ . (C) We study  $I_{AB}$  vs the volume of AB for the ground state and the thermalized quenched state. For small system sizes, the quenched state exhibits smaller correlations than the adiabatically prepared ground state and is nearly vanishing. When probed on a scale near the system size, the highly entangled quenched state exhibits much stronger correlations than the ground state. Throughout this figure, the entanglement entropies from the last time point in Fig. 3 are averaged over all relevant partitionings with the same subsystem volume; we also correct for the extensive entropy unrelated to entanglement<sup>7</sup>. All solid lines are theory with no free parameters.

key feature of a thermal state is the very absence of these long-range correlations. A useful metric for correlations, both classical (statistical) and quantum, between two subsystems  $A$  and  $B$  is the mutual information  $S_A + S_B - S_{AB}$ <sup>??</sup>. The mutual information demonstrates that the amount of correlation in the presence of a volume law is vanishing for subsystem volumes that sample less than half the full system, which is where the entropy growth is nearly linear (Figure 5.11B,C). Furthermore, even though the thermalized quantum state carries more entanglement entropy than the ground state, small subsystems display smaller correlations than in the superfluid ground state. Once the subsystem volume is comparable to the system size, which is where the entanglement entropy deviates from the volume law, the quantum correlations entailed by the purity of the full system become apparent (Figure 5.11C). The mutual information, therefore, illustrates how the volume law in the entanglement entropy yields an absence of correlations between sufficiently local observables, even though the quantum state retains a large amount of entanglement.

## 5.10 CONCLUSIONS

In this chapter we have shown how a subsystem of a pure many-body state, undergoing quantum unitary dynamics, dynamically loses its purity through entanglement with the rest of the system. This leads to the thermalization of the subsystem. We also show the correspondence between the classical thermal entropy and quantum entanglement entropy in the system. Our observations speak to a natural translation between thermalizing quantum and classical systems.

Classical statistical mechanics relies on a fundamental assumption: a system in thermal equilibrium can be found in any microstate compatible with the thermodynamic constraints imposed on the system, and, as such, is described by an ensemble of maximal entropy<sup>??</sup>. Although it is vastly successful, classical statistical mechanics does not itself justify this entropy maximization for closed systems<sup>??</sup>, and an open systems approach only defers the question of thermalization to the union of the bath and system<sup>?</sup>. While ergodicity and time-averaging can provide a justifica-

tion for entropy maximization in closed classical mechanical systems, ergodicity is not applicable on the same scale that statistical mechanics is successful, and time-averaging can require exponentially long times<sup>???</sup>. The latter also obscures the fact that there is in reality only one system, which, nevertheless, is well-modeled by an entropic ensemble<sup>?</sup>. Our studies, and beautiful recent theoretical work<sup>???</sup>, hint towards a microscopic origin for entropy maximization in a single quantum state, namely that induced by the entanglement we measure.

# 6

## Many-body localization

### 6.1 LOCALLY BORING GLOBALLY EXCITING

As we have seen in the previous chapter, isolated interacting many-body systems undergoes thermalization dynamics when taken out of equilibrium. This causes the subsystem's degrees of freedom to be ultimately described by a thermal ensemble, even if the full system is in a pure state<sup>???</sup>. As a consequence, the local information about the initial state of the subsystem gets scrambled and transferred into non-local correlations that are only accessible through global observables<sup>???</sup>.

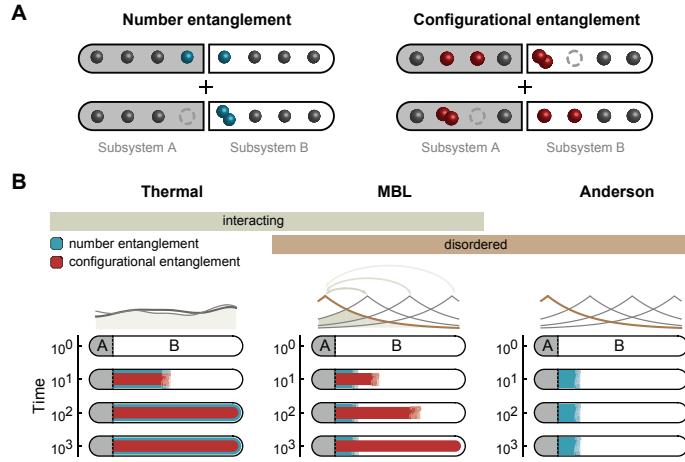
Disordered systems can provide an exception to this paradigm of quantum thermalization. As was first pointed out by Anderson, in the presence of disorder the eigenstates of non-interacting

particles localize exponentially<sup>?</sup>. This results in the long-term memory of initial density distribution, which goes in odds with the concept of thermalization. This phenomenon has been observed on a number of different platforms<sup>??????</sup>. However, the interacting problem appeared to be more complicated, and for a long time, the question of thermalization in interacting disordered systems remained elusive.

Recent theoretical works have identified the presence of the localized phase called many-body localization (MBL)<sup>??????</sup>. First experimental studies of interacting disordered system were focused on the low laying part of the spectrum studying glassy dynamics<sup>????</sup>, however, this phenomenon is qualitatively different from MBL, as the later describes the properties of the eigenstates of the system in the middle of the spectrum. Later on, MBL has been experimentally identified through the persistence of the initial density distribution<sup>??????</sup>, demonstrating the breakdown of ergodicity in the system.

While ergodicity breakdown is an important property of MBL, the presence of interaction in the system enables peculiar quantum many-body dynamics. In the absence of particle transport, the system generates non-local correlations<sup>??</sup>, which are inaccessible to local observables<sup>??</sup>. These dynamics are considered to be the hallmark of MBL and distinguish it from its non-interacting counterpart, called Anderson localization<sup>???????</sup>. Experimental studies have observed a buildup of a two-point correlation function during transient dynamics<sup>?</sup>, however, an unambiguous observation of non-local nature of correlations has remained elusive. Since the studies of this effect requires an exquisite control over the system's coherence over a long evolution time.

We study these many-body dynamics by probing the entanglement properties of an MBL system with fixed particle number<sup>????</sup>. We distinguish two types of entanglement that can exist between a subsystem and its complement (see fig. 6.1): *Number entanglement* implies that the particle number in one subsystem is correlated with the particle number in the other. It is generated by tunnelling across the boundary between the subsystems. *Configurational entanglement* implies



**Figure 6.1:** Entanglement dynamics in non-equilibrium quantum systems. (A) Subsystems A and B of an isolated system out of equilibrium entangle in two different ways: *number entanglement* is a superposition of states with different particle numbers in the subsystems and is generated through particle motion across the boundary; *configurational entanglement* is a superposition of states with different particle arrangement within the subsystems and requires both particle motion and interactions. (B) In the absence of disorder, both types of entanglement rapidly spreads across the entire system due to delocalization of particles (left panel). The degree of entanglement and the timescales change drastically when applying disorder (central panel): particle localization spatially restrict number entanglement, yet interactions allow configurational entanglement to form very slowly across the entire system. A disordered system without interactions shows only local number entanglement while the slow growth of configurational entanglement is completely absent (right panel).

that the configuration of the particles in one subsystem is correlated with the configuration of the particles in the other. It arises from a combination of particle motion and interaction. The formation of particle and configurational entanglement changes in the presence or absence of interactions and disorder in the system (Fig. 1B). In thermal systems without disorder, interacting particles delocalize and rapidly create both types of entanglement throughout the entire system. Contrarily, for Anderson localization, number entanglement builds up only locally at the boundary between the two subsystems. Here the lack of interactions prevents the substantial formation of configurational entanglement. In MBL systems, number entanglement builds up in a similarly local way as for Anderson localization. However, notably, the presence of interactions additionally enables the slow formation of configurational entanglement throughout the entire system.

