Many-Body Entanglement in Classical & Quantum Simulators

Johnnie Gray

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics & Astronomy

University College London

January 15, 2019

I, Johnnie Gray, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Entanglement is not only the key resource for many quantum technologies, but essential in understanding the structure of many-body quantum matter. At the interface of these two crucial areas are *simulators*, controlled systems capable of mimicking physical models that might escape analytical tractability. Traditionally, these simulations have been performed *classically*, where recent advancements such as tensor-networks have made explicit the limitation entanglement places on scalability. Increasingly however, analog *quantum* simulators are expected to yield deep insight into complex systems. This thesis advances the field in across various interconnected fronts. Firstly, we introduce schemes for verifying and distributing entanglement in a quantum dot simulator, tailored to specific experimental constraints. We then confirm that quantum dot simulators would be natural candidates for simulating many-body localization (MBL) - a recently emerged phenomenon that seems to evade traditional statistical mechanics. Following on from that, we investigate MBL from an entanglement perspective, shedding new light on the nature of the transition to it from a ergodic regime. As part of that investigation we make use of the *logarithmic negativity*, an entanglement measure applicable to many-body mixed states. In order to tie back into quantum simulators, we then propose an experimental scheme to measure the logarithmic negativity in realistic many-body settings. This method uses choice measurements on three or more copies of a mixed state along with machine learning techniques. We also introduce a fast method for computing many-body entanglement in classical simulations that significantly increases the size of system addressable. Finally, we introduce `quimb`, an open-source library for interactive but efficient quantum information and many-body calculations. It contains general purpose tensor-network support alongside other novel algorithms.

Impact Statement

Quantum technologies have the potential to make a huge impact across many areas of science and industry. One such technology is *quantum simulation*, which aims to mimic crucial models of interacting quantum particles with other, more controllable quantum systems. Such 'many-body' models are likely to underpin all sorts of advanced materials and chemicals that we expect will essentially always be impossible to fully simulate on normal, or 'classical', computers. One of the reasons that current computers find this such a hard task is *entanglement* – correlations between quantum components which are stronger than anything we are used to in our 'classical', day-to-day lives. Beyond this, entanglement is also a key perspective for understanding a myriad of outstanding problems in fundamental physics. Clearly then, tools to address questions of entanglement in this many-body setting are highly desirable, and these are a central part of what this thesis provides.

One of these is experimentally viable schemes for both measuring and distributing entanglement in quantum simulators. Another is a method for computing entanglement in much larger systems than has previously been possible. These tools are relevant for both quantum physicists working experimentally on quantum simulation or those wanting to perform classical simulations. We also use entanglement to reveal new features of a highly unusual phenomenon – many-body localization – that has attracted much attention in the condensed matter community. These collaborations have resulted in various academic publications. Another part of this research has been the development of open-source software for tackling quantum entanglement and simulation problems. As well as being useful in this field, work on the underlying software has involved collaborations with quantum chemists and climate scientists among many others, and produced several software publications.

Publications

This thesis has resulted in the following publications, the first five of which form the basis for the main chapters:

- Johnnie Gray, Abolfazl Bayat, Reuben K Puddy, Charles G Smith, and Sougato Bose. Unravelling quantum dot array simulators via singlet-triplet measurements. *Phys. Rev. B*, 94(19):195136, 2016

- Johnnie Gray, Sougato Bose, and Abolfazl Bayat. Many-body localization transition: Schmidt gap, entanglement length, and scaling. *Phys. Rev. B*, 97(20):201105, 2018

- Johnnie Gray, Leonardo Banchi, Abolfazl Bayat, and Sougato Bose. Machine-learning-assisted many-body entanglement measurement. *Phys. Rev. Lett.*, 121(15):150503, 2018

- Johnnie Gray. Fast computation of many-body entanglement. *arXiv preprint arXiv:1809.01685*, 2018

- Johnnie Gray. `quimb`: a python library for quantum information and many-body calculations. *Journal of Open Source Software*, 3(29):819, 2018

- Daniel GA Smith and Johnnie Gray. `opt_einsum` - a python package for optimizing contraction order for einsum-like expressions. *Journal of Open Source Software*, 3(26):753, 2018

Acknowledgements

I would like to begin of course by thanking my advisor Sougato Bose, who has been a never ending supply of knowledge and ideas across such a huge range of interesting topics. He has completely made this thesis possible. A special thanks also to Abolfazl Bayat who has not only been a fantastic collaborator but offered so much practical guidance as well. I very much hope that we get to work together in the future!

The CDT in delivering quantum technologies has been a great, structured, well-rounded way to approach a PhD and I would like to thank its heads, Andrew Fisher and Dan Browne, as well as its life-force, Lopa Murgai. It has been an pleasure to have done the course with such a nice cohort – Andrew, Cameron, Carlo, Claudia, Josh, Lorenzo, Sherif, Tim, and Tom – learning how to program a vending machine would not have been nearly as bearable without you.

Going back further I would like to particularly thank two of my school physics teachers, Barbara "Bob" Justham and Duncan Nicholls, who sadly passed away in my first year at university. The last chat I had with Duncan in the pub was about how confusing I found 'vectors' after one term – I think he would have been proud at the profligate usage of them in this thesis. Both of them were enormously successful in managing to convey the essential excitement and magic of physics and that inspiration has clearly stuck with me.

It goes without saying that there are many friends, too many to name, who at various points have offered such great advice and support – I am forever grateful to know such an amazing group of people.

Finally, I would like to thank my family for ultimately providing everything that has made this thesis possible – they should probably take all the credit.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantum science promises to fundamentally change many areas of technology, including, possibly most prominently, computation. However, in its most universal form, practical quantum computation is still widely considered to be a significantly distant goal. A reprieve is that one of the most useful tasks a quantum computer can do does not need to be done in a universal manner, nor even without errors — this is the task of *simulating* other quantum systems. The desire to do this arises from the huge difficulty in making ab initio calculations, either using analytical or classical computational methods, of the properties and dynamics of general many-body quantum systems. Many fundamental questions remain open regarding such systems, which also underpin understanding of all complex materials and chemicals. As such, any method that allows one to better predict their properties and understand them is clearly attractive, and this is the essential offer of quantum simulation.

Simulation sits at the intersection of, and requires the study of, several distinct fields. Firstly, the target systems and phenomena of interest are largely motivated by condensed matter models and many-body quantum theory. Secondly, there is an aspect of computational complexity - where does quantum simulation fit 'below' quantum computation, and what are the upper limits of classical simulation that define when quantum simulation is useful? Finally, quantum simulations are at the moment necessarily constrained by the physical architecture they run on — what models and measurements are possible now is thus motivated by the latest experimental developments.

An important perspective to take across all of these areas is that of *entanglement*. Although the situation is more subtle than simply saying that entanglement is the only thing separating quantum from classical simulations, it can be seen as, broadly speaking, the major ingredient. A direct implication of this is that at the limit of both experimental and computational methods, there often exists an effort to include or maintain as much controlled entanglement as possible. In the experimental setting this might include suppressing coherence and improving the quality of multi-qubit gates. In the classical simulation setting, the clearest example might be tensor networks, which are explicitly constructed by specifying the geometry and amount of entanglement. Beyond this feature of entanglement in a loose sense 'delimiting' classical and quantum simulations, it can also act itself when quantified as a versatile probe of various many-body phenomena. Particularly, its very general and abstract nature can reveal fundamental features that more traditional quantities might otherwise miss.

Given these points above, there is a strong case to be made for developing additional tools that let us address questions of entanglement in a many-body setting – be it in actual quantum simulations, or classical computational simulations. Indeed, developing these two sides hand in hand is particularly important with regards to near term quantum devices, where it will be essential to verify and cross-reference observed behaviour with known behaviour for as long as possible. In other words, as well as offering insight of their own into complex systems, classical simulations will be needed to certify and 'trust' quantum simulations, and entanglement should be a key aspect of this.

In this thesis we explore and advance the topic of many-body entanglement in quantum and classical simulations in a number of ways, but we begin in the remainder of this chapter by reviewing a variety of topics necessary to understand the context and basis for this research. This includes quantum simulation, key many-body models, entanglement, tensor-networks and experimental realizations. These sections aim to concisely set the scene, with more in-depth introductions found at the beginning of each chapter.

Then in Chapter 2 we propose two experimental schemes, one for characterizing a ground-state and one for distributing entanglement, tailored to quantum-dot arrays – a promising candidate for *solid state* quantum simulations. We also note that the effective model relevant for quantum dot arrays is very similar to that of a many-body localized spin chain and we provide numerical evidence that indeed quantum dot-arrays should naturally exhibit a transition to non-ergodic behaviour.

In Chapter 3 we focus further on the posited many-body localization transition, by studying two entanglement related quantities: the *Schmidt gap* and the disjoint *logarithmic negativity*. In doing so we resolve two outstanding problems. Firstly, the identification of a quantity that scales with a critical exponent in accordance with theoretical predictions. Secondly, the identification of a length scale that appears to diverge from both sides of the transition. Both of these results are evidence that aspects of this unusual phase transition may yet be understood using traditional methods – though exactly how this approach might be made remains elusive.

Given the success of the logarithmic negativity in revealing features of the many-body localization transition, as well as its various uses elsewhere in the literature, we then move in Chapter 4 onto developing a scheme for measuring the logarithmic negativity in the laboratory. This scheme involves preparing copies of a state and then performing counter-propagating swap measurements in order to evaluate the 'partially transposed moments'. Crucially, by adopting a machine learning approach to extracting the logarithmic negativity from these moments, we show that in many situations three copies of the state is all that is needed for robust experimental quantification of the many-body entanglement present.

The approaches in Chapter 3 and Chapter 4 both require the classical computation of the logarithmic negativity for as large subsystems of many-body states as possible. Although the logarithmic negativity is one of the only 'efficient' entanglement measures to compute, this is only in a technical sense. In fact, in the general case, computing it naively one is still limited to relatively small subsystems. To overcome this, in Chapter 5 we develop a generic method for stochastically but accurately computing 'spectral quantities' for operators lazily represented as tensor

networks. This method can be trivially applied to the computing of the logarithmic negativity in arbitrary subsystems of densely represented states and block subsystems of matrix product states. In doing so we essentially double the system size for which many-body entanglement is a tractable quantity to classically compute.

Finally, in Chapter 6 we introduce and demonstrate `quimb`, a python library for **qu**antum **i**nformation and **m**any-**b**ody calculations developed as part of this thesis. The relatively ambitious goal for `quimb` was to make a library that on the one hand is very easy and interactive to use, thus promoting quick exploratory research, but on the other hand does not sacrifice anything in terms of performance. Structurally, `quimb` is split into two parts: an core module for 'exact' dense and sparse calculations, and a tensor network module for larger many-body simulations. We demonstrate basic usage of both whilst noting various novel design advances. All of the other chapters make heavy use of simulation data generated with `quimb`, which includes advanced implementations of algorithms not found elsewhere.

1.1 Quantum Simulation

As with many things, the idea of a quantum simulator begins with Feynman, and his proposal in the early 1980s that a computer should be made with with quantum components [7]. Although the exact nature of how a quantum computer would be useful for solving general problems came later [8] (most prominently with Shor's algorithm [9] for factorizing prime numbers in polynomial time), Feynman understood that a huge amount of power was to be gained, his clear example being the difficulty met when attempting to simulate a quantum model using classical bits. The last 20 years have seen an incredible push towards realizing universal quantum computation, with much progress despite huge challenges. More recently however, there has been also been rising interest in constructing controllable quantum systems with the sole goal of simulation - spurred on by the fact that this comes with far less stringent engineering requirements. Recent reviews of quantum simulation can be found in [10, 11, 12, 13, 14, 15, 16].

The distinction between computing and simulating is important, which we hope

to highlight in this section, but additionally, this emerging field has focussed on *analog* rather than digital operation, so it is worth dwelling on what that means. Here, the continuous dynamics of some physical system are used to mimic a model of interest, with both being described approximately by the same mathematical equations. History has many examples of classical analog devices, both for simulating — such as orreries that predicted the position of the solar system's planets — and for computing — for example electronic integrators and mechanical differential equation solvers. At one point there was hope that these would even provide a form of supercomputing, since they could potentially solve very hard problems (NP-Complete) in polynomial time, although this was shown to come at the price of an exponential requirement for resources [17] and huge sensitivity on the initial set-up of the physical parameters. Instead, the rapid rise of general purpose digital computers, with discrete gates and error correction, means that we now think largely of simulations taking place on these – this is what we will refer to as *classical simulations*.

For simulating quantum systems, the first problem that one encounters is that even storing a description of the state, $|\psi\rangle$, requires an exponential number of coefficients. Taking the simplest non-trivial case of a generic pure system of N spin-$\frac{1}{2}$ particles, 2^N complex coefficients are required for a complete description, putting sizes of $N > 30$ beyond the capability of modern computers to store completely in memory, let alone manipulate and solve. So the idea of an analog quantum simulator is to represent this quantum information with exponentially fewer quantum components, interacting in the same way as the target model.

Here, the model is essentially all contained in the Hamiltonian, $\mathbf{H}$, that governs the system's dynamics via Schrodinger's equation:

$$\frac{\partial}{\partial t}|\psi(t)\rangle = -\frac{i}{\hbar}\mathbf{H}\,|\psi(t)\rangle \tag{1.1}$$

where we set the reduced Planck constant $\hbar = 1$ hereafter unless stated. We may be interested in the static, equilibrium properties of such a system, derived from the thermal state:

$$\rho(\beta) = e^{-\beta\mathbf{H}} \tag{1.2}$$

where β denotes the inverse temperature $(k_B T)^{-1}$. The groundstate $|\psi_{\text{gs}}\rangle$, in particular is likely to be of interest:

$$\mathbf{H}\,|\psi_{\text{gs}}\rangle = E_0\,|\psi_{\text{gs}}\rangle \tag{1.3}$$

where E_0 is the minimum eigenvalue of the Hamiltonian. One can equivalently define it as the zero temperature thermal-state:

$$|\psi_{\text{gs}}\rangle\,\langle\psi_{\text{gs}}| = \lim_{\beta\to\infty} \rho(\beta) \tag{1.4}$$

when the system is non-degenerate. More ambitious might be to simulate the dynamics of system, where for a time-independent Hamiltonian the unitary operator governing evolution is given by:

$$\mathbf{U}(t) = e^{-i\mathbf{H}t} \tag{1.5}$$

such that at some later time the state is given by

$$|\psi(t)\rangle = \mathbf{U}(t)\,|\psi(0)\rangle\,. \tag{1.6}$$

Or written explicitly in terms of eigenpairs, $\{E_k, |k\rangle\}$, one finds

$$|\psi(t)\rangle = \sum_k |k\rangle\, e^{-iE_k t}\,\langle k|\psi(0)\rangle\ , \tag{1.7}$$

which makes clear the important role that *all* Hamiltonian eigenstates play when considering general dynamics – a theme we shall return to later.

In this thesis we will indeed mostly consider the simulation of finite dimensional, pure states with time-independent Hamiltonians, since these are most relevant for the idealised models of *closed* interacting quantum systems. We note that even in this 'simplified' regime, which nonetheless must in some way underpin open and driven dynamics, there is certainly no lack of open questions.

Before discussing what criteria an analog quantum simulator needs to fulfil in

order to accomplish the above, it is worth mentioning *digital* quantum simulation. As shown by Seth Lloyd in 1996 [18], it is possible to efficiently perform a quantum simulation on a digital quantum computer, for physically reasonable Hamiltonians — by which we mean finite dimensional systems with local interactions such that the Hamiltonian can be written as:

$$\mathbf{H} = \sum_{i=1}^{N} h_i \tag{1.8}$$

where each local term h_i only acts non-trivially on a bounded number of particles, say k. This type of Hamiltonian encompasses the vast majority of physical models. The essential argument is to break down the exponential needed for the unitary evolution operator $\mathbf{U}$ from Eq. (1.5) using the Suzuki-Trotter expansion:

$$\mathbf{U}(t) = \lim_{n\to\infty} \left(\prod_i e^{-ih_i \frac{t}{n}} \right)^n . \tag{1.9}$$

Since Lloyd considered a quantum computer with a universal set of gates, for each time step $\Delta t = \frac{t}{n}$, the k-local exponential terms $e^{-ih_i \Delta t}$ can always be applied individually using discrete gates in an error corrected fashion. The overall error accumulated originates from the discarded terms in the expansion, which can be arbitrarily reduced using smaller time-steps (for the first order Suzuki-Trotter expansion above the error reduces as Δt^2), or by using a higher order expansion with correspondingly more complex local exponential gates. However, although this is a very powerful scheme, the requirements are essentially as demanding as for a full quantum computer. Thus the quantum simulators we refer to here are implicitly analog, and this is the case in most current literature.

So far, we have mentioned that a simulator must be governed by approximately the same Hamiltonian as its target, with continuous dynamics, and have a degree of controllability. To expand on these conditions it is worth roughly sketching the five criteria proposed by Cirac and Zoller [13], in analogy to DiVincenzo's criteria for a quantum computer [19], which clarify to an extent the non-implementation specific details that a quantum simulator needs to satisfy.

1. *Quantum System:* It should be some isolated system of quantum particles, with many degrees of freedom, and able to sustain large-scale entanglement. This ensures both that a simulator will be applicable to real quantum models, and also that it is useful in the sense of performing 'calculations' beyond any classical computer.

2. *Initialisation:* It should be possible to prepare the simulator in a known state — pure states in particular being desirable.

3. *Hamiltonian Engineering:* Although a simulator will likely be limited to a particular class of models, the interactions and parameters within that model should be tunable. This is particularly important in the context of exploring the phases of a model.

4. *Detection:* A simulation if of course only useful if one can measure quantities after and possibly during its operation. Often the available methods of measurement will be limited, for example, only local or only collective properties might be accessible.

5. *Verification:* There should be methods available to verify the operation of the quantum simulator. In the end, and by definition, a useful run of a quantum simulator will not be able to be completely verified by classical means. However, the constituent parts should be verifiable, as should certain limiting cases of known models. For small sizes and time-scales, numerical methods should serve as a comparison as well.

We can see from these criteria already that 'large scale' entanglement is expected to play a part, and so we move on to introducing that next.

1.2 Entanglement

An increasing trend in the study of many-body quantum models is the application of ideas from quantum information [20, 21, 22]. Much of this is related to moving from classical notions of correlations to quantum entanglement, and in one sense,

entanglement can be thought of as 'super-correlation'. More specifically, entanglement is generally defined for a bipartite mixed state, ρ_{AB}, as the inability to describe a state as the sum of tensor product states over each subsystem:

$$\rho_{\mathrm{AB}} \neq \sum_i p_i \left(\rho_i^{\mathrm{A}} \otimes \rho_i^{\mathrm{B}}\right) \tag{1.10}$$

where p_i denote the weights or probabilities of each state. If Eq. (1.10) is satisfied we call the state *separable*. While this definition is quite clear, when it comes the *quantifying* the amount of entanglement present the story becomes trickier.

One case that is simple however is that of bipartitions of pure states, for which $\rho_{\mathrm{AB}} = |\psi_{\mathrm{AB}}\rangle \langle\psi_{\mathrm{AB}}|$. Here, once can compute the reduced density operator:

$$\rho_{\mathrm{A}} = \mathrm{Tr}_{\mathrm{B}}(|\psi_{\mathrm{AB}}\rangle \langle\psi_{\mathrm{AB}}|) \tag{1.11}$$

and then simply find its Von Neumann entropy:

$$S[\rho_{\mathrm{A}}] = -\mathrm{Tr}\left(\rho_{\mathrm{A}} \log_2 \rho_{\mathrm{A}}\right) \ . \tag{1.12}$$

Since in this case the entropy can only be generated through entanglement between A and B, it naturally quantifies the total entanglement between them. Indeed, for pure bipartitions, all entanglement measures are monotonically related to this *entanglement entropy* (though they don't necessarily reduce to the same value).

Instead, for mixed states there are a range of operationally defined measures which quantify how much of a helpful resource, such as singlet pairs, can be extracted [23]. These are generally hard to actually compute, particularly for many-body states, since they often are defined as a maximum over operational procedures – a tricky optimization problem for generic states. Similarly, entanglement measures defined relative to the 'closest' separable state must also be computed using an optimization – making them unwieldy.