## 6.2 DISORDER POTENTIAL

The properties of a system (especially a finite-size one) can be altered due to the presence of so called rare regions<sup>?????</sup>, when multiple sites next to each other have the same chemical potential. In order to avoid this issue, we chose to study the system, where disorder is given by the periodic potential with incommensurate period with respect to the lattice spacing, realizing interacting Aubry-André model for bosons in one dimension<sup>??</sup>, which is described by the Hamiltonian

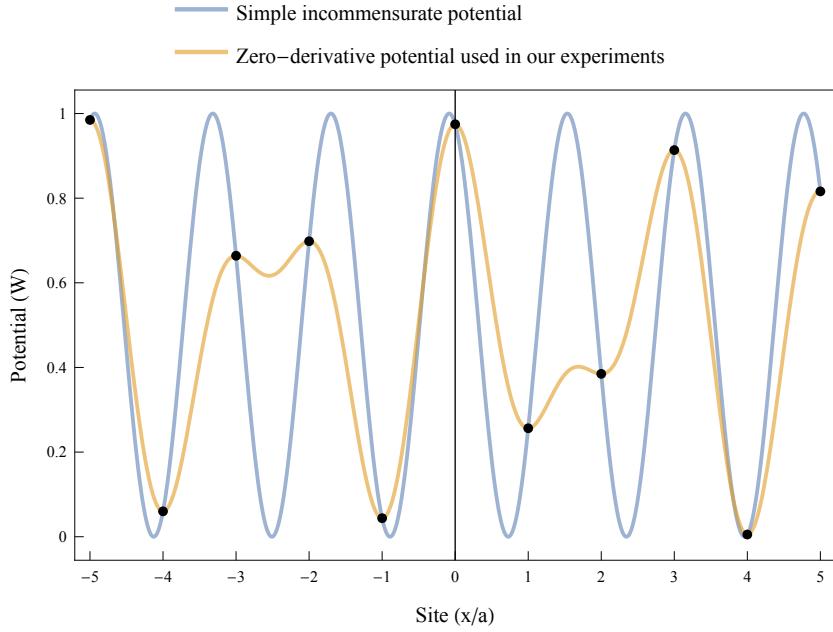
$$\hat{\mathcal{H}} = -J \sum_i (\hat{a}_i^\dagger \hat{a}_{i+1} + h.c.) + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) + W \sum_i h_i \hat{n}_i, \quad (6.1)$$

where  $\hat{a}_i^\dagger$  ( $\hat{a}_i$ ) is the creation (annihilation) operator for a boson on site  $i$ , and  $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$  is the particle number operator on that site. The first term describes the tunneling between neighboring lattice sites with the rate  $J/\hbar$ , where  $\hbar$  is the reduced Planck constant. The second term represents the energy shift  $U$  when multiple particles occupy the same site. The last term introduces a site-resolved potential offset, which is created with an incommensurate lattice  $h_i = \cos(2\pi\beta i + \phi)$  of period  $\beta \approx 1.618$  lattice sites (which is a very good approximation for a golden ratio), phase  $\phi$ , and amplitude  $W$ . In our experiment, we achieve independent control over  $J$ ,  $W$ , and  $\phi$  (see fig. ??).

We use our DMD to project a disorder potential onto our atoms. A conceptually simple way of producing such a potential is to project a second lattice with incommensurate periodicity onto the bare lattice holding the atoms, e.g. by adding a potential of the form

$$V_{\text{simple}}(x) = z \times W \cos^2 \left( \pi \frac{x}{\beta a} + \phi \right) \quad (6.2)$$

to the existing lattice. Since,  $\beta$  is very close to the golden ratio  $\beta - 1 \approx \frac{1}{\beta}$ , which means that, if the incommensurate potential is sampled when  $x$  is integer multiples of  $a$ , any integer difference



**Figure 6.2:** Disorder engineering. Plot showing a simple quasiperiodic disorder potential (blue) and the custom potential used in our experiments (yellow). The custom potential was engineered to yield the same on-site values as the simple one, but additionally possesses a vanishing first derivative at the position of the atoms (black dots), thereby making the system less sensitive to shaking-induced heating processes.

of  $\beta$  results in the same sampled potential.

The potential in Eq. 6.2 has one major drawback: if the lattice site sits on the slope of the potential, the offset sampled by the atoms varies as the relative position of the disorder potential with respect to the bare lattice is perturbed, thereby making the system susceptible to shaking-induced heating processes. In order to make the system more robust, we set out to find a potential  $V(x)$  that would provide the same on-site potential values as the potential in Eq. 6.2, while also having a vanishing first derivative at the individual sites of the bare lattice, i.e. a potential  $V(x)$  such that

$$V(x)|_{x=na} = V_{\text{simple}}(x)|_{x=na}, \quad \frac{\partial V(x)}{\partial x}\Big|_{x=na} = 0 \quad \forall n \in \mathbb{Z}$$

where  $a$  is the lattice constant of the bare lattice, and the position of an arbitrary lattice site is

defined to be  $x = 0$ .

This can be achieved by using the property of the golden ration  $\beta - 1 = \frac{1}{\beta}$  and the fact that we only sample the potential at the integer pints, resulting in:

$$\cos^2\left(\pi\frac{x}{\beta a}\right) = \cos^2\left(\pi\frac{(\beta - 1)x}{a}\right) = \cos^2\left(\pi\frac{(\beta - m)x}{a}\right), \quad \forall x = n \cdot a \quad (6.3)$$

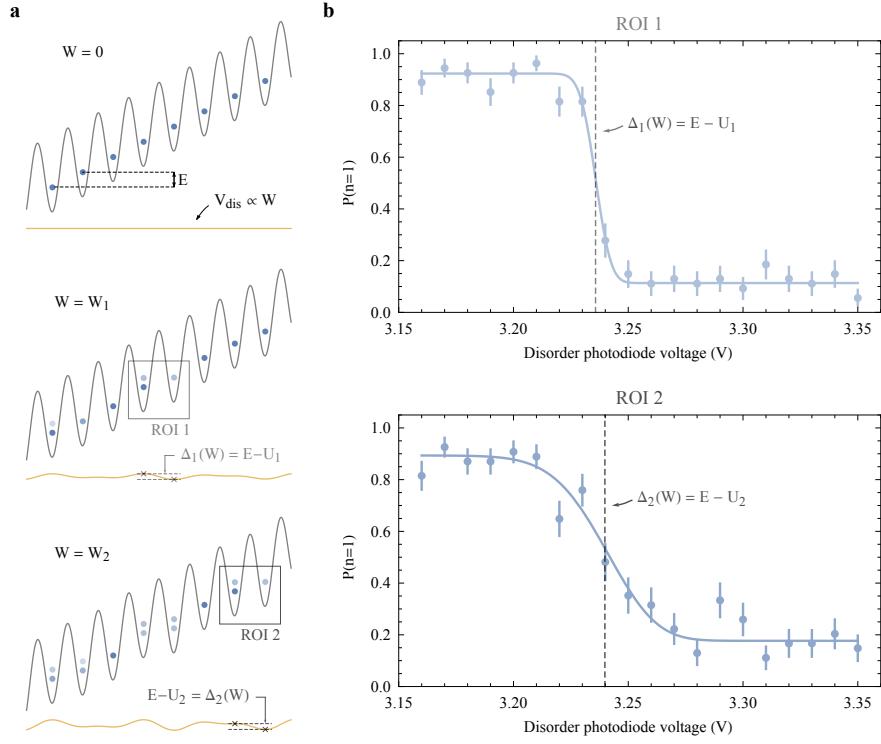
for any pair of integers  $m$  and  $n$ . One such potential is given by

$$V_{\text{dis}}(x) = 2 \times W \left[ (2 - \beta) \cos^2\left(\pi(\beta - 1)\frac{x}{a} + \phi\right) + (\beta - 1) \cos^2\left(\pi(\beta - 2)\frac{x}{a} + \phi\right) \right]. \quad (6.4)$$

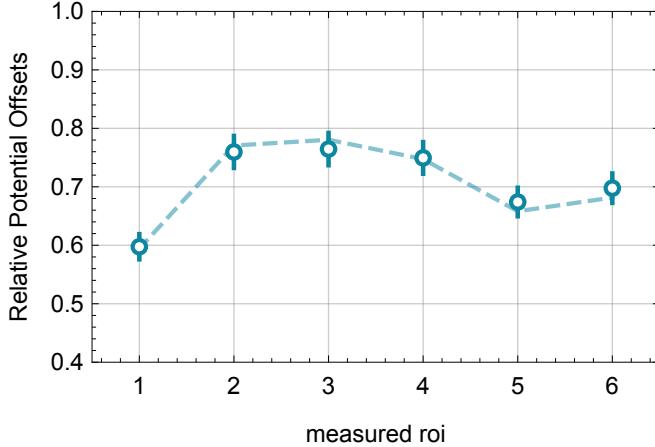
Here both cosine functions have the same value at the position of the lattice sites as the original potential, and their amplitudes are chosen such that the derivative at the position of the sites is canceled. Figure 6.2 shows the potentials given by equations 6.2 and 6.4 in units of  $W$  for  $\phi = 0.16$ .

### 6.3 CALIBRATION OF THE APPLIED DISORDER POTENTIAL

In order to be able to compare our experimental results with numerical simulations, we need a precise knowledge of the exact chemical potential offsets experienced by the atoms. To measure that, we prepare an  $n = 1$  Mott insulator and turn on a magnetic field gradient along the direction of the chain, that offsets neighbouring sites by energy  $E$ . Subsequently, we change the lattice depth to an intermediate value where tunnelling is principally significant but still suppressed by the tilt. We adiabatically ramp up the power of the disorder potential to various values and image the atom number distribution along the chain (see fig. ??a). When the potential difference between two neighbouring sites is equal to the applied tilt plus or minus  $U$  (depending on which way the potential offsets the sites), one of the atoms is allowed to tunnel freely between the sites,



**Figure 6.3: Calibration of the disorder strength  $W$  in our system.** (a) We prepare the same tilted lattice system as for the  $U$  calibration described in section ?? (top). Afterwards, we adiabatically ramp the optical disorder power to different voltages on the corresponding photodiode and analyze the resulting number statistics for various regions of interest. If the disorder power  $W$  is made sufficiently large, the potential difference  $\Delta_i(W)$  it induces within a given region of interest may partially compensate for the linear tilt and enable resonant tunnelling processes (centre, bottom). (b) We again post-select on measurement outcomes with exactly two atoms in a given region of interest and plot the probability  $P(n = 1)$  of finding them on separate sites. The resulting curves show a sharp decay around the photodiode voltages where  $\Delta_i(W) = E - U_i$ , a condition allowing for resonant tunnelling processes that deplete the initial  $n = 1$  population. Having calibrated the  $E - U_i$  values independently and knowing the constant ratios  $\Delta_i(W)/W$  from the disorder potential we project, we can link the corresponding photodiode voltages to the appropriate values of  $W$  in the plane of our atoms. Technically, data from a single region of interest would be enough to calibrate  $W$ . The independent calibration of various regions additionally allows us to verify that the potential seen by the atoms indeed obeys the shape we expect, which we found to be the case in our measurements.

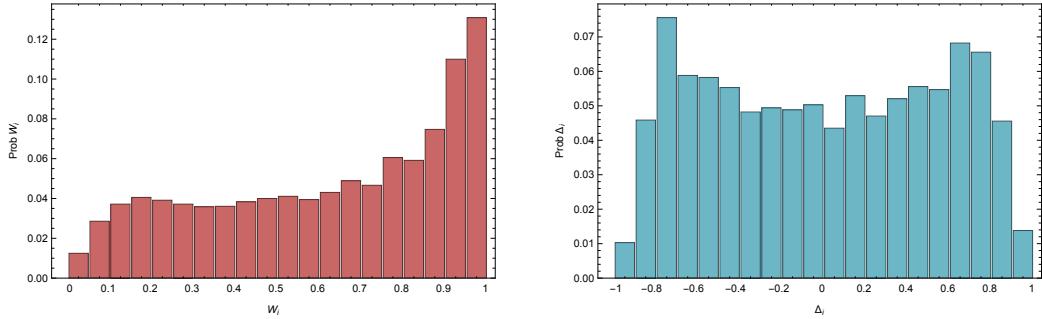


**Figure 6.4:** Relative potential offsets within one pattern. In order to benchmark the applied disorder potential, we compare the measured site offsets within one pattern to the numerical Fourier transform of the corresponding hologram (dotted line). The agreement between the data and numerical prediction confirms the high level of control over the applied optical potential. The relative potential offsets (both data and numerical prediction) are normalized to the largest potential difference in all applied patterns.

resulting in the deviation from the uniform density. By plotting the post-selected unity-filling fraction for a given region  $i$ , we can identify the potential offset between any pair of sites in terms of  $E$  and  $U$ , which we can calibrate independently in our system<sup>7</sup> (see fig. ??a).

In our experiment, we use 200 different disorder patterns. Calibrating each one of them would be an extremely time-consuming procedure. However, since DMD should provide us with the exact Fourier transform of the displayed pattern, we can perform numerical Fourier transform of all used patterns and determine the potential difference between the sites for each one of them. In order to verify that the potentials experienced by the atoms do not suffer from the aberrations of the imaging system and other possible imperfections, we measure all potential differences within one pattern and compare them to the result of the Fourier transform (see fig. 6.4). Excellent agreement of the measurement with numerical Fourier transform indicates the applicability of our method.