For these reasons, in this thesis we largely stick to working with the *logarithmic negativity* [24, 25, 26, 27], which can be calculated in a straightforward manner

and indeed is known rigorously to be one of the only measures for which that is possible [28]. It is defined as:

$$\mathcal{E}(\rho_{\mathrm{AB}}) = \log_2 |\rho_{\mathrm{AB}}^{T_{\mathrm{B}}}| \tag{1.13}$$

with $|\cdot|$ denoting the trace-norm, and $\cdot^{T_{\mathrm{B}}}$ the partial transpose. One can see with reference to Eq. (1.10) that in the separable case the partial transpose has no effect since ρ_i^B is Hermitian, the trace norm thus remains equal to one and we find $\mathcal{E} = \log_2(1) = 0$. We note that while there are so-called 'bound-entangled' states for which the logarithmic negativity is still zero, these are not expected to be especially important in the many-body setting – despite still being useful for quantum information processing tasks [29].

One possible conceptual picture of entanglement is that of information encoded non-locally, that is to say, *only* retrievable using a measurement that acts globally on the whole system. A naive quantity that one might then expect to quantify entanglement is the *mutual information*, defined as:

$$I(\rho_{\mathrm{AB}}) = S[\rho_{\mathrm{A}}] + S[\rho_{\mathrm{B}}] - S[\rho_{\mathrm{AB}}]\,. \tag{1.14}$$

However the difference between $I(\rho_{\mathrm{AB}})$ and $\mathcal{E}(\rho_{\mathrm{AB}})$, for example, is that the mutual information quantifies the *total* amount of correlation – be it classical or quantum – and thus also includes a contribution from information that *is* available locally. In order to illustrate the various points made above, in Fig. 1.1 we generate many random two-qubit states of varying purity, $\mathrm{Tr}(\rho_{\mathrm{AB}}^2)$, and plot the logarithmic negativity against the mutual information. We note that for all the pure states, there is indeed a monotonic relationship between the entropy and the logarithmic negativity. However, although there is a reasonable correlation between the mutual information and logarithmic negativity, there are many states with significant classical correlation but essentially no entanglement – as quantified by $\mathcal{E}$. For this reason we expect there to be fundamentally different aspects of many-body systems revealed by using a genuine entanglement measure.

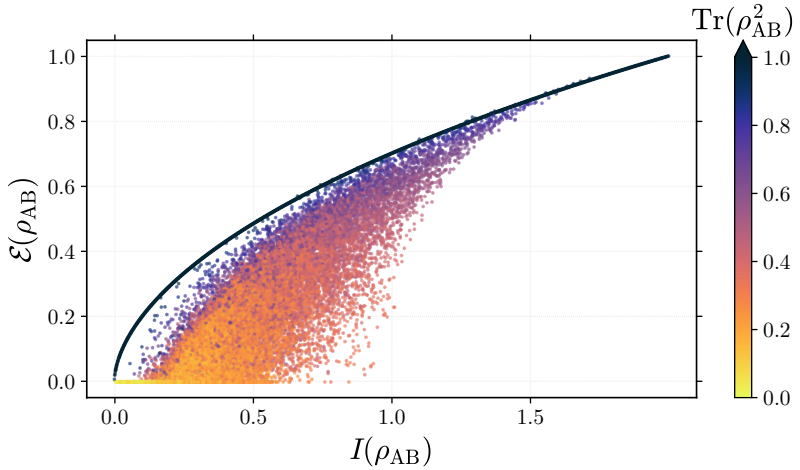


Figure 1.1: Comparison of the logarithmic negativity and the mutual information for randomly generated two-qubit states. The purity of each random instance is mapped to its colour.

In terms of moving from the canonical case of two qubits to many-bodies, we note that the logarithmic negativity is agnostic with regard to the underlying Hilbert space sizes, unlike the concurrence for example. For a pure many-body state, $|\psi\rangle$, to be 'highly entangled', as desirable in a quantum simulator, loosely speaking we would expect the entanglement to be high across any bipartition. Although the entanglement entropy is sufficient in that scenario, it is when we look at bipartite subsystems *within* the pure state,

$$\rho_{\mathrm{AB}} = \mathrm{Tr}_{\mathrm{C}} \left(|\psi_{\mathrm{ABC}}\rangle \langle \psi_{\mathrm{ABC}}| \right) \ , \tag{1.15}$$

that the need for a genuine entanglement measure emerges. Being able to ignore an environment, C, in this way allows one much greater flexibility when probing a many-body state, as we will later demonstrate. Note that the many-body entanglement we refer to is distinct from *multipartite* entanglement, though this can be defined easily by considering combinations of bipartite entanglements [30].

The study of entanglement in many-body systems has been crucial in understanding, for example, topological systems [31], where long range entanglement is an essential description of the different phases of matter that can occur. Many early studies began by considering the block-properties of entanglement in interesting

models [32, 33, 34] and this is still a recurring theme. More recent examples of using of other, slightly different quantum information measures include characterizing quantum phase transitions, for example using the quantum discord [35] or Schmidt gap [36]. The relationship of entanglement to entropy in field theories has also particularly flourished in recent years due to the development of the 'replica trick' for analytically calculating both the entanglement entropy and the logarithmic negativity [37, 38, 39]. This bridging of fields has also been fruitful the other way round, with quantum error correcting codes and topological quantum computing [40] both strongly dependent on input from the condensed matter community.

1.3 Many-Body Quantum Models

With these notions in place we move on to describing several canonical models from many-body quantum physics that strongly motivate the development of fully fledged quantum simulators and will act as reference points throughout the later chapters.

1.3.1 Hubbard Model

There is a vast range of possible models that motivate quantum simulation, but we introduce the two most common lattice models, which nonetheless provide a very rich range of behaviour. Firstly, the Hubbard model [41, 42], which aims to describe particles hopping in a periodic potential while interacting with each other. The model can be formulated for both bosons and fermions, but we show the latter, spin-$\frac{1}{2}$ version, since this applies to electron transport and thus has most relevance for the simulation of real materials and molecules. The Hamiltonian is written for N sites as:

$$\mathbf{H}_{\text{Hubb}} = -t \sum_{\langle i,j\rangle,\sigma} \left(f^{\dagger}_{i,\sigma} f_{j,\sigma} + f_{i,\sigma} f^{\dagger}_{j,\sigma} \right) + U \sum_{i=1}^{N} n_{i\uparrow} n_{i\downarrow}. \tag{1.16}$$

Let us break down the parts:

- The first term accounts for hopping between sites with tunnelling energy t, the sum running over nearest-neighbours $\langle i, j\rangle$ and spin states $\sigma = \{\uparrow, \downarrow\}$, with $f^{\dagger}_{i,\sigma}$ ($f_{i,\sigma}$) the fermionic creation (annihilation) operator acting on site i and spin state σ.

- The second term accounts for an on-site interaction, of strength U, with $n_{i\sigma} = f^{\dagger}_{i,\sigma} f_{i,\sigma}$ the occupation number of site i with spin state σ.

We note that the Bose-Hubbard model can be formulated in very similar terms – as the sum of kinetic and interaction contributions – with a couple of key differences. Firstly, unless one adopts *hardcore* bosons, there is no effective repulsion from the exclusion principle that limits the number of particles per local state. Secondly, the fermionic operators must satisfy *anti*-commutation relations on different sites, introducing a negative sign upon interchange of particles.

Despite the fact that Eq. (1.16) is possibly the simplest model we can formulate of interacting fermions, it displays a rich variety of behaviour, much of which been directly observed. For example, in two-dimensions and for small attractive U (< 0), its groundstate exhibits a crossover from a Bardeen-Cooper-Schrieffer superfluid to a Bose-Einstein Condensate [43]. For repulsive U (> 0), it instead can form a Mott insulator [44], or exhibit antiferromagnetic ordering with an effective exchange coupling $J_{\text{ex}} = \frac{4t^2}{U}$. However, depending on the choice of geometry and parameters, the full phase diagram of Eq. (1.16) is far from being completely solved analytically – though for one-dimensional systems is can be solved using the Bethe ansatz [45].

A few extensions of the model are notable in terms of quantum simulation for their possible practical relevance. With the addition of spin-interactions, yielding the $t - J$ model, this kind of system is a strong candidate for capturing the essential mechanism of high-temperature superconductivity [46, 47]. If we instead treat the 'sites' as molecular orbitals, by adding a two-electron hopping term we yield an effective model relevant for quantum chemistry [48].

1.3.2 Heisenberg Model

Another crucial lattice model to consider is the Heisenberg model [49] and its variants, which describe magnetic systems made up of lattices of spins. The Hamiltonian here, for N sites, is given by:

$$\mathbf{H}_{\text{Heis}} = \sum_{\langle i,j \rangle} \left(J_x S_i^x S_j^x + J_y S_i^y S_j^y + J_z S_i^z S_j^z \right) + B \sum_{i=1}^{N} S_i^z \tag{1.17}$$

where again the first sum runs over nearest neighbours, and the spin operators acting on site i for spin-half are given in relation to the Pauli operators, $\{X, Y, Z\}$, by:

$$S_i^x = \frac{1}{2}X_i \,, \quad S_i^y = \frac{1}{2}Y_i \,, \quad S_i^z = \frac{1}{2}Z_i \,. \tag{1.18}$$

For the second term B denotes the strength of an external magnetic field, here in the z-direction.

If we take $J_x = J_y = J_z = J$ then we recover the standard isotropic Heisenberg model, with antiferromagnetic order induced when $J > 0$, and ferromagnetic order when $J < 0$. Common variants of Eq.(1.17) include: (i) the Ising model, for which $J_x = J_y = 0$ and the global magnetic field is often taken in the 'transverse' X-direction; and (ii) the XY-model, for which $J_z = 0$. In its unmodified form, the Heisenberg model is analytically tractable using the Bethe ansatz [50], though this situation is complicated by the introduction of longer range interactions, or locally varying coupling strengths and magnetic fields.

Although notionally describing quite different underlying physical systems, a key aspect of Hubbard-like and Heisenberg-like models is that they can be *mapped* into each other and thus in fact share certain fundamental features. A good example of this is the Jordan-Wigner transformation [51], which maps spins to a system of spinless (unlike Eq. (1.16)) fermions. The initial step is to define raising and lowering spin operators acting on site j as:

$$S_j^+ = S_j^x + iS_j^y \tag{1.19}$$

$$S_j^- = S_j^x - iS_j^y \,. \tag{1.20}$$

These superficially look like good creation and annihilation operators, however, the problem with these two is that they commute on different sites. In order to reproduce the correct fermionic anti-commutation relations Jordan and Wigner introduced additional 'string' operators that produce the correct phase based on the occupancy of fermions to the 'left' of the creation/annihilation operators – in one dimension. If we also associate the z-component of spin with the fermionic occupation operator,

$f_j^\dagger f_j = n_j$, we get the full mapping

$$S_j^z = f_j^\dagger f_j - \frac{1}{2} \tag{1.21}$$

$$S_j^+ = e^{+i\pi \sum_{k<j} n_k} f_j^\dagger \tag{1.22}$$

$$S_j^- = e^{-i\pi \sum_{k<j} n_k} f_j \ . \tag{1.23}$$

Although the non-local string operators are cumbersome to deal with generally, for specific versions of Eq. (1.17) these phases cancel and a simple Hubbard-like model emerges. One such example is the XY-model, which in fact maps to the free, spin-less version of Eq. (1.16). It is in this sense that the XY-model (and therefore also the Ising model) is often referred to as *not* being a true interacting, many-body model, despite notionally consisting of many interacting particles. With reference to digital quantum simulation, it is also worth noting that a generalization of the Jordan-Wigner transformation is what allows fermionic systems to be efficiently simulated on quantum computers with non fermionic components [52]. In this thesis we do in fact work mostly with spin chains, but it is important to keep in mind that many phenomena, such as many-body localization, are just as relevant for Hubbard-like models.

As hinted at, for both Hubbard-like and Heisenberg-like models, the geometry and dimension of the underlying lattice plays an very important role in determining the properties of the system. By introducing varying or random couplings and interactions per-site, frustrations can emerge in these models which make them even harder to treat analytically. Both of these facts motivate quantum simulators as an important tool for exploring their rich behaviour when trying to understand material-orientated properties.

1.4 Many-Body Quantum Phenomena

Beyond understanding of real materials, there are many fundamental questions that quantum simulators would likely shed light on [53, 54], and in this section we briefly summarise three topics that appear throughout the later chapters – quantum

phase transitions, questions of equilibration and thermalization, and many-body localization.

1.4.1 Quantum Phase Transitions

The many phases of the Hubbard and Heisenberg models were alluded to, and a key concept of interest is the nature of quantum phase transitions [55] between them. Unlike classical phase transitions, these occur at zero temperature, as some parameter of the Hamiltonian is varied, triggering a sudden change in the character of the ground-state. Specifically, if a Hamiltonian, $\mathbf{H}(g)$, is parametrized by a scalar g, then we expect its groundstate, $|\psi_0\rangle$ for some range of g to have similar entanglement properties. If however we approach some critical point g_c, delimiting two phases, we expect excited states of the Hamiltonian to approach the energy of the groundstate. At this point one of these excited states might become the groundstate, or an avoided level crossing might occur. In either case, the nature of the groundstate's entanglement structure at that point is drastically changed. A canonical example is the transverse Ising model as we vary the external transverse magnetic field B_x. Additionally, the spin-half Heisenberg chain with $J_x = J_y = J_z$ in fact sits at a critical point.

Mapping where these transitions occur in the parameter space of a model is one primary aim of quantum simulators. This fact, and that the defining features of a phase are generally global quantities, drives the statement that quantum simulators can tolerate a moderate degree of error, since if one were simply identifying phases, small local errors would probably not be able to drive the system into a different phase. At the critical point between phases, the state of the system is of particular interest, with no universal framework for understanding the structure and dynamics of the system present yet [56]. Specifically, Landau's theory of symmetry breaking and universal scaling exponents does not seem to be applicable to all critical points. Again, it is hoped that analog quantum simulators will allow a deeper exploration of these systems.

1.4.2 Equilibration & Thermalization

Another key area of interest is the relationship between initially out of equilibrium quantum systems and the emergence of statistical mechanics – a process as yet not completely understood. By inspecting Eq. 1.5, it is clear that globally an initial state in a closed system undergoes very constrained evolution. Specifically, for time-evolved state $\rho(t) = |\psi(t)\rangle \langle\psi(t)|$, the matrix-elements for two different energy eigenvectors $|i\rangle$ and $|j\rangle$ evolve:

$$\langle i| \rho(t) |j\rangle = e^{-i(E_i - E_j)t} \langle i| \rho(0) |j\rangle \tag{1.24}$$

but the diagonal populations ($i = j$) of each are static, since $e^{-i(E_j - E_j)t} = 1$ and thus:

$$\langle j| \rho(t) |j\rangle = \langle j| \rho(0) |j\rangle \ . \tag{1.25}$$

The system will also return arbitrarily close to its initial state given a long enough (albeit likely enormous) amount of time. These facts already hint at the gap between starting with an arbitrary quantum state and an 'arrow-of-time'-like decay to the thermal, Boltzmann state that we expect. Work has focussed on considering *subsystems* of pure quantum states, such that globally the system can act as its own bath. From this point a number of questions can be formulated in order to subdivide the problem. Firstly, do subsystems *equilibrate* to a particular state and stay there for almost all time? Substantial progress has been made in showing that this is the case for generic systems [57, 58, 59, 60]. A second important question is what do the equilibrium states look like? This is the question of *thermalization* [61, 62]. Statistical mechanics tells us that they should lose all information about their initial conditions, instead taking on solely global thermal properties such as temperature. The perspective of entanglement here is important, as it offers a way of understanding how this might happen. Specifically, we can think of entanglement as being information describing a physical state that is *not* locally accessible. If entanglement has grown, or de-localized, to very large scales, as generically happens with interacting many-body systems, then it has essentially been lost, despite being technically retrievable.

Another important concept to mention here is the Eigenstate Thermalization Hypothesis (ETH) [63, 64, 65, 61, 66, 67, 68] - which posits some constraints on the spectral properties of a Hamiltonian if it is to demonstrate equilibrium properties given by statistical mechanics. Put most simply, ETH requires that subsystems of *individual* eigenstates match thermodynamic quantities. Although it has led to understanding of what classes of model do thermalize, there are now known several examples of when if fails. One of these such cases is that of *integrable* systems. Although there is some debate as to a precise definition, integrability might loosely be defined as the appearance in certain models of an extensive number of local quantities that commute with the Hamiltonian [69]. These conserved quantities prevent relaxation to a standard thermal state, forcing one instead to introduce a generalized Gibb's ensemble [70], that represents the maximum entropy state for these given symmetry constraints. It should also be noted that weakly perturbed integrable systems do still in fact thermalize, albeit on a much longer time-scale and after an initial relaxation to a 'pre-thermalized' plateau [71, 72].

1.4.3 Many-Body Localization

A topic that has garnered much interest recently is that of many-body localization (MBL), which in fact spans particularly the topic of thermalization but also quantum phase transitions. Broadly speaking, many-body localized systems are interacting many-body systems with quenched randomness (i.e. inherent to the Hamiltonian) that robustly resist thermalization – which is to say that even if perturbed, it is expected that MBL systems will still never thermalize. The dynamical nature of thermalization implies that MBL is is a feature imprinted across all eigenstates, which moreover, cannot obey the ETH. What is also unusual is that it seems, by varying the strength of the quenched randomness, that there is an eigenstate phase transition from thermalizing to MBL behaviour.

Anderson first proved that randomness could have the effect of localizing the eigenstates of single particles [73], however it was thought for a long time that this phenomena could not survive in the presence of interactions. It was not till much more recently that this thinking was overturned [74]. Evidence then

emerged that, unlike in the single particle case, there was a transition to this non-thermalizing case [75] and for one-dimensional systems the phenomena is now rigorously proved [76, 77]. A recent review can be found in [78].

With reference to the Hubbard and Heisenberg models introduced earlier, MBL can be induced by including an on-site random energy modified by overall random 'strength' h. The critical value of h for MBL to set in is still an open question, but thought to be in the range $h_c = 3.7 - 5.0$ [79, 2]. Deep in the MBL phase, the structure of eigenstates is now quite well understood, with the emergence of quasi-local integrables of motion [80, 81, 82, 83, 84, 85], leading to logarithmically slow growth of entanglement. However, unlike standard integrable systems, it is not clear that the generalized Gibb's state approach is valid for MBL models. Additionally, the nature and location of the transition, as well as whether MBL survives in higher dimensions, all remain as significantly open questions.

MBL serves as a very useful probe of how whatever mechanism causes thermalization breaks down, and one that quantum simulators can realize. Moreover there is a fundamental question about how one engineers the robust 'insulation' of quantum information and whether this can be harnessed practically for quantum memories.

1.5 Entanglement in Classical Simulations - Tensor Networks

Understanding the entanglement structure of states has also led to huge advances in numerical techniques for classically simulating many-body systems, the prime example being tensor network algorithms. We first recall that for most numerical schemes, just storing the full description of any state requires exponentially many coefficients. This poses a problem even before one tries to factor in the computational cost of operations such as diagonalisation However, in recent years, it was realized that many states satisfy an *area law* [86, 87, 88, 89, 90, 91, 92]. This posits that any subsystem of a state has an entropy that scales as its *boundary* rather than its volume. Using this assumption allows one to form a compressed representation of only the 'natural' states that satisfy it. The requirements for it to be 'natural' are that

the Hamiltonian is local, as in Eq. 1.8, and for the groundstate to be gapped, which technically rules out for example, critical systems. An intuitive picture of why area laws work is that if only local entanglement exists, only a constant depth ring around the subsystem's border can be entangled with the outside environment.

The representation that picks out only these states from the whole Hilbert space is that of tensor networks [93, 94, 95] — in their simplest form for 1D systems matrix product states (MPS). An important connection to make was that MPS underlie [96] the enormously successful density matrix renormalization group (DMRG) methods introduced by Steven White [97, 98]. Subsequent work expanded [99, 100] MPS into the much more general framework of tensor networks, complete with a very concise and intuitive graphical representation that we shall introduce later in Chapter 5. This tensor diagram notation is also notable for naturally encoding many quantum information identities that are otherwise laborious to work with, a fact we also briefly demonstrate in Chapter 4.