Our numerical studies show, that the exact potential differences do not exactly coincide with



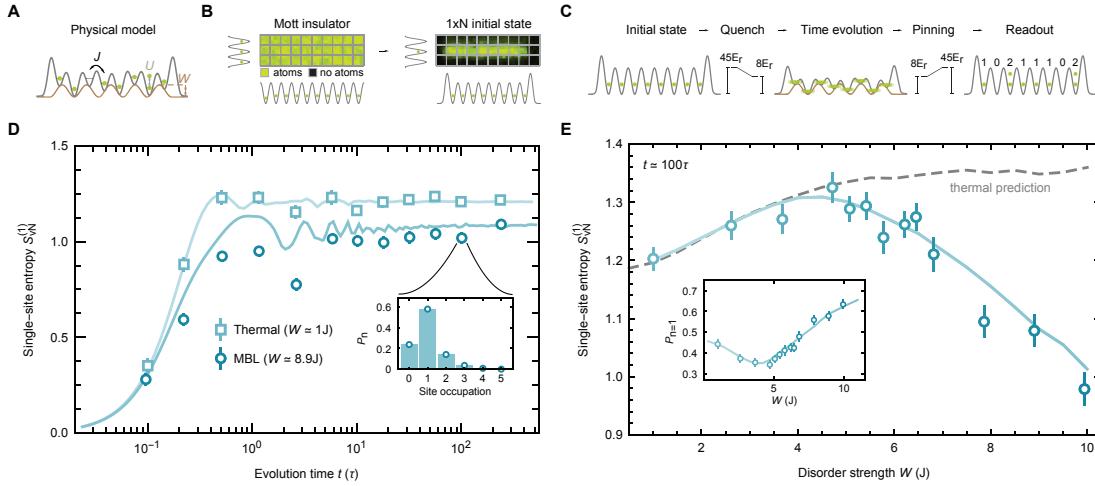
**Figure 6.5:** Distribution of the disorder potential. The left plot is a histogram of the DMD generated on-site quasi-periodic potential ( $W_i$ ) which is normalized to 1. The right plot is a histogram of the inter-site difference of the DMD generated quasi-periodic potential ( $\Delta_i$ )

true quasi-periodic lattice potential (see fig. ??). For all numerical simulations in this chapter, we use the disorder values extracted from the Fourier transforms of the holograms.

#### 6.4 BREAKDOWN OF THERMALIZATION

In order to investigate the breakdown of thermalization, we focus on the subsystem that consists of a single lattice site. The conserved total atom number enforces a one-to-one correspondence between the particle number outcome on a single site and the number in the remainder of the system—entangling the two during tunnelling dynamics. Ignoring information about the remaining system puts the subsystem into a mixed state of different number states. The associated number entropy is given by  $S_n^{(1)} = -\sum_n p_n \log(p_n)$ , where  $p_n$  is the probability of finding  $n$  atoms in the subsystem. Since the atom number is the only degree of freedom of a single lattice site,  $S_n^{(1)}$  captures all of the entanglement between the subsystem and its complement and is equivalent to the single-site von Neumann entanglement entropy  $S_{vN}^{(1)}$ .

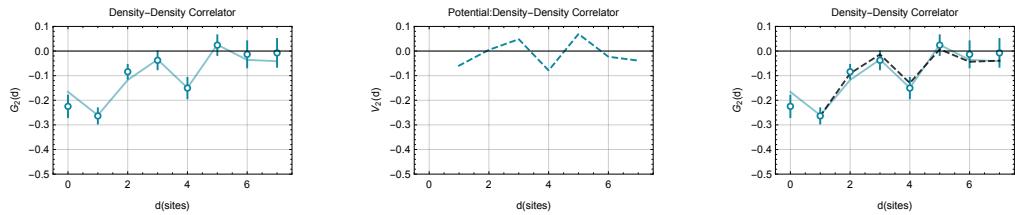
Counting the atom number on an individual lattice site in different experimental realizations allows us to obtain the probabilities  $p_n$  and compute  $S_{vN}^{(1)}$ . We perform such measurements for



**Figure 6.6: Breakdown of thermalization on a single site level** (A) We compute the single-site von Neumann entropy  $S_{\text{vN}}^{(1)}$  from the site-resolved atom number statistics (inset) after different evolution times (scaled with tunneling time  $\tau = \hbar/J$ ) in the presence of weak and strong disorder. (B) Probability  $p_1$  to retrieve the initial state (inset) and  $S_{\text{vN}}^{(1)}$  for different  $W$ , measured after  $100\tau$  evolution. The deviation from the thermal ensemble prediction for strong disorder signals the breakdown of thermalization in the system. All solid lines show the prediction of exact diagonalization calculations without any free parameters. Each data point is sampled from 197 disorder realizations (SI). Error bars denote the s.e.m.

various evolution times. At low disorder depth ( $W = 1.0(1)J$ ), the entropy grows over a few tunnelling times and then reaches a stationary value (see fig. ?? A). The stationary value is reduced for a deep disorder ( $W = 8.9(1)J$ ) and remains constant over two orders of magnitude, up to several hundred tunnelling times. The lack of entropy increase indicates the absence of heating in the system. The excellent agreement of the measured entropy with *ab initio* calculations up to the longest measured evolution times suggests a highly unitary evolution of the system.

We perform measurements of  $S_{\text{vN}}^{(1)}$  at different disorder strengths following an evolution of one hundred tunneling times (see fig. ?? B). To evaluate the degree of local thermalization, we compare the results with the prediction of a thermal micro-canonical ensemble for our system. For weak disorder, the measured entropy agrees with the predicted value, whereas the entropy is significantly reduced for strong disorder—signalling the absence of thermalization in the system.



**Figure 6.7: Potential and Wavefunction: Two-Point Correlations** The left plot shows the measured  $G_2(r)$  data for an 8 site, 8 atom system after 100 tunneling times with disorder strength  $W = 5.8J$  and  $U = 2.7J$ . This solid line is from simulation performed by exact diagonalization for the same experimental parameters. The middle plot is the calculated two-point correlator for the applied quasi-periodic potential and averaged over many different disorder realizations – the qualitative features in the left plot mimic the correlation function of the applied potential. The right plot is the same data with the fitted correlation function (black) made from the applied potential correlations ansatz (6.6).

As a consequence, the system retains some memory of its initial conditions for arbitrarily long evolution times. We indeed find that the probability to retrieve the initial state of one atom per site increases for the strong disorder (see fig. ?? B).

The thermal prediction, shown in the figure ?? B is calculated using equal probability statistical mixture of 11 eigenstates of the evolution Hamiltonian  $H_{evo}(W)$ , that are closest to the average energy of the initial state, given by  $E_o = \langle \psi_o | H_{evo}(W) | \psi_o \rangle$ . We verify that the results do not depend on the exact number of included eigenstates in the vicinity of the chosen value.

## 6.5 SPATIAL LOCALIZATION

The breakdown of thermalization is expected to be a consequence of the spatial localization of the particles. Previous experiments have determined the decay length of an initially prepared density step into empty space<sup>7</sup>. We measure the localization by directly probing density-density correlations within the system. These correlations are captured by  $G^{(2)}(d) = \langle n_i n_{i+d} \rangle - \langle n_i \rangle \langle n_{i+d} \rangle$ , where  $\langle \dots \rangle$  denotes averaging over different disorder realizations as well as all sites  $i$  of the chain. The particle numbers on two sites at distance  $d > o$  are uncorrelated for  $G^{(2)}(d) = o$ . If a particle moves a distance  $d$ , the sites become anti-correlated, and the correlator decreases to  $G^{(2)}(d) < o$ .