The key idea of tensor networks is to simply treat a finite wavefunction not as a vector but as a tensor, with each index specifying a subsystem:

$$\psi_i \to \psi_{a,b,\ldots,z} \ , \tag{1.26}$$

and to then make use of various factorizations of this tensor, guided by the expected entanglement structure.

For example if one partitions these indices into left and right groups, one can then perform a singular value decomposition (SVD) on the now generically rectangular matrix. The resulting form is identical to the Schmidt decomposition:

$$|\psi\rangle = \sum_i \lambda_i \, |L_i\rangle \, |R_i\rangle \tag{1.27}$$

with $|L_i\rangle$ ($|R_i\rangle$) forming an orthonormal basis for the left (right) subsystems. This decomposition is so far exact, but by *truncating* the sum to a maximum of χ Schmidt values λ_i, which represent the entanglement spectrum across the cut, the number of parameters stored representing the state can be drastically reduced. To see this

consider that the entanglement entropy of the left or right subsystem is given by

$$S[\rho_L] = S[\rho_R] = -\sum_i^{\chi} \lambda_i^2 \log_2 \lambda_i^2 \,, \tag{1.28}$$

which is upper bounded by $\log_2 \chi$. In the case of gapped one-dimensional systems, the 'area' of a subsystem's boundary is simply constant, and thus the entanglement entropy across any cut is also bounded by a constant with respect to system size. The implication is thus that we can keep χ *independent* of the size of L or R whilst still accurately describing the total state.

One can perform this procedure for every local subsystem, yielding a network of tensors that is not only compact in terms of storage but efficient to calculate, for example, operator expectation values. Of course, this is usually not the way one constructs such states since it initially requires a full state — instead the tensor network structure is used as a variational ansatz with which to optimise for the lowest energy of a Hamiltonian for example. It is important to note two here distinct aspects of tensor networks: (i) the ability to efficiently *represent* quantum states; and (ii) the ability to efficiently *compute* quantities such as the energy with said representation. While in one-dimensional systems both these criteria are generally satisfied, in two or more dimensional systems the latter criterion is often broken.

Two tensor network algorithms are worth briefly describing in this introduction due to their widespread usage. Firstly, a one-dimensional groundstate search using the aforementioned DMRG. Here, each local site is optimized individually, where the actual Hamiltonian (in matrix product operator form) and the 'environment' of all other sites act *together* as an effective local Hamiltonian. One then sweeps several times over all the sites, with this iterative minimization hopefully (and usually) leading to the groundstate. The second algorithm for a one-dimensional real or imaginary time evolution is time-evolving-block-decimation (TEBD), introduced by Vidal [99]. Here one makes use of the exact same Suzuki-Trotter expansion of Eq. (1.9), where one now only needs to act with the gate on the relevant local tensors rather than the full state. Additionally, after each application of a gate, the tensor

network ansatz must be locally decomposed back into MPS form using the SVD. Since gates generally introduce entanglement, one must let the bond dimension χ grow here, and in this way the entanglement explicitly limits the tractable times of TEBD to $\mathcal{O}(1)$.

This promotion of local entanglement to a starting point for states has allowed not only a swathe of new numerical simulations, but among other things the analytical classification of gapped quantum phases [101]. It has also cemented the idea that what makes quantum computation powerful is long-range entanglement, since otherwise it could be efficiently simulated using tensor network algorithms. Thus, in the context of quantum simulation, this tells us that those states and systems which are useful to simulate are those which have more than just local entanglement. Two examples of such systems are those introduced in Sec. 1.4.1 and Sec. 1.4.2, namely critical systems and out of equilibrium systems. Systems at a critical point are gapless, and expected to logarithmically violate the area law, this can be seen partly as a result of the diverging length scale that emerges in such systems. Out of equilibrium systems, on the other hand, inherently involve many excited states. In other words, they have significant support on mid spectrum eigenvectors which in general support a *volume law*. On top of these cases, two or more dimensional tensor networks, despite still being relatively compact *representations*, encounter problems with scaling when quantities such as expectation values need to be actually calculated.

To conclude this section, tensor networks represent the state of the art for classical simulation of states with local entanglement. Using them carefully, nearly all 1-dimensional equilibrium many-body quantum systems are amenable, however, this leaves a wide class of systems for which quantum simulation seems to be a necessity, fundamentally due to long-range entanglement.

1.6 Experimental Realizations

Finally we review the considerable amount of experimental progress in realising quantum simulators thus far. Optical lattices with ultra cold atoms [102, 103]

have provided some of the most striking simulations so far and already driven understanding of several many body systems. Here, a very perfect lattice is created using counter propagating laser beams, which trap atoms by inducing, and then interacting with, a periodic dipole moment. The use of lasers allows a great deal of freedom in choosing the geometry and depth of the lattice potential. Additionally to being able to tunnel from site to site, many interactions between atoms in the same and neighbouring sites are possible, and bosonic or fermionic species can be chosen for a given statistics.

Arguably the first many-body quantum simulation was performed using cold atoms in optical lattices, exploring the transition between a superfluid and Mott insulator [104]. More recently, by tuning off the tunnelling, an anti-ferromagnetic spin chain has been realised [105], and MBL has also been explored [106]. Although they can scale to hundreds of atoms and enjoy long coherence times, there are a number of limitations to optical lattice simulations. Firstly, individual measurement of the atoms is difficult due to the small size of the lattice, governed by the necessarily small laser wavelength. This limits detection and verification schemes. Secondly, although a wide number of models are realizable, often the energy ranges that are required impose very slow dynamics on the system.

Another very successful system for quantum simulations is that of trapped ions [107, 108, 109]. Here the charged nature of the ions allows very strong, long-range Coulomb interactions to occur. The ions also exhibit very long coherence times and these two facts together have produced some of the highest fidelity quantum gate operations. Disadvantages include the difficulty in scaling to many ions or complex geometries. On top of this, the long range and strength of the inter-ion interactions make it hard to simulate models with *only* nearest neighbour interactions.

Other systems worth mentioning include superconducting circuits [110, 111, 112], and NMR addressed spins [113, 114, 115], both of which naturally simulate Ising-like models, and photonic systems [116], which however are constrained by the ability to engineering couplings.

Finally, there are many strong arguments for pushing towards *solid-state* simula-

tors, though the noisy nature of such systems has limited progress so far. Specifically, if one is to simulate electrons as with the Hamiltonian in equation Eq. (1.16), then a natural choice for the components of your quantum simulator is electrons themselves, rather than trying to mimic their itinerant, spin-half characteristics artificially. Another advantage would be input and technology from the existing and huge industry in silicon chip manufacturing for example. For these reasons *quantum dot arrays* are the starting point for the next chapter.

Chapter 2

Unravelling Quantum Dot Array Simulators

This chapter is adapted from the publication:

- Johnnie Gray, Abolfazl Bayat, Reuben K Puddy, Charles G Smith, and Sougato Bose. Unravelling quantum dot array simulators via singlet-triplet measurements. *Phys. Rev. B*, 94(19):195136, 2016.

2.1 Abstract

Recently, singlet-triplet measurements in double dots have emerged as a powerful tool in quantum information processing. In parallel, quantum dot arrays are being envisaged as analog quantum simulators of many-body models. Thus motivated, we explore the potential of the above singlet-triplet measurements for probing and exploiting the ground-state of a Heisenberg spin chain in such a quantum simulator. We formulate an efficient protocol to discriminate the achieved many-body ground-state with other likely states. Moreover, the transition between quantum phases, arising from the addition of frustrations in a $J_1 - J_2$ model, can be systematically explored using the same set of measurements. We show that the proposed measurements have an application in producing long distance heralded entanglement between well separated quantum dots. Relevant noise sources, such as non-zero temperatures and nuclear spin interactions, are considered.

2.2 Introduction

Quantum simulators [117] are one of the hotly pursued topics of current quantum technology research. Analog quantum simulators directly mimic another physical quantum system in order to explore its behaviour in greater depth. In doing so, they provide a wide range of applications, for instance, addressing challenges in smart material design which could potentially revolutionize medicine and energy provision in the future. While already accessible, quantum simulators will scale to much larger sizes in the near future, in doing so becoming a significant technological step on the path to full quantum computation. A key question for such simulators is the certification of the states realized within them. For example, simple questions such as whether the state is a genuinely quantum, pure and entangled many-body state need to be answered with available measurement schemes. For an experimentalist who has realised a candidate state it is crucial to discriminate it from the closest classical counterpart (e.g. the Neel state for antiferromagnets), random, thermal and energetically proximal quantum states. Here we address the question with respect to the emerging field of solid state quantum simulators [118, 119, 120, 121, 122, 123].

So far, neutral ultra-cold atoms [103] and trapped ions [108] have been predominantly exploited for serving as quantum simulators thanks to their high controllability and long coherence times. Nevertheless, in order to simulate solid state systems, the presence of both particle hopping and long-range charge interactions are needed, and these are not not readily available in trapped ion and cold atom systems respectively. Additionally, the spin exchange couplings realised in these systems tend to be small, such that any dynamics take place over long ($\sim$ ms) time-scales. It is therefore timely, thanks to recent advances in fabrication of quantum dot arrays [124], to think about a real solid state quantum simulator. These advances have largely been fuelled by seminal work of Loss and DiVincenzo [125], who proposed single electron spins as qubits. Such quantum dot arrays have also been proposed for quantum state transfer [120] and adiabatic many-body state preparation [121]. A two-site quantum Hubbard model has been successfully simulated with dopant atoms in silicon [122], which are qualitatively equivalent to quantum dot arrays as far as their prospects for

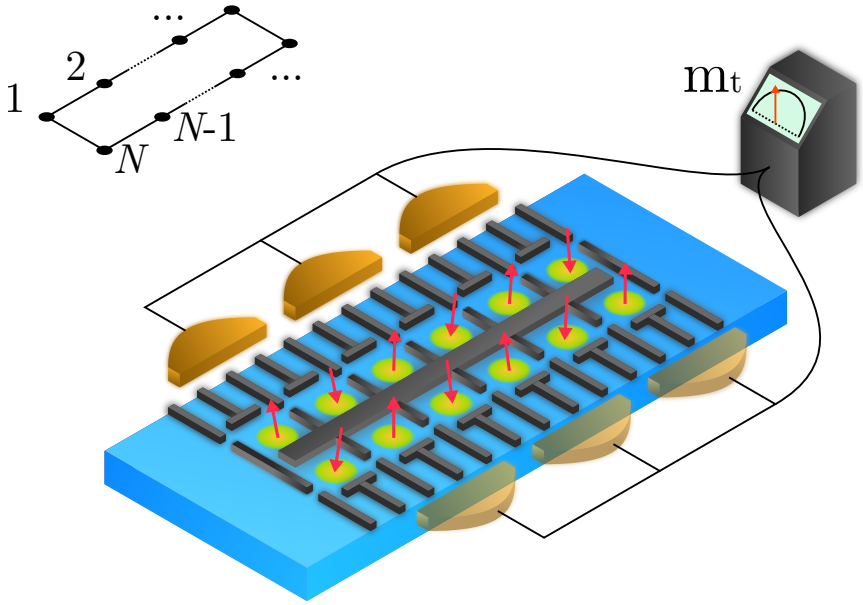


Figure 2.1: Quantum dot array spin model simulator and triplet profile readout: A simplified schematic of a quantum dot array simulating a Heisenberg $N = 12$ closed chain with Singlet-Triplet measurements. Here, grey bars represent voltage gates, each arrow a single confined electron, and the gold detectors Singlet Triplet measurements. These are simultaneously performed such that the total number of triplets present, m_t is recorded.

quantum simulations are concerned [123]. Unlike cold atoms and ions, quantum dot arrays naturally have more types of interaction, such as spin-orbit [126], and thus can simulate a wider range of interactions. Moreover, their compactness allows for stronger interactions resulting in faster operations. Nevertheless, there are still challenges worth mentioning: (i) there are also strong interactions between the electrons and the environment (such as proximal nuclear spins) which decohere the simulator, and (ii) the small scale of the fabrication, and the required number of gates, makes it currently difficult to scale up to complex arrays.

Recently, Singlet-Triplet (ST) measurement in double quantum dots has emerged as the dominant tool for spin information readout. Originally this was achieved through charge measurements [127], motivated by decoherence free singlet-triplet qubits [128]. Radio Frequency (RF) reflectometry has since emerged as the primary method of accomplishing this [129, 130, 131, 132, 133, 134]. The same

measurement tool now extends beyond double dot systems to donor-dimers [135]. These measurements discriminate between only the singlet state, and the remaining Bell-states. Nevertheless, it is known that these measurements, in combination with particular initial states, are sufficient for universal quantum computation [136]. The convenience and popularity of the ST-measurements motivate us to investigate their usefulness as a tool for probing and exploiting the many-body state realized in a quantum dot array.

Independent of the physical set-up, in order to verify the performance of a quantum simulator ideally one has to fully characterise the quantum state. The difficulty here is that by definition, a useful quantum simulator (i.e. with a large number of qubits) will have no exact, classically computable reference system. Additionally, full quantum state tomography requires an exponentially large number of distinct measurements [137, 138]. Recently, there have been proposals [139] for efficient tomography schemes which are applicable for those states satisfying a matrix product state ansatz, though one has to be able to perform complex multi-qubit unitary operations and measurements which are not necessarily available in the lab.

In this chapter, we consider quantum dot arrays simulating the ground-state of a Heisenberg spin chain. To characterise the state, we rely only on singlet-triplet measurements performed over nearest neighbour electron pairs, as has been experimentally demonstrated [129, 130, 131, 132, 133, 127, 140, 141]. This allows us to build up a probability distribution over outcomes that discriminates between our target state, i.e. the Heisenberg ground-state, and contaminated versions. In the presence of next-nearest neighbour interactions, realizable in recently developed multiplexed dot ladders [124], our setup can capture the quantum phase transition to a gapped, dimerized phase. Moreover, as another application, we show that the same set of measurements can be exploited to generate heralded entanglement between distant qubits. We investigate the performance of both applications under the influence of likely noise sources such as thermal fluctuations and hyperfine interactions with nuclear spins in the bulk.

The structure of this chapter is as follows: in Section. 2.3 we introduce the

model used to describe the system and the *triplet profile* that one can obtain from singlet-triplet measurements only. In Section. 2.4 we explore the possibilities of characterizing states using these measurements only, including a quantification of how *distinguishable* various states are from each other. We demonstrate that the quantum phase transition at $J_2/J_1 \sim 0.24$ for the $J_1 - J_2$ Heisenberg chain can be clearly observed. In Section. 2.6 we explore using singlet-triplet measurements only to *localize entanglement* between two ends of an open chain. In Section. 2.7 we investigate the effect of the two dominant noise sources in quantum dots — non-zero temperature and hyperfine interactions with proximal nuclei. Finally, in Section. 2.8 we propose a feasible experimental realization that could establish the validity of these methods.

2.3 Model

A key model in condensed matter physics is the Heisenberg Hamiltonian — used in many contexts including magnetism [142] and quantum phase transitions [55]. It describes the interaction between N spin-1/2 particles as

$$H_1 = J_1 \sum_{i=1}^{N} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1}; \tag{2.1}$$

where $\vec{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ is a vector of Pauli operators acting on site i, and J_1 represents the nearest neighbour spin coupling. We have assumed periodic boundary conditions, i.e. $\vec{\sigma}_{N+1} = \vec{\sigma}_1$, however, our analysis is equally applicable to open chain where increased dimerization makes the ground-state even more distinct. We set $J_1 = 1$ throughout the chapter, unless specified, considering it the energy scale of the system. This anti-ferromagnetic Heisenberg model has a unique SU(2) symmetric ground-state for even lengths, N, known as a global singlet since it has total spin $S = 0$. The lowest lying excitations are three degenerate 'triplet' states, with the energy gap to these closing as $1/N$ in the limit of large N.

In order to simulate the ground-state of the Heisenberg Hamiltonian in a controlled way we propose a quantum dot array with exactly one electron in each

quantum dot as schematically shown in FIG. 2.1(a). A similar structure has recently been realized for multiplexing quantum dots [124]. The spin sector of the interaction between the electrons is explained by the Hamiltonian (2.1) and the coupling J_1 can be tuned by applying appropriate gate voltages to the gates controlling the potential barrier between neighbouring electrons. By cooling this quantum system below its energy gap it can be initialized in its ground-state $|\psi_0\rangle$. The central object of interest in this chapter is $|\psi_0\rangle$ due to its highly entangled and non-trivial structure, described by a many-body global singlet, as well as its application for practical tasks in quantum technologies such as quantum state transfer [143, 144]. The first stage of verifying the operation of a quantum simulator is to characterize and certify its achieved state — hopefully the ground-state $|\psi_0\rangle$. Ideally this could be done using full quantum tomography [145] or other more efficient methods [139, 138] but for quantum dot arrays a current limitation is that only Singlet-Triplet (ST) measurements on adjacent sites are feasible. The question to be addressed here is to what extent characterisation and certification of a state is possible under this restriction.

The ST-measurement can be described by the following projectors

$$\begin{aligned} \mathbf{P}_s &= |\psi_-\rangle \langle\psi_-| , \\ \mathbf{P}_t &= \mathbf{1} - \mathbf{P}_s, \end{aligned} \tag{2.2}$$

where $|\psi_-\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ is the singlet with $|\uparrow\rangle$ and $|\downarrow\rangle$ representing spin up and down respectively. If the quantum simulator operates perfectly, i.e. initializes in the ground-state $|\psi_0\rangle$, then thanks to the SU(2) symmetry of the system the reduced density operator of any pair spin qubits will be a Werner state [146]

$$\rho = \alpha \mathbf{P}_s + (1-\alpha)\frac{\mathbf{P}_t}{3}, \tag{2.3}$$

with $0 \leq \alpha \leq 1$. In this sense ST-measurements are picked out as a preferred 'basis' for all SU(2) symmetric states.

Let's assume that the system is described by the density matrix ρ, ideally $|\psi_0\rangle \langle\psi_0|$. Performing ST-measurements on all $N/2$ consecutive pairs of spins, i.e.

qubits $(1,2)$, $(3,4)$, $\ldots$, $(N-1,N)$, results in $2^{N/2}$ different outcomes according to the singlet or the triplet output of each measurement. For example in a chain of length $N = 4$ any of the outcomes ss, st, ts or tt may occur with a certain probability. For any string of outcomes $x = x_1 x_2 ... x_{N/2}$ (with each x_i being s or t) the total projection operator is

$$\Pi_x = \bigotimes_{i=1}^{N/2} \mathbf{P}_{x_i}^{2i-1,2i} \tag{2.4}$$

where $\mathbf{P}_{x_i}^{2i-1,2i}$ are the same projectors as in Eq. (2.2) acting on qubits $2i-1$ and $2i$. Thus, the probability of getting the string x as the outcome of the measurements is $\mathrm{Tr}(\Pi_x \rho)$. For example the probability of getting the result $x = stts$ for a $N = 8$ state is $\mathrm{Tr}(\mathbf{P}_s^{12}\mathbf{P}_t^{34}\mathbf{P}_t^{56}\mathbf{P}_s^{78}\rho)$. We can further compress the number of outcome results by grouping together all result strings featuring the same number of measured triplets, thus creating a *triplet profile*:

$$p(m_t) = \sum_{x \in X_m} \mathrm{Tr}(\Pi_x \rho) \tag{2.5}$$

where X_m denotes the set of all result strings with *exactly* m triplet occurrences. This yields a concise characterisation of a state that is both easy to measure experimentally and to compute numerically. Performing the sum in Eq. (2.5) loses all information about how 'grouped' triplet excitations are, nevertheless, a surprising amount information can be gleaned from $p(m_t)$, including features heralding many-body entanglement. For example, one such feature that arises is that $p(m_t = 1) = 0$ for all global singlets. This arises from their spin-0 nature — they can have no overlap with the spin-1 subspace which includes all configurations of a single triplet. Indeed, under the reasonable restriction of translational invariance, classical states can only ever produce a binomial distribution for $p(m_t)$, and any deviations such as oscillations herald entanglement.

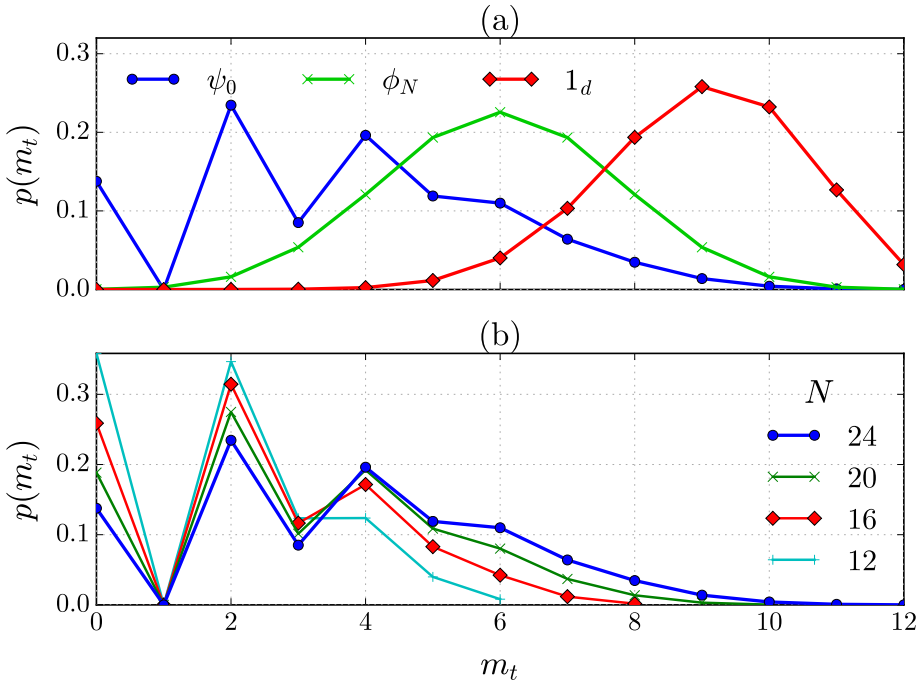


Figure 2.2: Discrimination of the ground, classical, and random states: (a) Triplet probability profiles for a number of states, namely the anti-ferromagnetic Heisenberg ring ground-state, ψ_0, the classical anti-ferromagnetic Neel state, ϕ_N, and the normalised identity, $\mathbf{1}_d$, all of size $N = 24$ (thick lines). **(b)** Scaling of the triplet profile with N for ψ_0.

2.4 Characterization of Simulator

In order to characterize the quantum state of the simulator we first calculate the full triplet profile $p(m_t)$ of the ground-state $|\psi_0\rangle$ and other likely states which may occur due to imperfections or malfunctioning of the quantum simulator. In particular, we consider the classical Neel state $\phi_N = |\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow \ldots\rangle$ and the maximally mixed state $\mathbf{1}_d$, which represents an infinite temperature thermal state. In FIG. 2.2(a) we plot the triplet profile $p(m_t)$ as a function of the number of triplet occurrences m_t for a chain of length $N = 24$ for all the three states. As can be easily calculated, ϕ_N and $\mathbf{1}_d$ are both characterised by binomial distributions centred on $\frac{1}{2}$ and $\frac{3}{4}$ respectively, whereas $|\psi_0\rangle$ produces a highly non-trivial oscillatory shape. For example, the zero-probability $p(m_t = 1)$ dip is very prominent, and also forms part of an oscillatory structure between odd and even occurrences of triplets.

The scaling of $p(m_t)$ for ψ_0 with size of system N is also shown in FIG. 2.2(b)

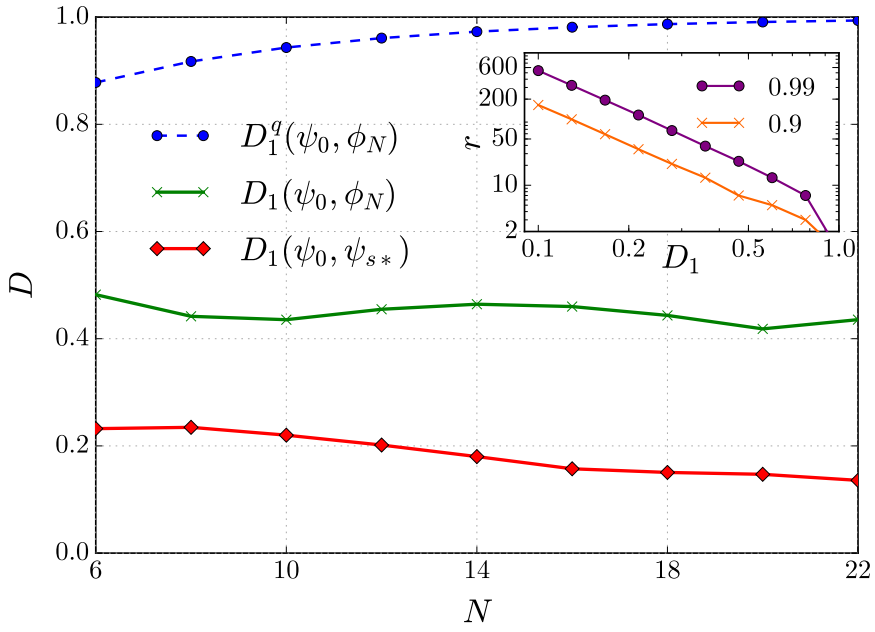


Figure 2.3: Comparison of the single shot distinguishability under an optimal measurement vs. triplet profile measurements: Scaling with system size of the single shot distinguishability between Heisenberg ground-state ψ_0, first excited singlet state ψ_{s*}, and the Neel state ϕ_N. Full lines denote distinguishability under triplet profile measurements, whereas dashed lines denote the full quantum distinguishability D_1^q. Note that $D_1^q(\psi_0, \psi_{s*})$ is not shown since it is always 1 — the states being orthogonal. **Inset:** number of repeat measurements, r, required to distinguish two states with probability (0.9, 0.99) for varying D_1.

— one can see that overall the features change slowly, with the average m_t increasing with N under the 'oscillating' envelope. As such, although the first 'fringe' contrast decreases with N slightly, the second increases and so on such that they should not be washed out in the thermodynamic limit.

Full quantum tomography is usually very demanding either in terms of sheer number of measurements or the complex many-body basis of such operations. Instead, we wish quantify the extent to which our ST-measurements can distinguish between likely quantum states (i.e. selected based on some prior intuition). A fundamental quantity here is what we shall call the single-shot-distinguishability, D_1, which quantifies the advantage a single measurement gives when guessing between two equally probably states such that the overall chance of success is $\frac{1}{2}(1 + D_1)$. If $D_1 = 0$ then the measurement yields no information at all about which state is

present, whereas if $D_1 = 1$ it perfectly discriminates them. If a measurement gives rise to two possible probability distributions, $p_1(a)$ and $p_2(a)$, over outcomes a then D_1 is given by [147]:

$$D_1 = \frac{1}{2} \sum_a |p_1(a) - p_2(a)| \tag{2.6}$$

which essentially formalizes the strategy of guessing whichever state is more likely to give result a_i each time. It has been shown [148] that for two quantum states that ρ and σ the maximum distinguishability is given by:

$$D_1^q = \frac{1}{2} \|\rho - \sigma\|_{\text{tr}}, \tag{2.7}$$

where $\|A\|_{\text{tr}} = Tr(\sqrt{AA^\dagger})$ is the trace norm. It is worth mentioning that the optimal measurement needed to yield D_1^q is likely to be a globally entangled projective measurement that is again not feasible.

An important aspect to investigate is whether the triplet profile's ability to distinguish scales well with system size. In FIG. 2.3 we present the single shot distinguishability between two states under both a triplet profile measurement, D_1, and an optimal quantum measurement, D_1^q, as it scales with system size N. First consider the case of ψ_0 and ϕ_N — D_1^q rises to 1 with N while D_1 for the triplet measurement hovers at just under half this, with possibly a slight decrease with N. In this sense, a value of $D_1 \sim 0.45$ is decent. As an illustration of two states that are almost worst-case scenario, we also present the distinguishability of the ground-state ψ_0 and the first excited global singlet, ψ_{s*}, which represents the smallest energy, symmetry preserving excitation that could occur. Clearly these two states are orthogonal and thus $D_1^q(\psi_0, \psi_{s*}) = 1$, but in character they are very similar. Nonetheless the triplet profile produces a non-zero distinguishability, as can be seen from FIG. 2.3, which also only decreases slowly with N — not surprising since these two states are becoming closer relatively within the Hilbert space. To give a sense of what these values of D_1 mean in practice, the inset of FIG. 2.3 shows the number of required measurements, r, in order to achieve a total probability of successfully distinguishing two states, given the naive strategy of guessing independently which

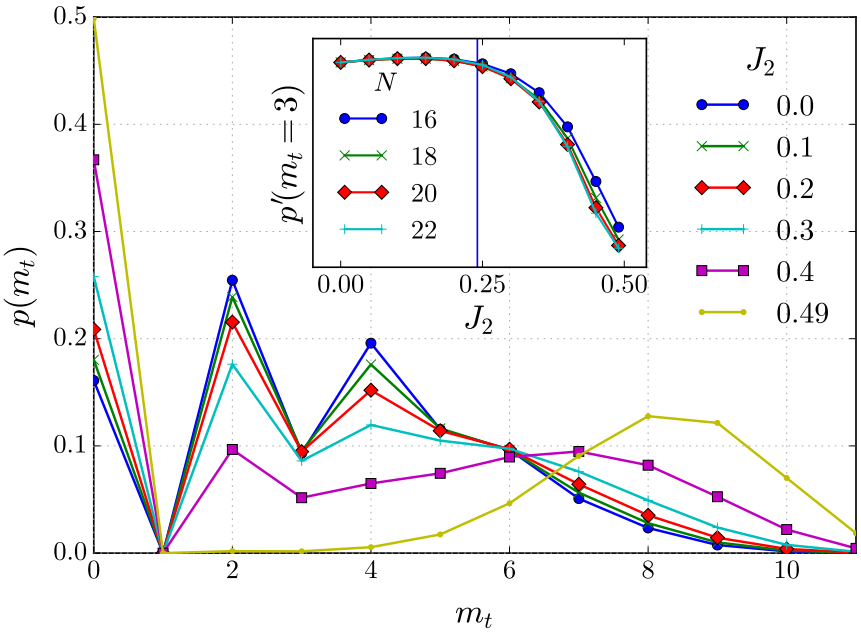


Figure 2.4: Observing the $J_1 - J_2$ quantum phase transition with the triplet profile: Triplet profile for the $N = 22$ Heisenberg ring ground-state, ψ_{J_2}, across the $J_1 - J_2$ phase transition. **Inset:** normalised probability of measuring three triplets varying with J_2 for various N. The vertical line denotes the exact critical point.

state was present each repeat . This sub-optimal scheme casts the overall distinguishability as that of between two binomial distributions. For example, if we take $D_1^q(\psi_0, \phi_N) \sim 0.43$, then 27 measurements would be required to guess which state was present with 99% success, as shown in the inset of Fig. 3.

2.5 Quantum Phase Transition in the $J_1 - J_2$ Model

In some condensed matter systems long range interactions are not negligible and play a crucial role in the character of the system. The simplest example is the $J_1 - J_2$ model with the Hamiltonian

$$H_2 = J_1 \sum_{i=1}^{N} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} + J_2 \sum_{i=1}^{N-1} \vec{\sigma}_i \cdot \vec{\sigma}_{i+2}, \tag{2.8}$$

where J_2 is the next nearest neighbour coupling strength. This model exhibits a quantum phase transition from a gap-less Heisenberg phase to a gapped dimer-

ized phase through increasing J_2. This infinite order quantum phase transition, in the Berezinskii–Kosterlitz–Thouless universality class, happens at $J_2/J_1 \simeq 0.24$. No standard quantities behave non-analytically across the transition, and instead properties such as the ground-state 'fidelity susceptibility' [149] or excited state fidelity [150] must be used to locate the critical point. Another interesting point in the dimerized phase is the Majumdar-Ghosh point at $J_2/J_1 = 0.5$. Here the ground-state is fully dimerized and can be explained as an equal superposition of $\bigotimes_{i=1}^{N/2} |\psi_-\rangle$ and its equivalent, but one site translated, form. The model in Eq. 2.8 could well be realised in future quantum dot arrays via a 'zig-zag' ladder geometry.

In FIG. 2.4 we show the triplet profile for the ground-state of H_2 for a number of J_2 values across the phase transition. As J_2 approaches the Majumdar-Ghosh point (i.e. $J_2/J_1 = 0.5$) the structure of an equal superposition of dimerizations becomes apparent — half of the state is exactly singlet pairs aligned with the measurements, and the other half appears as the identity since it is singlet pairs between the measurements. We note also that the rate and quality of change is different on either side of the critical point. This is shown more clearly in the inset of FIG. 2.4, where only the probability of getting three triplets (i.e. $m_t = 3$) is plotted versus J_2. In order to have a better perception of the effect of length N we have normalized the probabilities to the $J_2 = 0$ case, $p'(m_t = 3)$, for various lengths. Finally, this effect is not limited to $p(m_t = 3)$ — other values and combinations of m_t also give the same behaviour. However, due to the continuous nature of the transition, no such quantities are expected to show very sharp features, especially for short chain lengths.

2.6 Heralded entanglement of distant spins

We now show a potential quantum information application of using solely singlet triplet measurements in the form of *long-distance* entanglement. Generating perfect entanglement over arbitrary distances will likely be required for many quantum information tasks. In a many-body system, it has been shown [151, 152], that performing measurements on part of a system can localize entanglement between

the remaining, unmeasured parts. This is known as *localizable entanglement*, and we demonstrate here that singlet-triplet measurements on the ground-state of the Heisenberg chain can probabilistically localize entanglement between any two qubits. For applicability, to a quantum bus for example, we consider now an *open* chain, where the first and last quantum dots are desired to be entangled. The nature of global singlet states guarantee that if all but one pair of spins is measured and found in the singlet state, the final pair must also be in the singlet state. As previously described, this follows from the fact that all $m_t = 1$ states have spin-1, and no overlap with the SU(2) subspace. The generation of a perfectly entangled singlet is therefore reliant on the probability of finding this all-singlet outcome, $q(m_t = 0)$, but is certain to be there (i.e. *heralded*) if the measurement succeeds. Note that we use the symbol $q(m_t)$ for the probability of finding m_t triplet outcomes in our ST-measurements for the heralded entanglement scheme, which leaves one pair unmeasured, to discriminate it from $p(m_t)$ in the previous section in which all qubits are measured. Compared to a dynamic, gate based-scheme, the simultaneous nature of the measurement minimizes the time required and thus exposure to decoherence.

Since any ground-state with SU(2) symmetry displays this feature, we can also think about engineering the exchange coupling strengths along the chain to promote this configuration. One option is to weaken the coupling of just the end spins as

$$H_e = J_e(\vec{\sigma}_1 \cdot \vec{\sigma}_2 + \vec{\sigma}_{N-1} \cdot \vec{\sigma}_N) + J_1 \sum_{i=2}^{N-2} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} \tag{2.9}$$

where J_e is the ending coupling and is smaller than J_1. For $J_e \ll J_1$ it is known that a very high entanglement is established between the outermost spins in the ground-state of the system [153]. However, this entanglement is thermally unstable due to a vanishing energy gap. We combine this scheme, using larger values of J_e, and a heralding ST-measurement to achieve perfect entanglement with a higher rate. As another way to improve the probability $q(m_t = 0)$ we may also consider a

Hamiltonian with alternating couplings as

$$H_a = J_1 \sum_{i=1}^{N-1} [1 - (-1)^i \delta] \, \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} \tag{2.10}$$

where δ is the dimensionless anisotropy parameter.

In FIG. 2.5, we show the probability of an all singlet result, $q(m_t = 0)$, for these three cases, as the length of chain, N, varies for the case of $J_e = 0.5J_1$ and $\delta = 0.1$. Weakening the end-bonds yields a consistent improvement in long-distance entanglement over the normal Heisenberg chain, but both still decrease exponentially with length. The heralded nature of the entanglement means that for small enough chains repetition could still make the procedure viable. Moving to the ground-state of the H_a, we find that $q(m_t = 0)$ becomes almost constant with N at a value of ~ 0.3. A subtlety here is that engineering a Hamiltonian in this way can reduce the size of the energy gap, making the ground-state harder to prepare. This is in fact the case for both H_e and H_a above. Although this means that reaching the ground-state via direct cooling becomes more difficult, adiabatic state preparation has been shown to much alleviate the issue [121].

Finally, we point out that if full Bell-state measurements are possible on nearest neighbour spins, then perfect entanglement is *always* achieved between the ends. The state is not always the singlet Bell-state, but can be identified or corrected simply by counting the number of each Bell-states found and requiring the whole state to still be spin-0. This is essentially the same mechanism as addressed in [154].

2.7 Imperfections

The goal of our simulator is to create the ground-state of the Heisenberg Hamiltonian. In reality, thermal fluctuations spoil the quantum state of the system resulting in a thermal state

$$\rho^{\text{th}}(\beta) = \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})} \tag{2.11}$$

where $\beta = 1/k_B T$ and k_B denotes the Boltzmann constant. Performing the characterization ST-measurements on a thermal state result in a triplet profile $p(m_t)$

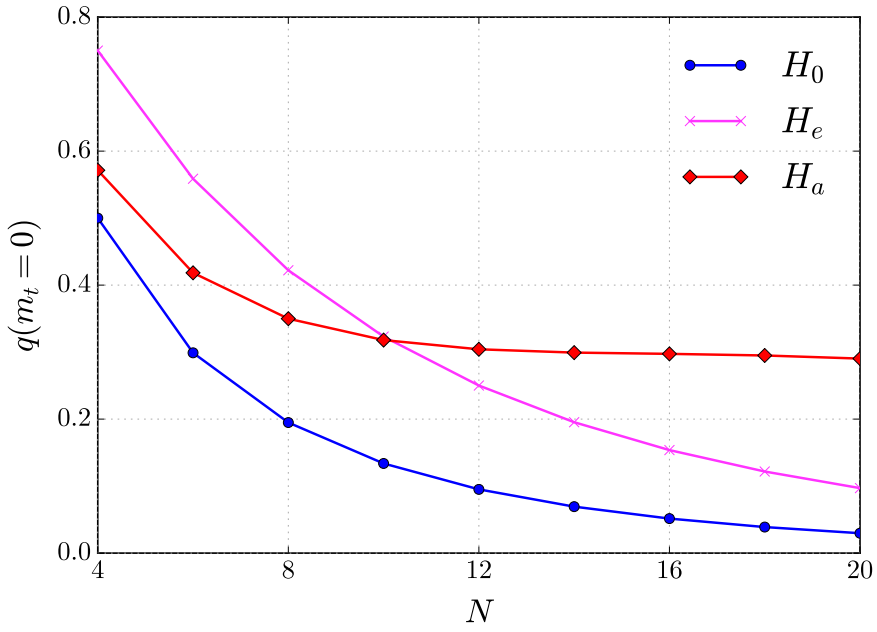


Figure 2.5: Localizing entanglement to the ends of a chain using singlet-triplet measurements on the middle $N-2$ qubits: Probability of finding all singlets after measuring the middle $N-2$ spins of the ground-state of an open Heisenberg chain with various coupling configurations. From this outcome, a perfect singlet in the remaining two spins at either end is heralded. The three Hamiltonian configurations are: H_0 — constant coupling, H_e — end couplings weaker by 50%, H_a — alternate couplings weaker by 20%.

which is shown in FIG. 2.6(a). From the figure, we find that up to $k_BT/J_1 = 0.2$ (approximately the gap of the Hamiltonian) the observed triplet profile is largely unchanged. Between $k_BT/J_1 = 0.5$ and $k_BT/J_1 = 1$ the oscillations suggesting many-body entanglement die out, and above the state appears largely classical. Since we know that for temperatures smaller than the energy gap the thermal state has very close to unit fidelity with the ground-state, what the result in Fig. 2.6(a) shows is that our singlet-triplet profile is sensitive to any rise in temperature that would significantly affect the state. Although we can positively identify a departure from the ground-state in this way, attributing the noise specifically to thermal fluctuations or identifying the temperature poses a greater challenge, though one worth investigating in the future.