For a moderate disorder strength, we observe, that the measured correlation function has some structure in it (see fig. 6.7 A). This can be explained by taking into account the correlations that are inherent to the disorder potential itself. Consider  $V(x, \phi) = \cos(2\pi\beta x + \phi)$ , we can define two-point correlation function of the potential as

$$V_2(d) = \langle V(x, \phi)V(x + d, \phi) \rangle_{d, \phi} - \langle V(x, \phi) \rangle_{d, \phi} \langle V(x + d, \phi) \rangle_{d, \phi}, \quad (6.5)$$

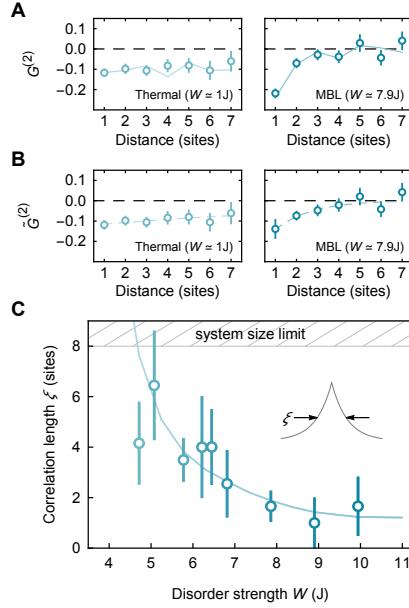
where  $\phi$  runs over all possible phases of the disorder (see fig. 6.7 B). In order to unbias our results from the effects coming from the potential, we introduce a fitting ersatz, that takes into account the correlations of the potential

$$G_2^{fit}(d) = (A + B \times V_2(d))e^{-d/\xi}, \quad (6.6)$$

where  $A$ ,  $B$ , and  $\xi$  are fitting parameters. Subtracting those effects leads to  $\tilde{G}^2(d) = G^{(2)}(d) - B \times V_2(d)e^{-d/\xi}$ , that we use to interpret our results.

We measure the density-density correlations  $\tilde{G}^2(d)$  for different disorder strengths in the stationary regime (see fig. 6.8). For low disorder, we find the correlations to be independent of distance and below zero. This indicates that the particles tunnel across the entire system and hence are delocalized. On the other hand, at the strong disorder, only nearby sites show significant correlations, signalling the absence of particle motion across large distances. We thus conclude that the particles are localized. We extract the correlation length by fitting an exponentially decaying function to the data (see fig. 6.8 A). For increasing disorder, the correlation length decreases from the entire system size down to around one lattice site (see fig. 6.8 B).

Our observation of localized particles is consistent with the description of MBL in terms of local integrals of motion<sup>??</sup>. It describes the global eigenstates as product states of exponentially



**Figure 6.8:** Spatial localization of the particles. (A) The density-density correlations  $G^{(2)}(d)$  as a function of distance  $d$  at weak and strong disorder after an evolution time of  $100t$ . The alternating nature of the density-density correlations is imprinted by the autocorrelation function of the quasiperiodic potential. (B) We retrieve the exponential decay of the correlation function by subtracting the influence of the quasiperiodic potential (SI). (C) Particle motion is confined within the correlation length  $\xi$ . We use a fit to extract  $\xi$  for different disorder strengths. The fit function is a product of an exponential decay with the autocorrelation function of the quasiperiodic potential (SI). Each measurement is sampled from 197 disorder realizations (SI). The solid lines show the prediction of exact diagonalization—calculated without any free parameters. Error bars denote the s.e.m in (A-B), and the fit error in (C).

localized orbitals. The correlation length extracted from our data is a measure of the size of these orbitals. Since the latter form a complete set of locally conserved quantities, this picture connects the breakdown of thermalization in MBL with non-thermalizing, integrable systems.

## 6.6 DIFFERENT TYPES OF ENTANGLEMENT

We now turn to a characterization of the entanglement properties of larger subsystems, starting with a subsystem covering half the system size. As for the case of a single lattice site, the particle number in the subsystem can become entangled with the number in the remaining system

through tunnelling dynamics. However, subsystems which extend over several lattice sites, with a given particle number, offer the particle configuration as an additional degree of freedom for the entanglement.

In the Schmidt basis the density matrix for the subsystem  $\rho$  has only diagonal elements, and the von Neumann entropy can be written as

$$S_{\text{vN}} = \sum_i \rho_{ii} \log (\rho_{ii}). \quad (6.7)$$

If the global particle number is conserved, there can not be any coherence between the states with different particle number within the subsystem, since the particle number in the subsystem's complement is uniquely determined. Therefore the density matrix can be written as  $\rho_{ii} = p_n \rho_{ii}^{(n)}$ , where  $p_n = \sum_{ii \subset n} \rho_{ii}$  is the probability of having  $n$  particles within the subsystem and  $\rho_{ii}^{(n)}$  describe the relative probability of being in the certain state within the given particle number sub-sector, such that  $\sum_{ii} \rho_{ii}^{(n)} = 1$ . Hence the von Neumann entropy can be written as:

$$\begin{aligned} S_{\text{vN}} &= \sum_{n=0}^N \sum_i p_n \rho_{ii}^{(n)} \log \left( p_n \rho_{ii}^{(n)} \right) = \sum_{n=0}^N \sum_i p_n \rho_{ii}^{(n)} \left( \log (p_n) + \log \left( \rho_{ii}^{(n)} \right) \right) = \\ &= \sum_{n=0}^N p_n \log (p_n) \sum_i \rho_{ii}^{(n)} + \sum_{n=0}^N p_n \sum_i \rho_{ii}^{(n)} \log \left( \rho_{ii}^{(n)} \right) = \\ &= \sum_{n=0}^N p_n \log (p_n) + \sum_{n=0}^N p_n \sum_i \rho_{ii}^{(n)} \log \left( \rho_{ii}^{(n)} \right) = \\ &= S_n + S_c \end{aligned} \quad (6.8)$$

The first term  $S_n = \sum_{n=0}^N p_n \log (p_n)$ , we call number entanglement, describes how many sub-sectors with different particle number in the subsystem are populated. This entanglement is created through particle tunneling between two subsystems. The second,  $S_c = \sum_{n=0}^N p_n \sum_i \rho_{ii}^{(n)} \log \left( \rho_{ii}^{(n)} \right)$ , we call configurational entanglement, measures how many states within the same sub-sector are

occupied. This type of entanglement only builds up substantially in interacting systems, since configurational correlations require an arrangement of several particles in the different subsystems. An analogous split can be done of any conserved quantity and for example also exists for spin systems with conserved total magnetization instead of the particle number.

## 6.7 MBL AS A SYSTEM OF LOCAL INTEGRALS OF MOTION

The universal properties of entanglement dynamics in the MBL phase could be understood from the phenomenological model<sup>???</sup>. It states that eigenstates of an MBL system can be decomposed into a product state of local integrals of motion, that we refer to as localized orbitals. Although their exact structure in terms of original particle operators can be quite complicated<sup>?</sup>, the knowledge of exact mapping is not necessary in order to understand the resulting dynamics.

Each of those orbitals can be in an arbitrary superposition of different occupation states, determined by the initial state projection. Furthermore, the combination of original particles tunnelling and interactions result in the effective Hamiltonian for the localized orbitals:

$$\hat{\mathcal{H}} = U_{\text{eff}} \sum_{i,j} e^{-|i-j|/\xi} \hat{o}_i \hat{o}_j, \quad (6.9)$$

where  $\hat{o}_i$  is the occupation number operator of the localized orbital  $i$  and  $U_{\text{eff}}$  is a constant given by the parameters of the original system. The exponential decay of interactions strength is due to the localized nature of the orbitals and  $\xi$  is their effective localization length.

In order to understand the dynamics of the system under this model let's consider a simple example of two orbitals  $i$  and  $j$ . As we mentioned before and as the form of the Hamiltonian suggests, the occupation number of each of the orbitals is a conserved quantity. First, notice that the only effect the time evolution has on the state is to advance the phase of some of the states depending on the orbitals occupations. Second, if the orbitals start in the single occupational

state the time evolution has a trivial effect of overall phase accumulation and, hence, can not lead to the entanglement between the orbitals.