Another dominant form of noise[155] arises from each electron's hyperfine

interaction with proximal nuclei. This manifests as an isotropic, normally distributed random static magnetic field for each site, which we can model with the Hamiltonian

$$H^{\text{nuc}}(B_n) = J_1 \sum_{i=1}^{N-1} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} + \sum_{i}^{N} \mathbf{B}_i \cdot \sigma_i \tag{2.12}$$

where $\mathbf{B}_i$'s are effective magnetic fields with random directions. The amplitude of these fields are determined by a Gaussian probability distribution as

$$P(\mathbf{B}) = \frac{1}{(2\pi B_n)^{3/2}} \exp\left(-\frac{\mathbf{B}\cdot\mathbf{B}}{2B_n^2}\right) \tag{2.13}$$

where B_n is the variance of the distribution and quantifies the strength of the nuclear field noise. The noise is quasi-static (changes slowly relative to the electron dynamics) and thus we can think of each experimental run as having a fixed set of random fields and simply average over many runs until convergence is reached.

In FIG. 2.6(b), we find that the nuclear noise quickly changes the triplet profile such that $B_n < 0.1$ would likely be required for a decent characterisation. Above $B_n \sim 0.3$ the oscillations disappear. Actual values for the bare value of B_n/J_1 estimate it below 0.1 [127], which hardly affects our triplet profile characterisation. Moreover, many successful avenues exist for reducing the effect of the nuclear noise, such as dynamical decoupling [156], and moving to Si/SiGe quantum dots [157], though these both introduce their own challenges for scaling to dot array simulators.

The effect of noise on long distance entanglement could be two-fold, it could change the probability of getting an all-singlet measurement, $q(m_t = 0)$, and it could also make the resultant state shared between the end qubits less entangled. In practice we find that $q(m_t = 0)$ is roughly constant across the region of interest for both temperature variation and hyperfine interaction. So, it suffices to consider only the *remaining* entanglement, E, which we characterise with the concurrence [158] on the reduced density matrix of the ends post-measurement.

In FIG. 2.7(a) we show how this remaining entanglement varies as a function of temperature. As the figure shows, there is a plateau of low temperature for which perfect entanglement still remains, though this drops with N and can be again be

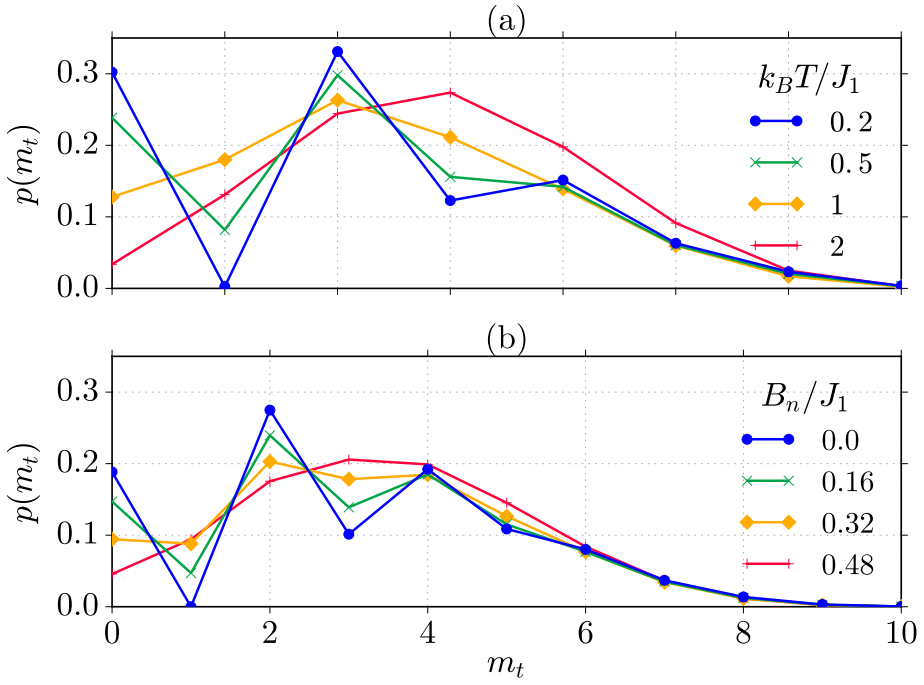


Figure 2.6: Characterising the Heisenberg chain under the influence of temperature and hyperfine interactions: **(a)** Triplet profile of the Heisenberg ring (H_0) ground-state for varying temperature T, here for chain length $N = 14$. **(b)** The same but for varying random nuclear field strength B_n, here for chain length $N = 20$. It is worth mentioning that a realistic (but pessimistic) estimation of the hyperfine interaction is $B_{\rm n}/J_1 \sim 0.1$ [127]. This will have little effect on state discrimination.

linked with the Hamiltonian's gap. Similarly to the case of state characterization, we find that nuclear noise has a much more immediate effect on the long-distance entanglement rather than on $q(m_t)$. In FIG. 2.7(b) we plot entanglement versus B_n. As the figure shows for $B_n < 0.1J_1$, which as mentioned is a conservative estimation based on experiment [127], the entanglement remains high even for chains as long as $N = 20$.

Another potential source of error in quantum dot array simulators are fluctuations in the charge potential landscape. The overall effect can be modelled to first order as a random fluctuation of J_1 about its mean value [159]. Since this type of noise maintains the SU(2) symmetry of the system, the essential arguments regarding oscillations in the triplet profile and localizing heralded entanglement remain intact. Indeed, there is evidence that the overall groundstate of a system with moderately

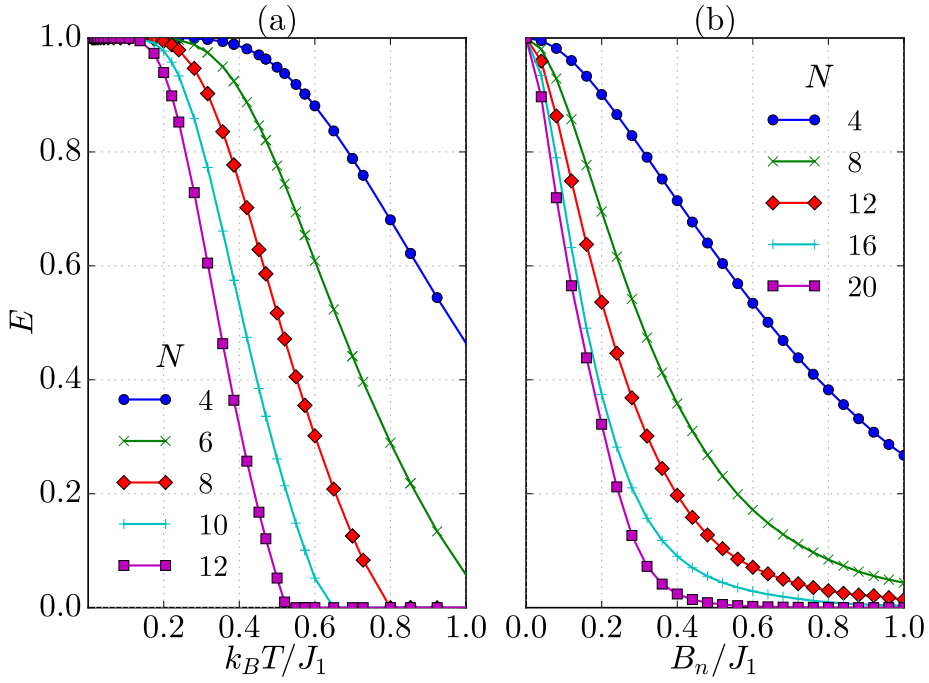


Figure 2.7: Entanglement localization to the ends of an open chain under the influence of temperature and hyperfine interactions: **(a)** Entanglement (as measured by concurrence) between the two furthest spins after an all-singlet measurement result on the remaining middle section of the chain, varying with temperature $1/\beta$ and length N. **(b)** The same, but now varying with the strength of the random nuclear field B_n. It is worth mentioning that a realistic (but pessimistic) estimation of the hyperfine interaction is $B_\mathrm{n}/J_1 \sim 0.1$ [127]. For all considered lengths, this yields high entanglement.

random couplings is very similar in terms of character and utility [121, 160, 161]. In Fig. 2.8 we show the effect of this noise on state characterization as well as the average fidelity between the ideal ground-state and many realizations of the erroneous ground-state, $\bar{f}(\psi_0, \psi_0^{\sigma_J})$. One can see that the average fidelity remains above 85% for $\sigma_J < 0.1J_1$, which is a high level of randomness. The corresponding change in the triplet profile also becomes noticeable with increasing σ_J, and as expected, $p(m_t = 1)$ remains zero throughout. Since even with this noise, the conditions for entanglement localization using singlet-triplet measurements are met, that scheme in its basic form is not affected. One observation is that the slight randomization of J_1 actually on average raises the chance of perfect entanglement, $q(m_t = 0)$, when compared to a Heisenberg chain (data not shown).

A final source of potential error worth discussing is the singlet-triplet readout

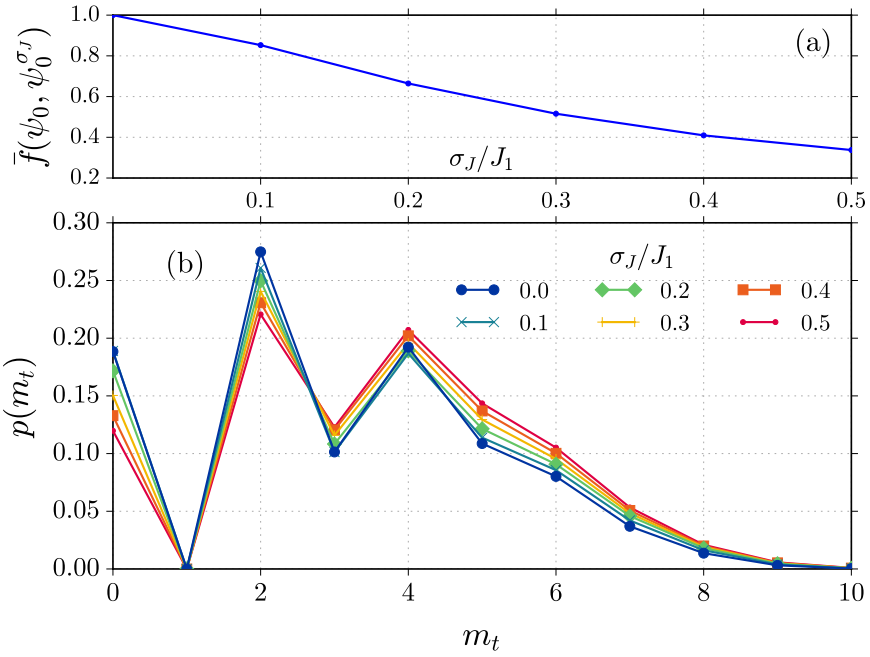


Figure 2.8: Effect of random couplings caused by charge fluctuations. **(a)** Average fidelity of the ideal groundstate, ψ_0, with many realizations of the erroneous ground-state generated due to random coupling noise, $\psi_0^{\sigma_J}$, as a function of the strength of those fluctuations, σ_J, for chain length, $N = 20$. **(b)** Averaged triplet profile for the ground-state of a Heisenberg chain with fluctuating couplings, varying with the strength of that fluctuation, σ_J.

fidelity, which for a RF-reflectometry based method reduces to the error in distinguishing two levels of capacitance. We have assumed this readout to be perfect throughout, for two main reasons. Firstly, this measurement is already very sensitive [162] in comparison to the other sources of error. Secondly, our measurement is single-shot, which means that its sensitivity can be increased simply by extending the integration time.

2.8 Experimental Realization

In this section, we discuss a potential experimental realisation. A SEM image of a gate-defined dot array, recently developed in [124], is shown in Fig. 2.9(a) in which fourteen quantum dots interact in a 2×7 array. Similar structures are being developed in other groups [163, 164]. The ladder structure, shown in Fig. 2.9(a), is capable of realizing a $N = 14$ open chain, but in principle a ring geometry is possible and

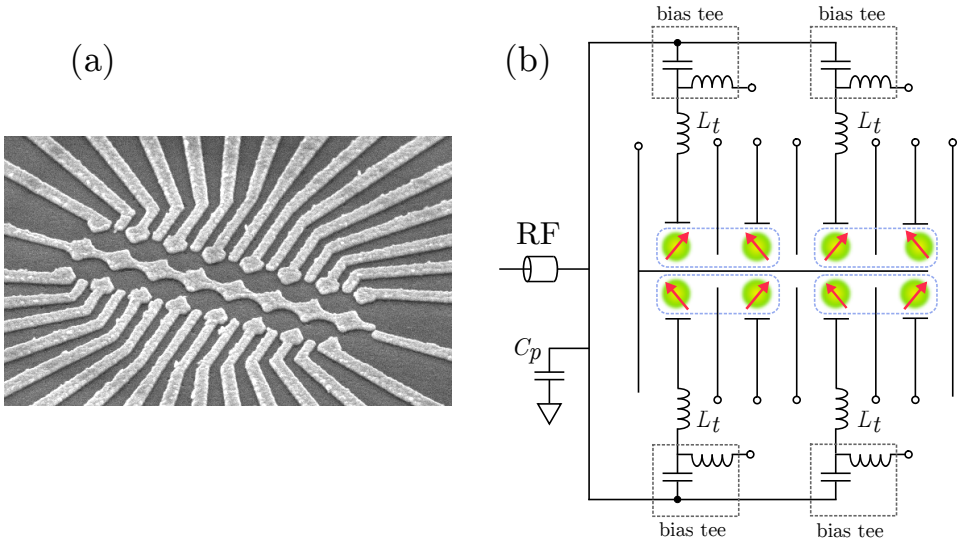


Figure 2.9: A realistic quantum dot array with triplet profile measurement: (a) SEM image of an example quantum dot array, as realised in [124]. **(b)** Circuit schematic of an eight dot device with dispersive gate sensors which forms a $N = 8$ open spin chain. The gate sensors are formed by coupling an RF-signal to gate electrodes via bias tees. L_t are chosen so that a resonant tank circuit is formed in combination with the dot system and the parasitic capacitance, C_p. The reflected RF-signal is used to read out the total capacitance of all double dots in parallel and hence m_t. The dashed blue lines denote pairings of the quantum dots for this ST-measurement.

both yield qualitatively similar results. Although the exchange coupling J_1 can be tuned to very large values, due to the limitations imposed by electronics speeds, a $\sim$ 1GHz value is preferable. In fact, in Ref. [127] J_1 up to 3μeV (0.75GHz) has been reported. In order to initialize the system in its ground-state solely through cooling, the energy gap, ΔE, has to be larger than the temperature of the fridge, typically around $T \sim 50$mK (i.e. $k_B T = 4.3\mu$eV for dilution fridges. This currently limits direct initialization to short chains ($N \sim 6$). However, for longer chains, a series of double dot singlets can be adiabatically welded to form the ground-state even in higher temperatures in a time-scale much less than the thermalization time [121].

We now describe how the triplet profile is measured once the target state is realised in the quantum dot array. We rapidly (with respect to $1/J_1$) raise voltage barriers to isolate pairs of quantum double dots, each of which can act (when connected to an appropriate circuit) as a capacitor, with capacitance dependent on whether the spin state is a singlet or a triplet. It is important that isolating the pairs is

rapid in order to avoid any adiabatic evolution towards the new effective Hamiltonian, which would change the state. When all these capacitors are connected in parallel in a LC-circuit a single reading of the total capacitance measures their sum, from which m_t can be deduced. Such a circuit is shown in Fig. 2.9(b), which measures the total capacitance of 4 quantum double dots ($N = 8$) in a single shot. The set-up uses dispersive gate sensors coupled to DC gate electrodes via bias tees [134]. The inductors, together with the parasitic capacitance C_p, form the resonant circuit with the dot array and thus one can sense the capacitance through the phase and amplitude of the reflected RF signal.

2.9 Relation to Many-Body Localization

The noise model presented in Eq. (2.12) is strikingly similar to that of the paradigmatic many-body localized (MBL) spin chain. Specifically, the slow nature of the average nuclear moment's movement leads to an effective, quenched disorder per site. Although the standard MBL model considers disorder in the z-direction only – and uniformly distributed – we might still expect a transition to MBL behaviour with the quantum dot array model. A now well known method to identify the transition point is to use the value of the mean energy level spacing ratios of the Hamiltonian. A single such ratio is defined as

$$r = \min\left(\frac{E_{k+1} - E_k}{E_k - E_{k-1}}, \frac{E_k - E_{k-1}}{E_{k+1} - E_k}\right) \tag{2.14}$$

for adjacent energy eigenvalues $E_{\{k-1,k,k+1\}}$. Averaging over of many eigenvectors near energy density ϵ and many instances of randomness yields a number, $\langle r_\epsilon \rangle$ that can be related to random matrix theory [165]. It can either display Gaussian statistics, implying repulsion and thus interaction between eigenvectors, or Poissonian statistics, implying no repulsion. in Fig .2.10 we show how this value changes across a wide range of random nuclear fields strengths, for varying chain lengths N.

There is clear evidence of a phase transition that sharpens with increasing lengths. Since these systems are small, they experience a pseudo-critical point of randomness, $\tilde{B}_n^c$, significantly smaller than the thermodynamic limit case. By

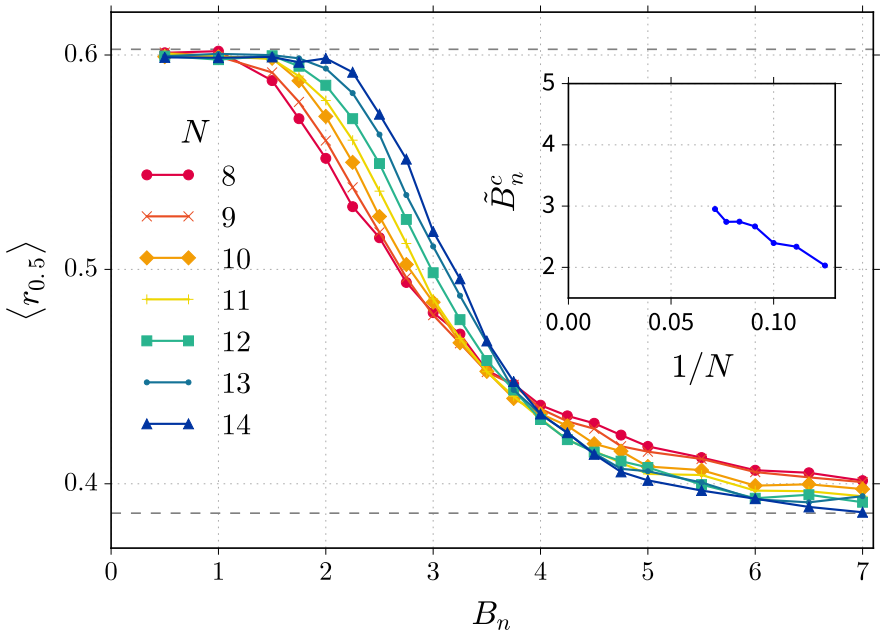


Figure 2.10: Mean energy level spacing ratio of the Hamiltonian from Eq. 2.12 in the middle of spectrum $\langle r_{0.5} \rangle$, as a function of random nuclear field strength, B_n for varying chain lengths N. The dashed lines denote the expected values for Gaussian Unitary Ensemble statistics (upper), which implies ergodic behaviour, and Poissonian statistics (lower) which implies an MBL phase. For each Hamiltonian instance, 50 eigenvalues are extracted from the middle of the spectrum $\epsilon = 0.5$ from which the statistics are calculated. This is further average over 1000-200 instances of noise, depending on system size. **Inset:** Scaling of the pseudo critical point, $\tilde{B}_n^c$, with inverse system size, a rough extrapolation of this line to infinite size yields $B_n^c \approx 4.5$.

extracting the location of the maximum derivative of these lines, one can then extrapolate roughly the value of the true critical point B_n^c which we find to be ≈ 4.5. Using other established methods such as the block entropy of bipartition gives similar values.

There are two important points to be drawn from this. The first is that it may well be interesting to further investigate different types of random noise in the context on MBL. Here, in the ergodic (but still random) phase with $0.5 < B_N < 2.0$, having noise present in all three directions removes any symmetry from the Hamiltonian, whereas if the noise is in the z-direction only there remains an abelian $\mathbb{Z}_2$ symmetry – rotation about the z-axis. A result of this is that Gaussian *unitary* ensemble (GUE)

statistics are encountered rather than Gaussian *orthogonal* ensemble. Having such a physical model that appears very similar to a GUE random matrix might well be a useful reference point in future.

The second important conclusion to take from this is that quantum dot arrays could quite possibly act as viable simulators of MBL. Here, though one cannot directly control the strength the of the random nuclear fields, B_n, one can engineer the strength of the coupling J_1. Since it is only the ratio of these parameters that matters, one now has an effective MBL model, where the main source of noise has now itself been harnessed as part of the simulation.

2.10 Conclusions

Motivated by established technology, we have explored the possibility of using solely singlet-triplet measurements to characterise the achieved ground-states of quantum dot arrays and found that a measurement of a triplet profile is largely sufficient for distinguishing the ground-state from other potential candidates. Features of this quantity can also indicate that the achieved state is highly non-classical. Our investigation fits with experimental accessibility as we only demand nearest neighbour measurements, do not demand a full Bell-basis measurement (although this can be achieved in principle with further single qubit rotations), and motivated by scalability do not even demand positional information of the outcomes: only m_t, as shown in Fig. 2.1. To demonstrate its utility, we investigated the $J_1 - J_2$ phase transition in the Heisenberg ladder.

Since our method is suitable for any models with isotropic antiferromagnetic couplings, one could consider in the future investigating 2D arrays and other more complex geometries. For non-Heisenberg Hamiltonians, such as the Ising model with transverse field, we expect that singlet-triplet measurements are still useful since different phases tend to have different local correlations. Another clear direction would be to consider the extra information currently missed by only recording the total number of triplets. For example, if information regarding the clustering of the triplet occurrences was retained, this could serve as a second axis on the probability

profile. Such an increase in the probability distribution space would clearly aid in distinguishing quantum states, and would also likely reflect physical traits of the system such as correlation length.

As well as characterization, we showed that singlet-triplet measurements have a quantum information processing application in localizing entanglement between the opposite ends of an open SU(2)-symmetric chain. Engineering the couplings slightly allows this effect to be amplified, though the effect on the Hamiltonian's resultant spectrum must be considered.

Finally, we considered the relevant noise sources for practical application of these techniques in GaAs quantum dot arrays for example. Interestingly, the nature of this noise lends quantum dots to potentially be a viable direct quantum simulator for the phenomenon of many-body localization, and so we have also noted several possible areas of research in that direction.

Chapter 3

Entanglement & the many-body localization transition

This chapter is adapted from the publication:

- Johnnie Gray, Sougato Bose, and Abolfazl Bayat. Many-body localization transition: Schmidt gap, entanglement length, and scaling. *Phys. Rev. B*, 97(20):201105, 2018.

3.1 Abstract

Many-body localization has become an important phenomenon for illuminating a potential rift between non-equilibrium quantum systems and statistical mechanics. However, the nature of the transition between ergodic and localized phases in models displaying many-body localization is not yet well understood. Assuming that this is a continuous transition, analytic results show that the length scale should diverge with a critical exponent $\nu \geq 2$ in one dimensional systems. Interestingly, this is in stark contrast with all exact numerical studies which find $\nu \sim 1$. We introduce the Schmidt gap, new in this context, which scales near the transition with a exponent $\nu > 2$ compatible with the analytical bound. We attribute this to an insensitivity to certain finite size fluctuations, which remain significant in other quantities at the sizes accessible to exact numerical methods. Additionally, we find that a physical manifestation of the diverging length scale is apparent in the entanglement length computed using the logarithmic negativity between disjoint blocks.

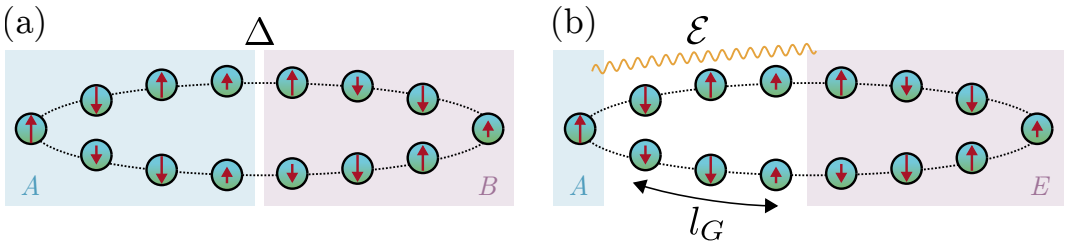


Figure 3.1: Schematic of the two main quantities studied here: a) the Schmidt gap, Δ, across a bipartition of the system and b) the logarithmic negativity, $\mathcal{E}$, between disjoint blocks separated by length l_G.

3.2 Introduction

It has become apparent that the Anderson Localization [73] of disordered models can survive in the presence of interactions [74], with rigorous proof now found for 1D systems [76, 77]. This phenomena, known as many-body localization (MBL), has attracted much interest [166, 167, 81, 168, 169, 75] in fundamental physics due to the fact that such systems generically break ergodicity and fail to thermalise — thus lying beyond the scope of statistical mechanics. Additionally, MBL occurs throughout the energy spectrum, implying that its fingerprint can be observed at all temperatures. These facts combined have significant practical implications for quantum transport [74] and information storage [170, 171, 172]. Experimental advances have allowed the controlled observation of MBL phenomena [106, 173], further driving interest.

Considerable progress has been made in understanding the strongly localized phase, particularly in terms of local integrables of motion [80, 81, 82, 83, 84, 85] , which permit a MPS description of all eigenstates [174, 175, 176, 177, 178, 179]. However, eigenstates in the ergodic phase generally have volume law entanglement, restricting one to exact diagonalization techniques and small system sizes (up to ~ 20 spins) — this has constrained the development of a clear picture of the nature of the transition from ergodic to MBL (the MBLT). For example, questions that still require attention include: (i) Which quantities can best characterize the transition? (ii) Is it valid to treat the MBLT using the same framework, based on the emergence of a diverging length-scale, developed for zero-temperature quantum phase transitions? (iii) If so, what is the universal critical exponent, ν, governing this length-scale? And

(iv) what is the physical picture of the said length-scale?

An extensive exact numerical analysis of the MBLT, using a variety of quantities, can be found in [79], in which finite size scaling analysis throughout the spectrum allows the observation of a mobility edge. In fact, it is now commonplace to diagnose the MBLT with the mean energy level statistics and the block entanglement entropy [75, 167, 79, 180, 181, 182, 183]. These works are largely based on the assumption that the MBLT is continuous, and their exact numerical analyses have consistently found $\nu \sim 1$. This is in striking contrast with analytic results, found by Chayes-Chayes-Fisher-Spencer [184] and Chandran-Laumann-Oganesyan [185], which would demand $\nu \geq 2/d$ for system dimension d (the CCFS/CLO bound). A recent explanation [182] posits that at the finite system sizes available for exact studies, the fluctuations in these quantities are not yet dominated by the true disorder. Thus it is highly desirable to use a new quantity better able to capture the real disorder induced transition properties.

In this chapter, we bring in new tools to understand the nature of the MBLT. Firstly, the Schmidt gap, which has been successfully employed as an order parameter in quantum phase transitions [186, 187, 36]. Secondly, an entanglement length computed from the logarithmic negativity [188, 189, 25, 26, 190], quantifying the bipartite entanglement between two disjoint blocks [191, 192, 193, 194, 195], which has been previously used to probe the extension of the Kondo screening cloud [196, 197]. We find that, unlike previously used quantities, the Schmidt gap reveals a critical exponent $\nu \geq 2$, consistent with the CCFS/CLO bound, though, curiously as opposed to previous studies, it does *not* act as an order parameter. Moreover, we find that the entanglement length witnesses the emergence of a diverging length scale at the transition from ergodic to MBL phase.

3.3 Model

We consider a periodic spin-$1/2$ Heisenberg chain, with random magnetic fields in the z-direction:

$$H = \sum_{i=1}^{L} \left(J \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} - h_i S_i^z \right), \tag{3.1}$$

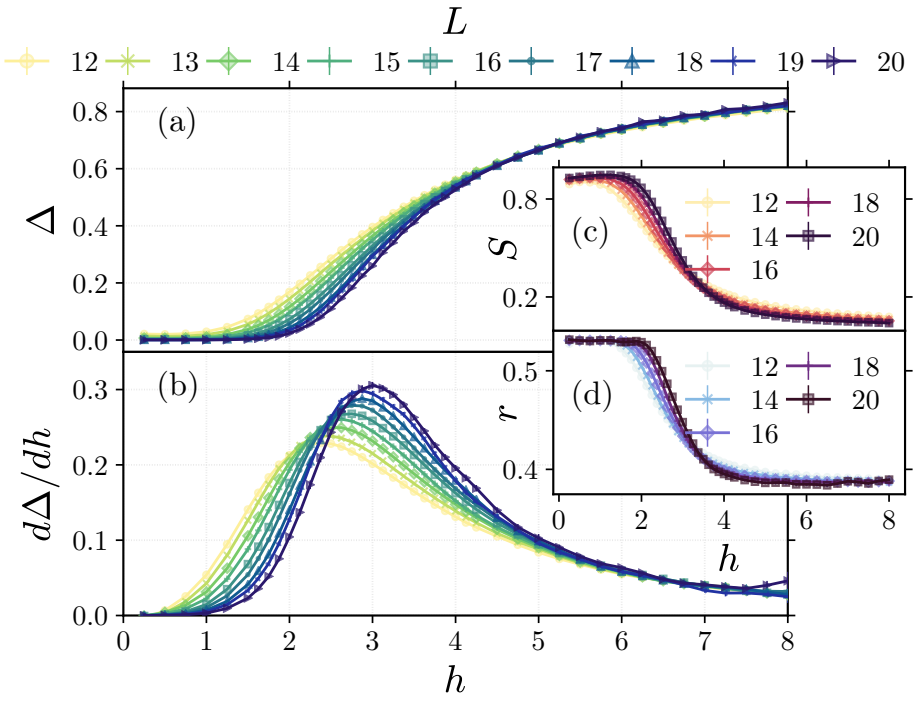


Figure 3.2: (a) and (b) : The Schmidt gap, Δ, and its derivative as a function of disorder, h, across the MBLT for varying chain length, L. (c) and (d) : The normalized half chain entropy, S, and the mean energy level spacing ratio, r, as a function of disorder for varying L. Error bars shown where visible.

with J the exchange coupling, $\boldsymbol{S}_i = \frac{1}{2}(\sigma_i^x, \sigma_i^y, \sigma_i^z)$ a vector of Pauli matrices acting on spin i and dimensionless parameter h_i the random magnetic field at site i drawn from the flat distribution $[-h, h]$. We diagonalize the Hamiltonian in either the spin-0 or spin-$\frac{1}{2}$ subspaces for even and odd L respectively. For each random instance we extract 50 eigenvectors, $\{|E_k\rangle\}$, in the middle of the energy spectrum [198, 199]. Since there is evidence of a mobility edge in MBL [79], at least for finite sizes, this targeting sharpens any transition observed. The choice of 50 is a reasonable compromise on numerical efficiency whilst being statistically representative.

3.4 Characterizing the MBLT

The main quantity we compute, new in the context of MBL, is the Schmidt gap. For two chain halves (or as close to for odd L), A and B, as shown in Fig. 3.1(a), an eigenvector's reduced density matrix is $\rho_{A,k} = \mathrm{Tr}_B(|E_k\rangle \langle E_k|)$ for a particular sample of the random fields. The disorder-averaged Schmidt gap is then defined

as $\Delta = \overline{\langle \lambda_1^k - \lambda_2^k \rangle_k}$, where λ_1^k, λ_2^k refer to the largest eigenvalues of the reduced density matrix $\rho_{A,k}$, $\langle \cdot \rangle_k$ denotes average over eigenstates and $\overline{\cdot}$ denotes the average over many samples. The Schmidt gap has previously been shown to act as an order parameter for quantum phase transitions [186, 36]. We explore the possibility of using it for characterizing the MBLT. Unlike entanglement entropy, the Schmidt gap ignores most of the spectrum of $\rho_{A,k}$, describing only the relationship between the two dominant states across the $A - B$ cut. This is pertinent in light of the recent finding that while the Schmidt values decay polynomially in the MBL phase [175], finite size corrections are stronger for small Schmidt values. In the ergodic phase we expect strong entanglement to produce multiple, equally likely orthogonal states, thus $\Delta \sim 0$. In the MBL phase, however, a single dominant state should appear on either side of the cut, with Δ rising towards 1 as $h \to \infty$, implying a tensor product. This behaviour is shown in Fig. 3.2(a) and becomes becomes sharper with increasing L. To see this more vividly, we plot the derivative of Δ with respect to h in Fig. 3.2(b). The derivative has a peak at $h = \tilde{h}_c$, which not only becomes more pronounced but also shifts to the right with L. We infer this to be the finite size precursor to the transition point, which suggests that in the thermodynamic limit, $L \to \infty$, the derivative of the Schmidt gap diverges at the MBLT and $\tilde{h}_c$ asymptotically approaches the transition point h_c.

For reference, we consider the normalised half chain entropy, widely employed to herald the MBLT [75, 167, 79, 180, 181, 182]. The von Neumann entropy of subsystem A is defined as $S_{\text{vn}} = -\text{Tr}(\rho_{A,k} \log \rho_{A,k})$. This is normalized by the Page entropy [200], $S_P = (1/\log 2) \sum_{i=n+1}^{mn} \frac{1}{i} - \frac{m-1}{2n}$, with m, n the Hilbert space dimensions of subsystems A and B, yielding the disorder-averaged $S = \overline{\langle S_{\text{vn}} \rangle_k}/S_P$. S_P is the expected entropy for a subsystem of a random pure state; since these overwhelmingly have entropy that scales as their enclosed volume, S gives a measure of how far $|E_k\rangle$ has departed towards area law behaviour. In Fig. 3.2(c) the behaviour of S across the MBLT is shown. In the ergodic phase its value approaches 1 (showing the volume law), whereas in the MBL phase it falls to 0 (representing the area law), as expected.

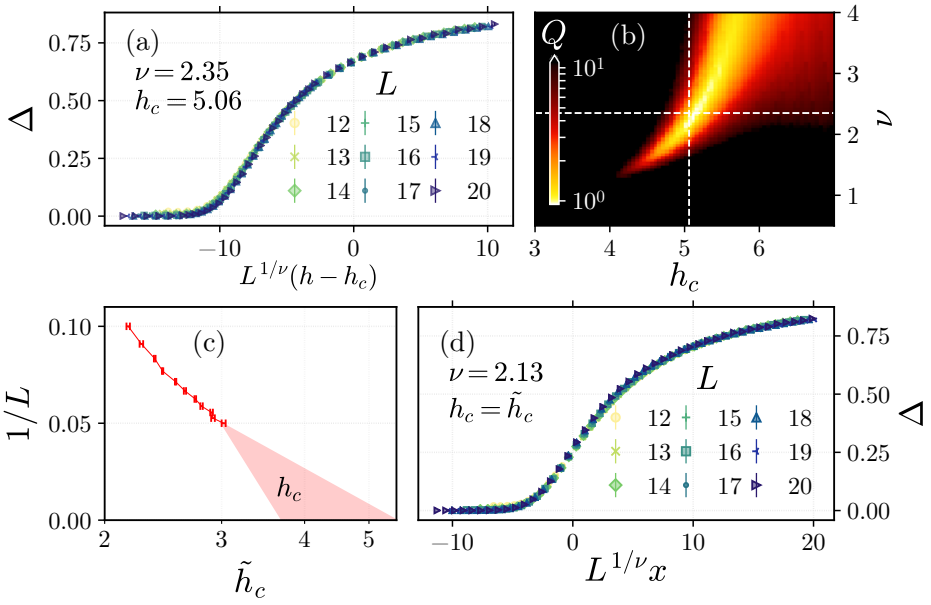


Figure 3.3: (a) Schmidt gap data collapse with fitting parameters as shown. (b) Quality of data collapse, Q, (lower is better) across the whole parameter space. The dashed lines denote the minimum point which yields the parameters shown in (a). (c) Pseudo-critical points $\tilde{h}_c$ as a function of inverse length $1/L$. (d) Schmidt gap data collapse using $\tilde{h}_c$ directly, and optimizing for ν only - the value of which is shown. Error bars shown where visible.

For reference we also compute the mean energy level spacing ratio, r. For energy eigenvalues E_k, with gaps $\delta_k = E_n - E_{n-1}$, this is defined as $r = \overline{\langle \min(\delta_k, \delta_{n+1}) / \max(\delta_k, \delta_{n+1}) \rangle_k}$. In the ergodic phase, energy level repulsion yields statistics for r that match those of Gaussian Orthonormal Ensemble (GOE) random matrices [165] with $r = 0.5307(1)$. In the MBL phase however, the eigenenergies are no longer correlated, and the energy level are simply spaced according to Poisson statistics, giving $r \approx 0.38629$. In Fig. 3.2(d) the behaviour of r is shown across the MBLT, clearly varying between these two statistical regimes. For all of these quantities, we average over between 10000 for $L = 10$ and 1000 for $L = 20$ samples of random fields, and compute errors using statistical bootstrapping across these samples.

3.5 Scaling

The behaviour in Figs. 3.2(a)-(b) suggests that the MBLT is a continuous transition in which a diverging length scale $\xi \propto |h - h_c|^{-\nu}$ emerges near the transition

point, consistent with [75]. In order to estimate the exponent ν, finite size scaling analysis [201] has previously been employed for various quantities, including the entanglement entropy S. These analyses, based on exact numerical methods, find $\nu \sim 1$ [202, 79, 182], contradicting the CCFS/CLO bound. A recently proposed explanation [182], suggests that there are two universality classes at play here, with that of inter-sample randomness not yet dominant for the system sizes studied.

In order to estimate ν for both models we consider the following finite size scaling ansatz,

$$\Delta = f(L^{1/\nu}x), \tag{3.2}$$

where $f(.)$ is an unspecified function and x is ideally the scaled coordinate $h - h_c$. Given the ansatz of Eq. (3.2), one can then find the best fit of h_c and ν, using an objective function quantifying quality. We use such a quality measure, Q, as refined in [203], which is discussed in Appendix A. In Fig. 3.3(a) we show optimal data collapse of Δ for various L, which is found to occur for $h_c = 5.06 \pm 0.09$ and $\nu = 2.35 \pm 0.21$. Remarkably, this value for ν is consistent with the CCFS/CLO bound, in contrast to finite size scaling analyses for S and r, which previous studies [202, 79, 182] have generally shown to yield values of $\nu \sim 1$ – a finding also reproduced in our analyses (data not shown). We show the quality of collapse, Q, for all possible combinations of h_c and ν in Fig. 3.3(b), the minimum point of which defines the best fit values of ν and h_c. To define errors on ν and h_c, we perform the scaling with various subsets of data (see appendix A) and compute the variance among all those which achieve a good quality.