However, if both orbitals start in the superposition of different occupation states the system dynamics will produce entanglement between the two at certain time points. To see that consider an initially unentangled state

$$|\psi(0)\rangle = (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) = |00\rangle + |01\rangle + |10\rangle + |11\rangle. \quad (6.10)$$

Under the time evolution only the last term has non-trivial dynamics, such that after a time  $t$  the state reads

$$|\psi(t)\rangle = |00\rangle + |01\rangle + |10\rangle + e^{iU_{\text{eff}}t/\hbar} |11\rangle, \quad (6.11)$$

and after  $t_{ent} = \pi \frac{\hbar}{U_{\text{eff}}}$  we get

$$|\psi(t_{ent})\rangle = |00\rangle + |01\rangle + |10\rangle - |11\rangle = |0\rangle \otimes (|0\rangle + |1\rangle) + |1\rangle \otimes (|0\rangle - |1\rangle). \quad (6.12)$$

The final state is maximally entangled, since two states of each orbital are orthogonal to each other. Note, that this effect is purely quantum, since it stems from the coherent phase evolution of the superposition states, hence it can not be captured by any mean-field approach or classically interacting particles.

The dynamics of  $S_n$  and  $S_c$  in the MBL regime ?? can be understood in the picture of localized orbitals. Since the localized orbitals restrict the particle motion, the number entropy can only develop within the localization length and hence  $S_n$  saturates at a lower value than for the thermal case. In the MBL regime, disorder suppresses the tunnelling. Therefore, saturation is reached at a later time. However, the dynamics of  $S_c$  are strikingly different, as they are governed by the relative phase evolution between different many-body states. The effective interactions among

the orbitals decay exponentially with the distance between them. As a consequence, entanglement between distant orbitals forms slowly, causing a logarithmic growth of  $S_c$ , even after  $S_n$  has saturated????.

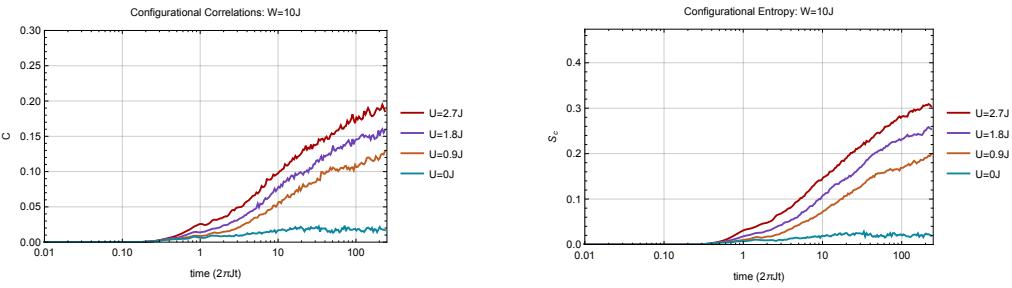
## 6.8 ACCESSING CONFIGURATIONAL ENTANGLEMENT THROUGH CORRELATIONS

In our experiment, we can independently probe both types of entanglement. We obtain the number entropy  $S_n$  through the probabilities  $p_n$  by counting the atom number in the subsystem in different experimental realizations. The configurational entropy  $S_c$ , in contrast, is challenging to measure in a many-body system since it requires experimental access to the coherences between a large number of quantum states?? . Here we choose a complementary approach to probe the configurational entanglement in the system. It exploits the configurational correlations between the subsystems, quantified by the correlator (SI):

$$C = \sum_{n=0}^N p_n \sum_{\{A_n\}, \{B_n\}} |p(A_n \otimes B_n) - p(A_n)p(B_n)|, \quad (6.13)$$

where  $\{A_n\}$  ( $\{B_n\}$ ) is the set of all possible configurations of  $n$  particles in subsystem A ( $N - n$  in B), and  $N$  is total number of particles in the system. All probability distributions are normalized within the subspaces of  $n$  particles in A and the remaining  $N - n$  particles in B. The configuration  $A_n \otimes B_n$  is separable if  $p(A_n \otimes B_n) = p(A_n)p(B_n)$ . The correlator therefore probes the entanglement through the deviation from separability between A and B.

In general, one does not a priori expect that there should be any linear relation between  $S_c$  and C. In fact, their asymptotic behavior is very different from one another. While  $S_c$  is a



**Figure 6.9:** Correlations and Configurational Entropy. The configurational correlations as measured in the fock basis is plotted on the left as a function of interaction strength and evolution time. The corresponding configurational entropy is plotted in the right as a function of interaction strength and evolution time. All of these simulations were performed by exact diagonalization on 6 sites with 6 atoms.

unbounded quantity, the correlator  $C$  has an upper bound of 2. To see that

$$\begin{aligned}
 C &= \sum_{n=0}^N p_n \sum_{\{A_n\}, \{B_n\}} |p(A_n \otimes B_n) - p(A_n)p(B_n)| \leq \\
 &\leq \sum_{n=0}^N p_n \sum_{\{A_n\}, \{B_n\}} |p(A_n \otimes B_n)| + |p(A_n)p(B_n)| = 2 \sum_{n=0}^N p_n = 2.
 \end{aligned} \tag{6.14}$$

However, in the localized regime for sufficiently small amounts of entanglement, our numerical studies show, that the two are proportional to one another (see fig. 6.9). This proportionality even hold for different values of intercalation strength, making it a good proxy for configurational entropy in such systems.

As pointed out in<sup>7</sup>, the total amount of entanglement in the MBL phase depends on the initial state of the system. Therefore, for any system size, the maximal amount of entanglement can

be adjusted such that the proportionality between  $S_C$  and  $C$  holds for the entire duration of time evolution, making the correlator a very powerful tool to study this phenomenon. The correlator  $C$  is also experimentally more accessible than  $S_c$ , since it involves projective measurements in the particle number basis. Therefore, it can be generalized to higher dimensions and different experimental platforms, e.g. trapped ions, neutral atoms, superconducting circuits, where a direct measurement of entanglement entropy remains challenging.

## 6.9 ENTROPY DYNAMICS AND SCALING

We study the time dynamics of  $S_n$  and  $C$  with and without disorder (Fig. 4B, C). Without disorder, both  $S_n$  and  $C$  rapidly rise and reach a stationary value within a few tunnelling times (insets).

In the presence of strong disorder, we find a qualitatively different behaviour for the two quantities:  $S_n$  reaches a stationary state within few tunnelling times, although after longer evolution time due to reduced effective tunnelling. Additionally, the stationary value is significantly reduced, indicating suppressed particle transport through the system. The correlator  $C$ , in contrast, shows a persistent slow growth up to the longest evolution times reached by our measurements.

The growth is consistent with logarithmic behaviour over two decades of time evolution. We conclude that we observe interaction-induced dynamics in the MBL regime, which are consistent with the phenomenological model<sup>???</sup>. The agreement of the long-term dynamics of  $S_p$  and  $C$  with the numerical calculations in the MBL regime confirms the unitary evolution of the system within its 6435-dimensional Hilbert space over 100  $\tau$ . The system remains in the finite-time limit, not in the finite-size limit since the spread of entanglement has not yet stopped at the longest studied evolution times.

Considering the entropy in subsystems of different size gives us insights into the spatial distribution of entanglement in the system: in a one-dimensional system, locally generated entanglement results in a subsystem size independent entropy, whereas entanglement from non-local

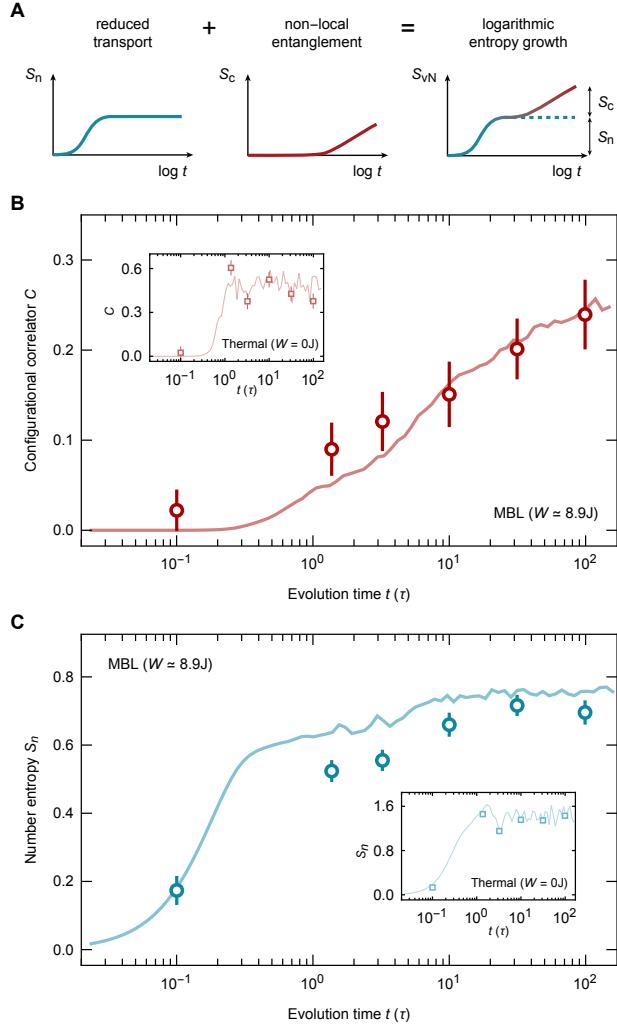
correlations causes the entropy to increase in proportion to the size of the subsystem. In reference to the subsystem's boundary and volume, these scalings are called area law and volume law. We find almost no change in  $S_n$  for different subsystems of an MBL system (Fig. 5A)—indicating an area law scaling due to localized particles and confirming that particle transport is suppressed. In contrast, the configurational correlations  $C$  increase until the subsystem reaches half the system size (Fig. 5B). Such a volume-law scaling is also expected for the entanglement entropy and demonstrates that the observed logarithmic growth indeed stems from non-local correlations across the entire system.

#### 6.10 CONCLUSION

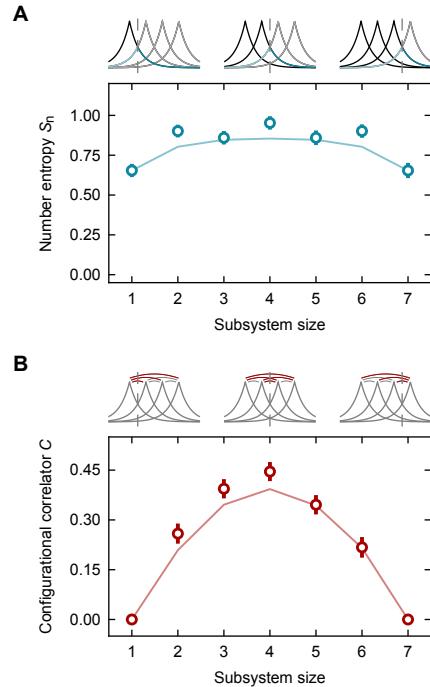
Investigating the growth of non-local quantum correlations has been a long-standing experimental challenge for the study of MBL systems. In addition to achieving exceptional isolation from the environment and local access to the system, such a measurement requires access to the entanglement entropy<sup>7</sup>. Our work provides a novel method to characterize the entanglement properties of MBL systems. Since it is based on measurements of the particle number fluctuations and their configurations, the method is experimentally accessible and can be generalized to higher dimensions and different experimental platforms, where direct measurement of entanglement entropy remains challenging, e.g. trapped ions, neutral atoms, superconducting qubits. The observation of slow coherent many-body dynamics along with the breakdown of thermalization coincides with the expected behaviour for larger systems and allows us to unambiguously identify and characterize the MBL state in our system.

In future, experiments at different system sizes will be of interest to shed light on the critical properties of the thermal-to-MBL phase transition, which are the subject of ongoing studies<sup>???</sup>. In our system, it is experimentally feasible to continue scaling the system size at unity filling to a numerically intractable regime. Additionally, we have full control over the disorder potential on

every site, which opens the way to studying the role of rare regions and Griffiths dynamics as well as the long-time behaviour of an MBL state with a link to a thermal bath<sup>??</sup>. Ultimately, these studies will further our understanding of quantum thermodynamics and whether such systems are suitable for future applications as quantum memories<sup>??</sup>.



**Figure 6.10:** Dynamics of the number and configurational entanglement. (A) The von-Neumann entanglement entropy is the sum of the number entropy and the configurational entropy, whose dynamics in an MBL system occur over different time scales (SI). (B) We probe the configurational entropy with the correlator  $C$ . At strong disorder, it shows a persistent slow increase that is consistent with a logarithmic growth in time, until the longest evolution times covered by our measurements. (C) The number entropy  $S_n$  reaches a stationary value within few tunneling times. Without disorder, the entanglement dynamics change: both  $S_n$  and  $C$  quickly reach a stationary value (insets). The solid lines show the prediction of exact diagonalization calculations without any free parameters. The above data was taken on a six-site system and averaged over four disorder realizations. Error bars denote the s.e.m.



**Figure 6.11:** Spatial distribution of the entanglement. Number entropy and configurational correlator in the MBL regime ( $W = 8.9J$ ) after an evolution time of  $100\tau$ . (A) The number entropy  $S_n$  does not depend on the subsystem size, i.e. follows an area law. (B) The configurational correlator  $C$  increases almost linearly with the subsystem size, showing a volume-law behavior. The solid lines show the prediction of exact diagonalization calculations without any free parameters. The above data was averaged over four disorder realizations. Error bars denote the s.e.m. and are below the marker size if hidden.

# 7

## Conclusions and outlook

In this thesis, I have presented techniques that allow us to image one-dimensional systems of atoms in the optical lattice with full number resolution on a single site level, extending the capabilities of quantum gas microscopes beyond the parity projection imaging. Together with the ability to prepare pure initial states with hight fidelity it enabled us to study coherent many-body dynamics over multiple decades of time evolution, more then an order of magnitude longer then has been previously demonstrated

Exquisite control over our system using digital micro-mirror device, that was pioneered in our group before, allowed us to experimentally verify the purity of a quantum state and observer how its subsystems undergo quantum thermalization during the time evolution. By comparing various local observables to the predictions of thermodynamic ensembles we were able to verify the

validity of eigenstate thermalization hypothesis in our system. Using a beamsplitter technique we studied the dynamics of the entanglement entropy in the thermalizing many-body system and were able to make a connection between it and the classical thermodynamic entropy in the thermal state.