The critical h we find with Δ is slightly higher than that generally reported. One possible explanation is that a lower effective ν fits best with a lower effective h_c, a relation that can be seen in Fig. 3.3(b). Thus it is possible that in other studies using S and r, where $\nu \sim 1$, h_c is artificially lower due to the finite size effects. We note that a standard method of extracting h_c independently – plotting the pseudo-critical points against inverse length, shown in Fig. 3.3(c) – does not give a decisive value for the real critical point. In fact $h_c \sim 3.7$ would seem to be a lower bound on the

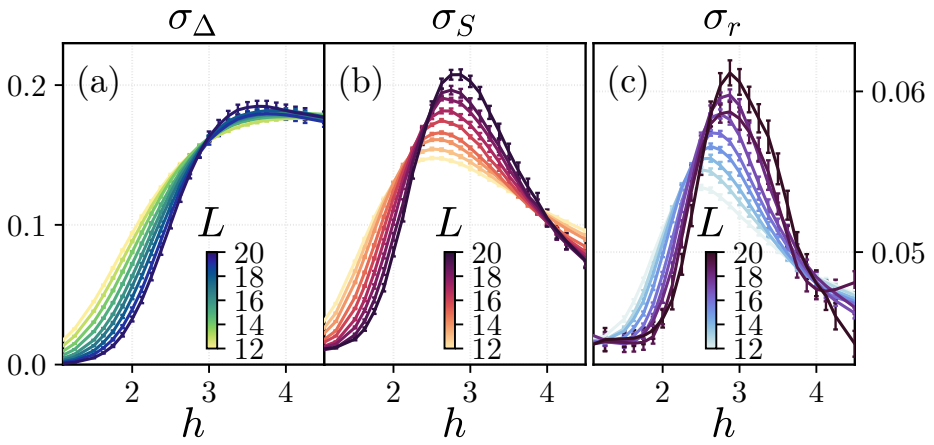


Figure 3.4: Standard deviation between samples for: (a) the Schmidt gap, σ_Δ; (b) normalized block entropy; σ_S; and (c) mean energy level spacing ratio, σ_r. Shown as a function disorder, h, and length, L. Error bars shown where visible.

transition point , with a value between 4.5 and 5.5 more consistent. Additionally, if one were to identify an intersection point for all lengths in Fig. 3.2 – which should occur at $h = h_c$ as implied by Eq. (3.2) – this would also be at $h \sim 5$. In contrast, the point of intersection for S and r shifts significantly as L increases - implying a deviation from the finite size ansatz. As a final cross-validation, to estimate ν independently from h_c, we take the pseudo-critical points $\tilde{h}_c$ directly to define the scaled coordinate x, and find the best quality of fit, Q, solely as a function of ν. This approach yields $\nu = 2.13 \pm 0.15$ – in accordance with the first estimate – for which data collapse is shown in Fig. 3.3. In contrast to previous groundstate quantum phase transitions[186, 36], here we find that the Schmidt gap is a scaling function rather than an order parameter. Namely it corresponds to $\beta = 0$ if Eq. (3.2) had pre-factor $L^{\beta/\nu}$.

3.6 Sample Fluctuations

In order to understand why the Schmidt gap is more successful than typical quantities, we study the fluctuation of Δ, S and r between samples. Motivated by Ref. [182], we consider how the size of these fluctuations scales with L. We define the standard deviations as $\sigma_\Delta^2 = \mathrm{Var}\left[\langle \lambda_1^k - \lambda_2^k \rangle_k\right]$, $\sigma_S^2 = \mathrm{Var}\left[\langle S_{\mathrm{vn}} \rangle_k / S_P\right]$ and $\sigma_r^2 = \mathrm{Var}\left[\langle \min(\delta_k, \delta_{n+1}) / \max(\delta_k, \delta_{n+1}) \rangle_k\right]$, with the variance $\mathrm{Var}[\cdot]$ taken across

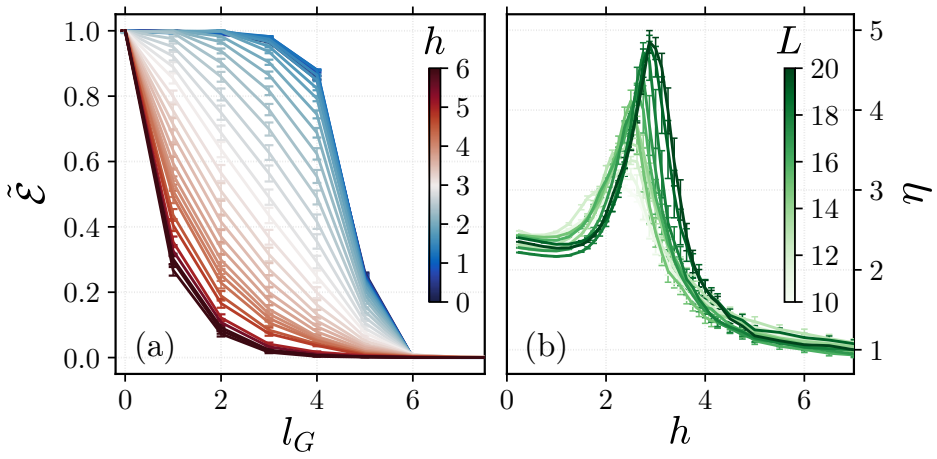


Figure 3.5: (a) Average logarithmic negativity as a function of the gap between two disjoint blocks as depicted in Fig. 3.1, the first block being a single spin, the second the rest of the system. Here, $L = 20$, which has a pseudo-critical point at $h \sim 3$. (b) Bipartite entanglement length, as computed with Eq. (3.3), across the MBLT for varying chain length L.

samples. These are shown across the MBLT for various system sizes in Figs. 3.4(a-c). All three quantities must lie between 0 and 1, thus their standard deviation is capped at 0.5. As the figures show however, the peaks of σ_S and σ_r are both still rising significantly with L and not yet saturated, whereas the peak of σ_Δ is almost constant. The implication is that for S and r, the effect of the small system sizes is to suppress the amount of fluctuations driven by the true disorder. On the other hand, changing the length L seems to have little effect on σ_Δ – suggesting that it already experiences the full, disorder driven, thermodynamic-limit fluctuations. A possible explanation is that finite size effects are dominantly confined to the smaller Schmidt coefficients, which still contribute significantly to σ_S. This could also be phrased in terms of the presence of various length scales, not yet very small compared to the correlation length, that the Schmidt gap is largely insensitive to.

3.7 Entanglement length

The nature of the diverging length scale ξ, in the context of MBLT, is mysterious and a physical picture is lacking. To shed light on this, we introduce an entanglement length, as previously used for detecting the Kondo screening cloud [196, 197]. Specifically, we consider the entanglement between a small subsystem A, here a

single spin, and an environment E, separated by a gap of length l_G, a geometry shown in Fig. 3.1(b). The reduced state of the two blocks is $\rho_{AE,k} = \mathrm{Tr}_G(|E_k\rangle \langle E_k|)$, where Tr_G is removes the $2l_G$ spins not in A or E. We use the logarithmic negativity [188, 189, 25, 26, 190] to quantify the entanglement between systems A and E, defining $\mathcal{E}(l_G) = \overline{\langle \log ||\rho^{\Gamma}_{AE,k}||_1 \rangle_k}$, with Γ the partial transpose, and $||\cdot||_1$ the trace norm. Since we are only concerned with the relative decay of entanglement we also define the normalized entanglement as $\tilde{\mathcal{E}}(l_G) = \mathcal{E}(l_G)/\mathcal{E}(0)$. This naturally gives information about bipartite entanglement over a range of scales, unlike the two-site concurrence for example (which quickly goes to zero for large separation), and unlike the widely used entanglement entropy (which cannot quantify the entanglement of mixed states – which inevitably arise when looking at two subsystems of a larger state). In the ergodic phase, due to volume law entanglement, the eigenstates are highly multipartite entangled between their spins. This implies that any reduced state of two small blocks is close to the identity and thus very weakly entangled. From this two features can be inferred: i) $\tilde{\mathcal{E}}(l_G)$ is initially expected to decay slowly with increasing l_G, and ii) $\tilde{\mathcal{E}}(l_G)$ must go to zero as $l_G \to L/2$. Since this precludes a linear type decay, it is expected that there is a distance at which $\tilde{\mathcal{E}}(l_G)$ rapidly decays - indeed we find this to be the case, with a sharp drop-off when half the system is traced out, i.e. $l_G \sim L/4$. In the MBL phase, however, A will be weakly entangled with only spins close to it, and thus $\tilde{\mathcal{E}}(l_G)$ should decay quickly even for small l_G. In Fig. 3.5(a) we plot $\tilde{\mathcal{E}}$ as a function of l_G for various disorder strengths h in a chain of length $L = 20$. As is clear from the figure the location of the main drop in $\tilde{\mathcal{E}}$ varies significantly with h. While in the ergodic phase $\tilde{\mathcal{E}}$ this decay is concentrated at $l_G \sim L/4$, in the MBL phase it is concentrated at $l_G \sim 1$. Interestingly, at the pseudo-critical point, ($\tilde{h}_c \sim 3$ for $L = 20$, see Fig. 3.2(b)), entanglement decays close to linearly — each spin lost contributes equally to the entanglement, implying that the bipartite entanglement is equally spread over many sites. This fits with a picture of a self-similar structure of entangled clusters [204, 181]. The detailed behaviour of $\tilde{\mathcal{E}}$ as a function of system size can be found in the appendix A.

To extract a length scale from $\tilde{\mathcal{E}}(l_G)$ we define a length, η, from the maximum

inverse gradient as such:

$$\eta = \left(\max_{l_G} |d\tilde{\mathcal{E}}/dl_G| \right)^{-1} . \tag{3.3}$$

To see why such a function works, first consider exponential decay, as appears in Fig. 3.5 on the MBL side of the transition. With an ansatz of

$$\tilde{\mathcal{E}}_{\text{MBL}}(l_G) = e^{-l_G/\lambda} \tag{3.4}$$

for some characteristic decay length λ, we find that

$$\left| \frac{\tilde{\mathcal{E}}_{\text{MBL}}(l_G)}{dl_G} \right| = \left| \frac{-1}{\lambda} \exp\left(\frac{-l_G}{\lambda} \right) \right| \tag{3.5}$$

which is maximised at $l_G = 0$, for which $\left| \frac{\tilde{\mathcal{E}}_{\text{MBL}}}{dl_G} \right| = \frac{1}{\lambda}$ and thus $\eta = \lambda$ as desired. If instead we take an ansatz of linear decay over some length λ, as observed in Fig. 3.5 at the transition point, we use

$$\tilde{\mathcal{E}}_{\text{crit}}(l_G) = 1 - l_G/\lambda \tag{3.6}$$

for which clearly $\left| \frac{\tilde{\mathcal{E}}_{\text{crit}}(l_G)}{dl_G} \right| = \frac{1}{\lambda}$ everywhere and thus again we find $\eta = \lambda$. Finally, in the ergodic phase we observe a plateau followed by decay over some length scale, λ, centered around length l_c. One ansatz that qualitatively captures such behaviour is the function

$$\tilde{\mathcal{E}}_{\text{erg}}(l_G) = \frac{1}{1 + \exp\left(\frac{4(l_G - l_c)}{\lambda} \right)} . \tag{3.7}$$

The absolute derivative of this is

$$\left| \frac{\tilde{\mathcal{E}}_{\text{erg}}(l_G)}{dl_G} \right| = \frac{4e^{4(l_G - l_c)/\lambda}}{\lambda(1 + e^{4(l_G - l_c)/\lambda})^2} \tag{3.8}$$

which is maximised for $l_G = l_c$ at the value $4/\lambda 2^2$ and so again we find $\eta = \lambda$. As such, Eq. (3.3) represents a robust way of finding the length over which the disjoint

entanglement decays for all noise strengths, as opposed to tailoring an ansatz to each.

The behaviour of η as a function of h for varying L is shown in Fig. 3.5(b), in which it can be seen to sharply peak at $h \sim \tilde{h}_c$ for each L across the critical region, – evidence that the diverging length scale ξ is closely captured by the length η. In the appendix A we show that taking the initial block as 2 spins yields almost identical results. A plausible explanation for the *increase* in η as one approaches the MBLT from the ergodic side is that proximal spins become off-resonant so that bonding (*bipartite* entanglement) takes place at increasingly longer scales – a process that is not possible if the spins are part of a large multi-partite entangled block. We note several interesting approaches that made use of the two site concurrence [205, 206] or mutual information [207], which despite revealing other interesting features, such as scaling, do not show a divergence in the localization length from both sides of the transition. An alternative approach to identifying the diverging length scale on the ergodic side based on the entanglement spectrum has been recently developed in Ref. [208]. It is an interesting open question whether that length is related to the entanglement length proposed here.

3.8 Conclusions

In this chapter we have explored the MBLT using the Schmidt gap and the entanglement length. We show that the Schmidt gap not only exhibits scaling at the MBLT, but does so with a critical exponent $\nu > 2$, compatible with analytic predictions. This compatibility is absent in all quantities studied with exact numerical methods thus far, a fact that we attribute to the presence of significant finite size effects which the Schmidt gap is less sensitive to. We have also considered an entanglement length computed using the logarithmic negativity across two disjoint blocks, which yields a diverging length scale at the MBLT.

Chapter 4

Machine-learning-assisted many-body entanglement measurement

This chapter is adapted from the publication:

- Johnnie Gray, Leonardo Banchi, Abolfazl Bayat, and Sougato Bose. Machine-learning-assisted many-body entanglement measurement. *Phys. Rev. Lett.*, 121(15):150503, 2018.

4.1 Abstract

Entanglement not only plays a crucial role in quantum technologies, but is key to our understanding of quantum correlations in many-body systems. However, in an experiment, the only way of measuring entanglement in a generic mixed state is through reconstructive quantum tomography, requiring an exponential number of measurements in the system size. Here, we propose a machine learning assisted scheme to measure the entanglement between arbitrary subsystems of size N_A and N_B, with $\mathcal{O}(N_A + N_B)$ measurements, and without any prior knowledge of the state. The method exploits a neural network to learn the unknown, non-linear function relating certain measurable moments and the logarithmic negativity. Our procedure will allow entanglement measurements in a wide variety of systems, including strongly interacting many body systems in both equilibrium and non-equilibrium

regimes.

4.2 Introduction

Entanglement is a key property for many emerging quantum technologies [209, 9, 210, 211, 212, 213], but also is essential for understanding the structure of strongly correlated many-body systems [214, 93]. Despite its paramount importance, only for the very limited case of a bipartition of a pure state can the entanglement, quantified by subsystem entropy, be measured in an efficient and state-independent way [215]. There are multiple proposals to carry out such a scheme in various physical systems, such as optical lattices [216, 217], quantum dot arrays [218] and Gaussian systems [219]. Recently, some of these have also been experimentally realised in simulated spin chains, for example in cold atoms [220] and photonic chips [221]. Nonetheless, pure states are very rare: they are not only difficult to prepare in realistic situations, but also difficult to maintain in the presence of an environment. For example, just consider the entanglement between: (i) optical modes traversing fibres, crucial for quantum communication; (ii) spatially separated parts of an extended many-body pure state, important for characterizing long range entanglement [222, 196, 223, 2, 224, 225]; (iii) two systems in a thermal state – in none of the above cases, ironically, can the entanglement entropy quantify the entanglement. Witnesses do exist for specific forms of entanglement, but these are state-dependent and provide only a simple yes/no answer [189, 226, 227], or bounds on the quantity of entanglement [228, 229, 230, 231, 232, 233, 234]. However, the crucial task of being able to accurately measure entanglement for mixed states in an experimental setting remains open.

While for pure states bipartite entanglement is uniquely defined by the entropy of the subsystems, for mixed states the landscape is far more complex [235, 23]. Aside from isolated special cases such as two qubit states [236] and bosonic Gaussian states [219, 237], only the (logarithmic) negativity [24, 25, 26, 27] is a computationally tractable quantity [28]. It bounds the distillable entanglement and teleportation capacity [26], and is a pivotally important quantity to estimate for both quantum

technologies [23, 235, 238] and condensed matter systems [38]. Nonetheless, there is no state-independent observable that can measure the logarithmic negativity, and thus its experimental measurement requires full state tomography [137] — demanding, in general, an exponential number of measurements in the system size. Recently, polynomial tomography schemes have emerged, such as tensor networks for lowly entangled states [139, 238], or breakthroughs in neural network state reconstruction [239, 240]. However, these may be insufficient for estimating entanglement, since many entanglement measures, such as the logarithmic negativity, are not continuous [241]. Namely, even if reconstructed state ρ_r approximates actual state ρ closely, the two may have significantly different negativities [242].

Here, we put forward a machine learning assisted scheme for accurately estimating the logarithmic negativity in a completely general and realistic setting, using an efficient number of measurements – scaling polynomially with system size. Our estimator works for a wide range of states, and is remarkably accurate for highly entangled states. Our method is based on measuring a finite number of moments of a partially transposed density matrix [243, 244, 245] from which we extract the entanglement negativity using machine learning. This *direct* estimation of negativity avoids approximate state reconstruction [139, 238, 239, 240], and represents a new front in applying classical machine learning to quantum problems [246, 247, 93, 248, 249]. Moreover, we propose a new method for measuring those moments, beyond [243, 244, 245], which is experimentally feasible in the many-body setting, since the individual building blocks have already been demonstrated in solid state [127] and cold atoms [220].

4.3 Logarithmic Negativity

Logarithmic negativity [24, 25, 26, 27] for a generic mixed state ρ_{AB} quantifies the entanglement between subsystems A and B. It is defined as:

$$\mathcal{E} = \log_2 \left|\rho_{AB}^{T_A}\right| = \log_2 \left|\rho_{AB}^{T_B}\right| = \log_2 \sum_k |\lambda_k| \tag{4.1}$$

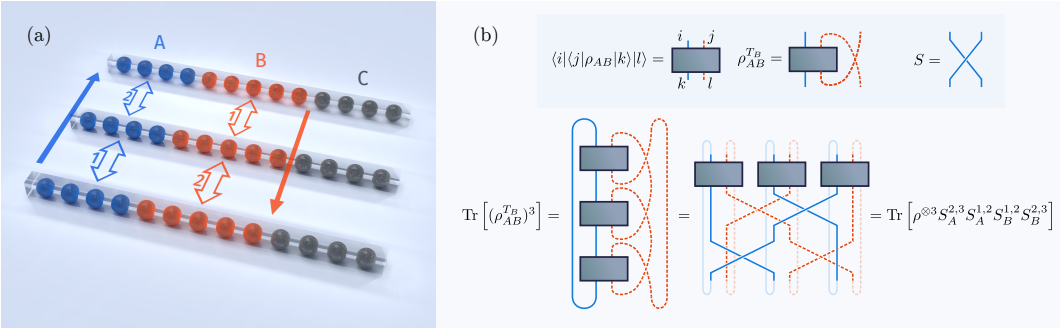


Figure 4.1: Schematics: (a) Example measurement set-up for the moments, $\mu_m = \mathrm{Tr}\left[(\rho_{AB}^{T_B})^m\right]$, here for $m = 3$, from which one can extract the logarithmic negativity $\mathcal{E}$ between A and B. The generic mixedness of ρ_{AB} could arise from entanglement with environment C. Here the subsystems contain N_A, N_B and N_C particles respectively. The scheme involves three copies of the original system, and two counter propagating sets of measurements on A and B, ordered by the shown numbers, with direction depicted by the filled arrows. (b) Diagrammatic proof (for $m = 3$) of the equivalence between the moments μ_m and expectation of two opposite permutations (decomposed as swaps) on A and B – from which a measurement scheme can be derived.

with $|\cdot|$ the trace norm, $\rho_{AB}^{T_X}$ the partial transpose with respect to subsystem X, and $\{\lambda_k\}$ the eigenvalues of $\rho_{AB}^{T_X}$. Because of the non-trivial dependence of $\mathcal{E}$ on ρ_{AB}, there is no state-independent observable that can measure it — generally demanding full state tomography. The $\{\lambda_k\}$ are the roots of the characteristic polynomial, $P(\lambda) = \det(\lambda - \rho_{AB}^{T_B}) = \sum_n c_n \lambda^n$, where each c_n is a polynomial function of the partially transposed moments:

$$\mu_m = \mathrm{Tr}\left[(\rho_{AB}^{T_B})^m\right] = \sum_k \lambda_k^m. \tag{4.2}$$

In this way, full information about the spectrum $\{\lambda_k\}$ is contained in $\{\mu_m\}$. It is known that these measuring these moments is technically possible using m copies of the state and controlled swap operations [243]. However, even if these experimentally challenging operations were available, the problem of extracting $\{\lambda_k\}$ from the moments is notoriously ill-conditioned [250], with a closely related problem being described as numerically catastrophic. Alongside this, an exponential number of moments respective to the size of AB are needed to exactly solve the equations. On the other hand, to estimate the logarithmic negativity, a precise knowledge of all

λ_k is not required. Since $-\frac{1}{2} \leq \lambda_k \leq 1$ for all k [251] and $\sum_k \lambda_k = 1$, generically, the magnitude of the moments quickly decreases with m, with the first few carrying the most information. Backing up this intuition, we will show that the moments required, $\{\mu_m : m \leq M\}$, to accurately estimate the entanglement can number as few as $M = 3$. We do this by employing machine learning to *directly* map moments to logarithmic negativity, avoiding reconstruction of the spectrum or state. Note that μ_0 is simply the dimension of the systems Hilbert space, while $\mu_1 = 1$ in all cases. Additionally, it can be easily shown that μ_2 is equal to the purity of the state $= \mathrm{Tr}\left[\rho_{AB}^2\right]$, and as such, $M \geq 3$ is needed to extract any information about $\mathcal{E}$. In this sense our method is optimal in terms of number of copies.