By adding disorder to the system we realized a many-body localized state – the only known robust exception for thermalization in the quantum system. We studied the breakdown of eigenstate thermalization hypothesis by comparing the entropy of a single site density matrix to the predictions of thermal ensemble. For the first time, we have observed the localization of interacting particles in the real space by measuring the density-density correlation length in the system with uniform density. Long coherent evolution in our experiment allowed us to observe the hallmark feature of such many-body localized state: the formation of non-local correlations in the absence of transport, qualitatively distinguishing it from non-interacting disordered systems.

In the future, the study of larger systems will provide the way to probe the scaling predictions of eigenstate thermalization hypothesis. Conversely, the study of integrable Hamiltonians, where thermalization fails due to the emergence of an extensive number of conserved quantities as a result of fine-tuning of the system parameters, could provide a direct connection between thermalization and conservation laws. The studies of fullcounting statistics after the quench will reveal the relations between full quantum dynamics and various effective descriptions of the system.

In the context of thermal to many-body localized transition it is of a great interest to understand the behavior of the system at the intermediate values of the disorder depth. By varying the system's size the nature of the transition can be revisited. An alternative way to gain knowledge of the transition is to study a many-body localized system locally coupled to a thermal bath. The full control over the disorder potential makes our system a perfect candidate to carry out such experiments.

# A

## Computing expectation values in thermalized systems

ETH implies an equivalence between the local expectation values of a quenched many-body state and those of the thermal density matrix with the same average total energy as the many-body state. For the reported experiments, our system is initialized into the ground state,  $|\psi_0\rangle$ , of an ini-

tial Hamiltonian,  $\mathcal{H}_o$ , with single particle per site. At  $t = o$ , we quench the system into a Hamiltonian,  $\mathcal{H}_q$ , after which the system is allowed to evolve for a variable amount of time. Comparing to the data at longs times (10 – 20 ms, where we observe saturation), we can compute predictions for the expectation values of various local observables based upon different thermodynamic ensembles, computed using the following procedures.

### MICROCANONICAL ENSEMBLE

The microcanonical ensemble is an equal probability statistical mixture of all the eigenstates that lie within an energy interval given by the initial state  $|\psi_o\rangle$ . In the quenched Hamiltonian, the initial state has an energy  $E^{(o)} \equiv \langle \psi_o | \mathcal{H}_q | \psi_o \rangle$  while the eigenstates of  $\mathcal{H}_q$ ,  $|\phi_i^{(q)}\rangle$ , have energies  $E_i^{(q)}$ . The microcanonical ensemble is then composed of the  $N_{MC}$  number of eigenstates for which  $|E_i^{(q)} - E^{(o)}| < \delta E$ . For our numerical data, we have chosen  $\delta E = 0.2J$ , but the ensemble predictions are insensitive to the precise value of  $\delta E$ . The microcanonical ensemble can be represented by the thermal density matrix

$$\rho_{ij}^{MC} = \begin{cases} \frac{1}{N_{MC}}, & \text{if } i = j \text{ and } |E_i^{(q)} - E^{(o)}| < \delta E \\ 0, & \text{else} \end{cases}. \quad (\text{A.1})$$

### CANONICAL ENSEMBLE

The canonical ensemble is a statistical mixture of all the eigenstates in the system weighted by each state's Boltzmann factor,  $\exp(-E_i^{(q)}/k_B T)$ . The temperature in the Boltzmann factor is fixed

through the stipulation that the average energy of this thermal ensemble matches the energy of the initial state, i.e. we choose  $T$  such that  $\text{Tr}(\mathcal{H}_q \rho^{CE}) = \langle \psi_o | \mathcal{H}_q | \psi_o \rangle$ , where the thermal density matrix  $\rho^{CE}$  has the following construction,

$$\rho_{ij}^{CE} = \begin{cases} e^{-\frac{E_i^{(q)}}{k_B T}}, & \text{if } i = j \\ 0, & \text{else} \end{cases}.$$

#### DIAGONAL ENSEMBLE

The diagonal ensemble is a statistical mixture of all eigenstates of the full Hamiltonian  $\mathcal{H}_q$ , with the weights given by their amplitudes after quench.

$$\rho_{ij}^D = \begin{cases} |\langle \psi_o | \phi_i^{(q)} \rangle|^2, & \text{if } i = j \\ 0, & \text{else} \end{cases}$$

It carries all information about the amplitudes of the eigenstates but ignores all their relative phases.

#### GRAND CANONICAL ENSEMBLE

The grand-canonical ensemble requires calculating the temperature and chemical potential for the subsystem associated to the observable of interest. For example, the top (bottom) row of FIG. 6C pertains to the subsystem consisting of the third site (the first three sites) of the chain. We

calculate the temperature and chemical potential for the subsystem as follows. Because the energy and particle number within the subsystem are not conserved during the quench dynamics, we must compute the average energy  $\langle E_A \rangle$  and average number  $\langle N_A \rangle$  within the subsystem numerically. We time-evolve the full many-body state to the thermalized regime, then compute the reduced density matrix for the subsystem, with which we can calculate  $\langle N_A \rangle$  and  $\langle E_A \rangle$ . We note that the average energy of nearly all the subsystems is very close to that of the full system (zero), while the average number is nearly consistent with unity particle density. For the single site subsystems, however, there is no tunneling term to offset the interaction energy, and therefore these subsystems have non-zero energy. We perform this full calculation to account for finite-size effects that cause small temporal energy and number fluctuations. If we neglect the energy fluctuations, the grand-canonical predictions (described below) are negligibly different.

After the above calculations, we can compute the chemical potential and temperature. Using each  $H_A^N$ , the subsystem Bose-Hubbard Hamiltonian with  $N$  particles, we compute the eigenstates ( $|E_A^{N,i}\rangle$ , where  $i$  indexes the eigenstate) and energies ( $E_A^{N,i}$ ) for each particle sector. We seek  $T$  and  $\mu$  such that,

$$\langle N_A \rangle = \langle N_{\text{GCE}} \rangle = \frac{1}{Z} \sum_{i,N} N e^{-(E_A^{N,i} - \mu N)/k_B T}$$

and,

$$\langle E_A \rangle = \langle E_{\text{GCE}} \rangle = \frac{1}{Z} \sum_{i,N} E_A^{N,i} e^{-(E_A^{N,i} - \mu N)/k_B T},$$

where for each particle number  $N$ , the index  $i$  is summed over all eigenstates within that number sector. The partition function,  $Z$ , is the overall normalization. These equations are numerically

solved to find  $\mu$  and  $T$ . With these in hand, we arrive at the grand-canonical ensemble,

$$\rho_{\text{GCE}} = \frac{1}{Z} \sum_{i,N} |E_A^{N,i}\rangle \langle E_A^{N,i}| e^{-(E_A^{N,i} - \mu N)/k_B T}.$$

## OBSERVABLES

For all statistical ensembles above, expectation values of an observable  $\mathcal{A}$  are calculated from the density matrix as

$$\langle \mathcal{A} \rangle = \text{Tr}(\mathcal{A}\rho),$$

where  $\rho$  is the full system density matrix corresponding to the appropriate ensemble.

## References



**T**HIS THESIS WAS TYPESET using L<sup>A</sup>T<sub>E</sub>X, originally developed by Leslie Lamport and based on Donald Knuth's T<sub>E</sub>X. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, *Science Experiment 02*, was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at [github.com/suchow/Dissertate](https://github.com/suchow/Dissertate) or from its lead author, Jordan Suchow, at [suchow@post.harvard.edu](mailto:suchow@post.harvard.edu).