4.4 Measuring the Moments of $\rho_{AB}^{T_B}$

The method for measuring the moments proposed in [243] based on 3-body controlled swaps is practically challenging in a many-body set-up where natural interactions are two-body. A simpler protocol, for 4 moments only, was provided in [245]. Here, we show that any moment in Eq. (4.2) can be measured using only SWAP-operators between the *individual* constituents of the m copies of the state ρ_{AB}, namely $\rho_{AB}^{\otimes m} = \bigotimes_{c=1}^{m} \rho_{A_cB_c}$. This general set-up is shown in Fig. 4.1(a), where the mixedness of ρ_{AB} arises from possible entanglement with a third system C, such that $\rho_{AB} = \mathrm{Tr}_C \left|\Psi_{ABC}\right\rangle \left\langle\Psi_{ABC}\right|$ with $\left|\Psi_{ABC}\right\rangle$ being a pure tripartite state. The first step is to write the matrix power as an expectation of a permutation operator, similar to Ref. [252, 215], but here on the partially transposed copies:

$$\begin{aligned} \mu_m &= \mathrm{Tr}\left[\left(\bigotimes_{c=1}^{m} \rho_{A_cB_c}^{T_{B_c}}\right)\mathbb{P}^m\right] \\ &= \mathrm{Tr}\left[\left(\bigotimes_{c=1}^{m} \rho_{A_cB_c}\right)(\mathbb{P}^m)^{T_B}\right], \end{aligned} \tag{4.3}$$

where $\mathbb{P}^m$ is any linear combination of cyclic permutation operators of order m and the second line makes use of the identity $\mathrm{Tr}(\rho_{AB}^{T_B}O)$=$\mathrm{Tr}(\rho_{AB}O^{T_B})$, valid for any operator O. A schematic of the equality in Eq. (4.3) for $m = 3$ is shown in Fig. 4.1(b). In Appendix B we provide a choice of $\mathbb{P}^m$ with a neat operational meaning, both

for spin and bosonic systems. For spin lattices, our choice of $\mathbb{P}^m$ to measure the moments μ_m results to the following steps in practice: (i) prepare m copies of the state ρ_{AB}; (ii) sequentially measure a 'forward' sequence of adjacent swaps, $S_A^{c,c+1}$ between neighbouring copies of system A from $c = 1$ to $m - 1$; (iii) sequentially measure a 'backward' sequence of adjacent swaps, $S_B^{c,c-1}$ between neighbouring copies of system B from $c = m$ to 2; (iv) repeat these steps in order to yield an expectation value. This procedure is also depicted for $m = 3$ in Fig. 4.1(a). For bosonic lattices, our procedure corresponds to the following steps: (i) prepare m copies of the state ρ_{AB}; (ii) Perform 'forward' Fourier transforms between modes in different copies for each site in A – this can be achieved using a series of beam splitters [253]; (iii) Perform 'backwards' (reverse) Fourier transform between modes in different copies for each site in B, via reverse beam splitter transformations; (iv) Measure the boson occupation numbers $n_{j,c}$ on all sites $j \in \{A, B\}$ and all copies c to compute $\phi = e^{i \sum_{j\in\{A,B\},c} 2\pi c n_{j,c}/m}$. (v) Repeat these steps to obtain the expectation value μ_m as an average of ϕ. Both procedures require $\mathcal{O}(N_A + N_B)$ measurements for each m between 2 and M, and are explained in detail in Appendix B. This is in stark contrast to tomography, which generically for qubit systems requires $2^{2(N_A+N_B)}$ measurement settings.

It is worth emphasizing the difference between our procedure, and recently proposed operational methods for measuring Renyi entropies [216, 218, 254]. First of all, Renyi entropies only quantify entanglement for pure states, and cannot be used in the more general mixed state scenario. Secondly, while for entropies the operations are only performed on a single subsystem, here, one performs both 'forward' and 'backward' operations on two subsystems at once, as explained above. Remarkably, even though partially transposed density matrices are generically unphysical, measurement of their moments is possible.

4.5 Machine Learning Entanglement

We focus now on estimating the logarithmic negativity from the information contained in the moments, μ_m. One approach using only the even moments has been

proposed in the quantum field theory literature [38, 255] by exploiting numerical extrapolation. However, this method neglects the odd moments and generally requires a large number of moments and thus copies. We have developed an alternative analytical method based on Chebyshev functional approximation, detailed in Appendix B, which takes into account these odd moments. Indeed with the same number of copies we find it produces more accurate estimates, and thus serves as a reference quantity. The Chebyshev expansion is analytically tractable, and becomes accurate for large enough M, as is shown in Appendix B. Nonetheless, this expansion is based on a linear mapping between the moments and the negativity, despite this relationship being inherently non-linear. Therefore it is natural to think that a non-linear transformation could be more optimal, and thus more efficient for smaller M – namely fewer copies.

Machine learning has recently emerged as a key tool for modelling an unknown non-linear relationship between sets of data. In the supervised learning paradigm, one trains a model with a set of known inputs and their corresponding outputs. Once trained, the model can then be used to predict the unknown output of new input data. Here, we take the moments μ_m as the input and the logarithmic negativity $\mathcal{E}$ as the output. Training is performed by taking a large set of states for which μ_m and $\mathcal{E}$ can be computed on a classical computer. This model can then be used to predict $\mathcal{E}$ from a set of experimentally measured moments. The experimental system under study motivates the choice of which training states to use, so that they share, for example, similar entanglement features. Among the most successful machine learning algorithms for non-linear regression are supervised vector machines [256], random decision forests [257], and deep neural networks [258, 259]. However, we have found that using the same training set for each, neural networks are superior when it comes to predicting logarithmic negativity for a wide range of states beyond the training set. As we show with our numerical results, neural networks provide a very accurate method for extracting the logarithmic negativity with as few as $M = 3$ copies. The details of our neural network construction can be found in Appendix B.

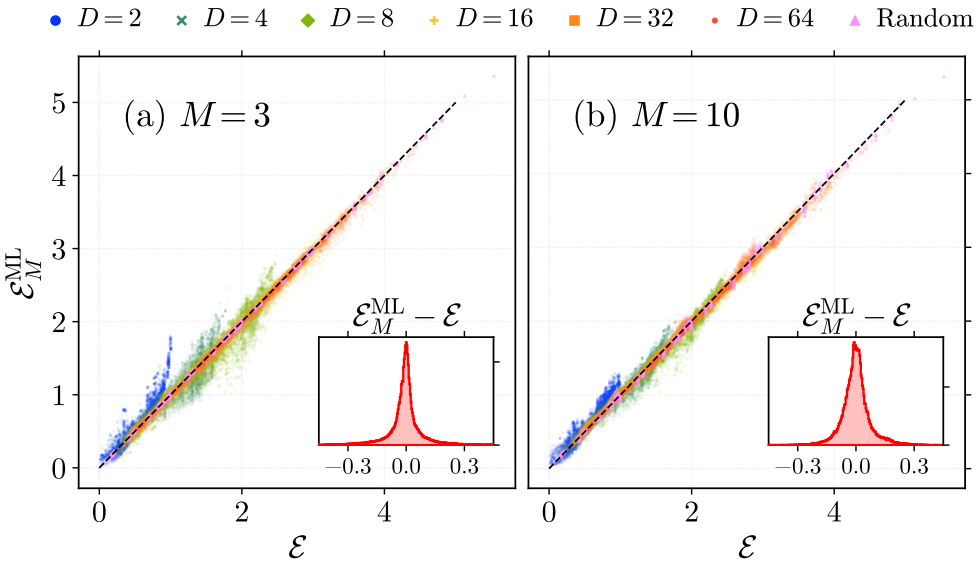


Figure 4.2: Machine learning entanglement. Estimated logarithmic negativity $\mathcal{E}_M^{\mathrm{ML}}$, using a machine learning vs. actual logarithmic negativity $\mathcal{E}$, for the same set of random states described in the main text. Training and prediction is performed using the moments μ_m generated from: (a) $M = 3$ copies; (b) $M = 10$ copies. The respective insets show the distribution of error, $\mathcal{E}_M^{\mathrm{ML}} - \mathcal{E}$.

4.6 Training with Random States

In order to train a neural network, a set of suitable training states are required for which both the moments and logarithmic negativities are known. From an entanglement perspective, relevant states in condensed matter physics can be classified as either area-law, or volume-law. In the first case, the entanglement of a subsystem A with the rest is proportional to the number of qubits along their boundary. In the second, this entanglement is instead proportional to N_A, the number of qubits in A. Area-law states arise as low energy eigenstates of local gapped Hamiltonians, with logarithmic corrections in critical systems. Volume-law states however, are associated with the eigenstates found in the mid-spectrum, and as such arise in non-equilibrium dynamics, e.g. quantum quenches [260, 261].

Rather than concentrate on a training with a specific model system, we initially consider the very general case of random states. To encompass both area- and volume-law states, we consider two classes of states $|\Psi_{ABS}\rangle$: (i) random generic pure states (R-GPS), e.g. sampled from the Haar measure, which typically have

volume-law entanglement [262, 263]; (ii) random matrix product states (R-MPS) with fixed bond dimension, which satisfy an area-law by construction [93]. In order to generate a training set with a wide range of entanglement features, subsystem sizes, and mixedness, we perform the following procedure: (i) For a fixed number of qubits N, take either a R-GPS, or R-MPS with bond dimension D. (ii) Take different tri-partitions such that $N = N_A + N_B + N_C$, and for each calculate μ_m and $\mathcal{E}$ for ρ_{AB}. (iii) Repeat for different random instances, while separately varying N and D. Further generation and training are provided in Appendix B.

4.7 Numerical Results for Random States

To check the performance of our neural network estimator, we take the set of random states described in the previous section and split this data in two, one half for training the neural network model, and the other as 'unseen' test data. In Fig. 4.2(a) we plot the machine learning model's predictions, $\mathcal{E}_M^{\mathrm{ML}}$, for the test data, using only $M = 3$ copies, in which a high degree of accuracy is achieved. In the inset of Fig. 4.2(a), we plot a histogram of the errors $\mathcal{E}_M^{\mathrm{ML}} - \mathcal{E}$, which displays a very sharp peak at zero error with standard deviation ~ 0.09 A further improvement, particularly in outliers, is achieved by increasing the number of copies M to 10, see Fig. 4.2(b), where the error standard deviation decreases to ~ 0.07 Regardless, the machine learning method is already very accurate for extracting entanglement using only three copies. The machine learning approach works particularly well for large bond dimension and volume-law like states – an important fact given that these are the exact cases where efficient tomography fails. A more detailed discussion about sensitivity and ascribing errors to machine learning predictions can be found in Appendix B.

4.8 Numerical Results for Physical Statesb

We now consider the more realistic setting of quench dynamics in a many-body system. We take a system of N spin-1/2 particles with nearest neighbour Heisenberg Hamiltonian $H = J\sum_{i=1}^{N-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1}$ with J the interaction strength and $\boldsymbol{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ the vector of Pauli matrices acting on site i. The system is initialised in the (separable) Neel-state $|\Psi(0)\rangle = |\uparrow\downarrow\uparrow \ldots\rangle$. As the chain unitarily evolves in time

as $|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle$, it becomes entangled, with an effective MPS description whose bond dimension increases until the state is essentially volume-law [260, 261].

In Fig. 4.3 we plot the evolution of $\mathcal{E}$ and three approximation methods, as functions of time for three different choices of subsystems. The three methods are the Chebyshev approximation with $M = 10$ and $M = 20$, discussed in Appendix B, and machine learning with $M = 3$, with respective approximate entanglements $\mathcal{E}^{\rm Cheb}_{M=10}$, $\mathcal{E}^{\rm Cheb}_{M=20}$ and $\mathcal{E}^{\rm ML}_{M=3}$. In Fig. 4.3(a) we consider a specific partition with $N_A = 2$, $N_B = 2$ and $N_C = 4$. Here, $\mathcal{E}^{\rm ML}_{M=3}$ and $\mathcal{E}^{\rm Cheb}_{M=20}$ are comparably accurate. For larger subsystems, as shown in Fig. 4.3(b) and (c), the machine learning approximation, using only $M = 3$ copies, significantly outperforms the Chebyshev approximations, using either $M = 10$ or $M = 20$ copies. It is remarkable that despite being trained on a arbitrary set of random states with no knowledge of the underlying physical system, the evolution of $\mathcal{E}$ is accurately captured by the neural network estimator for all partitions and times, with as few as $M = 3$ copies.

In Appendix B, we explore various other physical situations, including the groundstate of an XX-chain across its phase transition, the fully symmetric W-state, and a quench across the critical point of a transverse Ising chain.

4.9 Conclusions

The measurement of logarithmic negativity in generic multi-particle mixed states (where Renyi entropies are insufficient to quantify entanglement) has so far relied on the complete reconstruction of a quantum state, which in general requires an exponential number of measurements, and is thus limited to small system sizes. In this chapter, we have devised an alternative strategy, based on machine learning, by which we can extract the entanglement from very few measurements. These measurements are based on two counter-propagating series of swap operators on copies of the state – techniques for achieving this have already been demonstrated in a number of physical set-ups ranging from quantum dot arrays [127, 264] to cold atoms in optical lattices [265, 220]. Our method is based on learning the functional relationship between these measurement outcomes — the first few moments of the

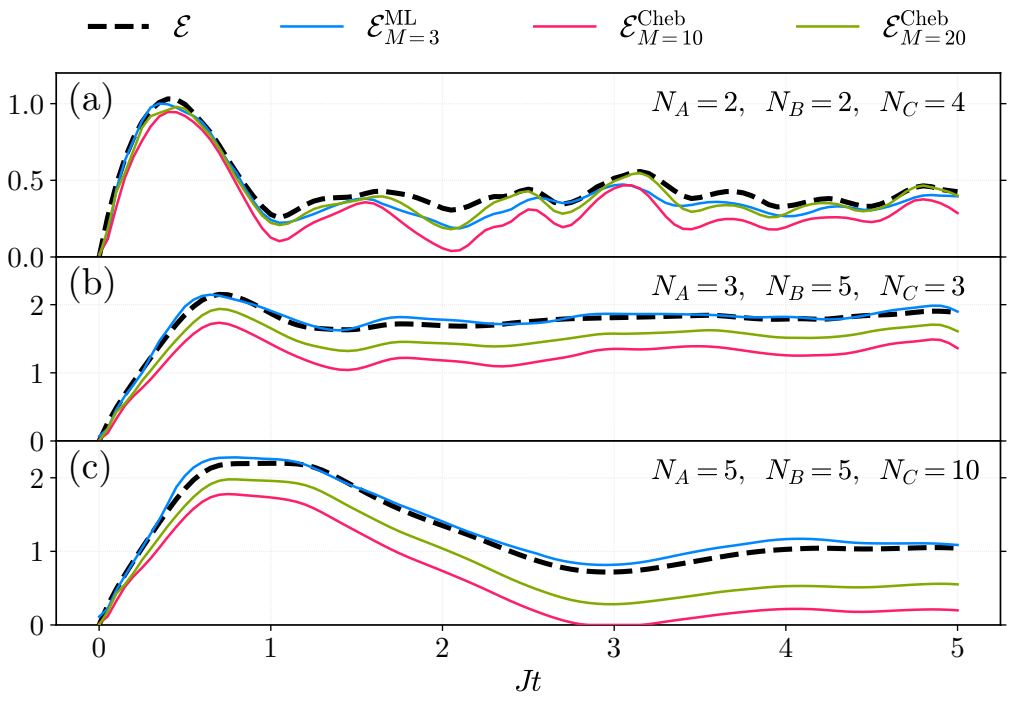


Figure 4.3: Estimating entanglement for physical states. Logarithmic negativity, $\mathcal{E}$, and its approximations, using machine learning ($M = 3$) and a Chebyshev expansion ($M = 10$, $M = 20$), as a function of time Jt for the quench dynamics of a Heisenberg spin-chain initialised in a Neel-state. A variety of system sizes with different partitions is shown here: (a) $N = 8, N_A = N_B = 2$. (b) $N = 11, N_A = 3, N_B = 5$; (c) $N = 20, N_A = N_B = 5$.

partially transposed density matrix — and the logarithmic negativity using a neural network. Remarkably, our method is already very accurate for as few as three copies – making it very resource efficient and desirable for practical applications – even for estimating the entanglement of highly entangled physical states, such as those arising in quantum quenches.

Chapter 5

Fast computation of many-body entanglement

This chapter is adapted from the preprint:

- Johnnie Gray. Fast computation of many-body entanglement. *arXiv preprint arXiv:1809.01685*, 2018.

5.1 Abstract

Mixed state entanglement measures can act as a versatile probes of many-body systems. However, they are generally hard to compute, often relying on tricky optimizations. One measure that is straightforward to compute is the logarithmic negativity, yet done naively even this is still limited to small system sizes. Here, we introduce a method to compute the logarithmic negativity for arbitrary subsystems of a densely represented state, as well as block subsystems of matrix product states. The method combines lazily evaluated, tensor network representations of the partially transposed density matrix with stochastic Lanczos quadrature, and is easily extendible to other quantities and classes of many-body states. As examples, we compute the entanglement within random pure states for density matrices of up to 30 qubits, explore scrambling in a many-body quench, and match the results of conformal field theory in the ground-state of the Heisenberg model for density matrices of up to 1000 spins. An implementation of the algorithm has been made available in the open-source library *quimb*.

5.2 Introduction

Entanglement not only plays an essential role across many aspects of quantum technologies [209, 9, 210, 211, 212, 213], but also in understanding the nature of many-body quantum systems [214, 93]. A prevalent quantity to study in this context is the *entanglement entropy*, either computationally or analytically [38, 266, 37]. However, this is only applicable to bipartitions of pure states: one can only control the ratio of subsystem sizes and cannot, for example, exclude any sort of environment. On the other hand, a true mixed-state entanglement measure allows full control over the sizes of two subsystems at once, and can thus be a much more refined probe for many phenomena [191, 196, 267, 38, 39, 268, 269, 270, 271, 2]. One drawback is that most true entanglement measures are inefficient to compute for many-body systems, even ignoring the exponential scaling of Hilbert space size, d, with system size, L. One quantity that is efficient [28], in a technical sense, is the *logarithmic negativity* [24, 25, 26, 27], though the naive computational effort still scales cubically with Hilbert space size, limiting practical calculations to $\lesssim 15$ qubits.

Here we demonstrate an efficient method to approximately but accurately compute the logarithmic negativity for subsystems of many-body quantum states. The method relies on treating the reduced density matrix as an implicit operator defined as a tensor network [93, 94, 95], and then using stochastic Lanczos quadrature [272, 273, 274] to estimate a spectral sum of this operator. We refer to the whole procedure as tensor network stochastic Lanczos quadrature (TNSLQ). The logarithmic negativity is a particular instance of the algorithm, which we target here, but other quantities such as entropy and thus mutual information are even simpler to compute. There are also many representations of many-body states amenable to a tensor network description, but we focus here on two key ones: (i) density operators derived from partially tracing densely represented pure states; and (ii) 'compressed' density operators derived from partially tracing matrix product states. Broadly speaking, the TNSLQ method enables the computation of logarithmic negativity for density matrices of $\lesssim 30$ qubits, without resorting to supercomputer-level resources. In terms of matrix product states, the equivalent limit for computing

entanglement between arbitrarily separated contiguous blocks, with open or periodic boundary conditions, is that the the bond dimension is initially $\lesssim 180$. Efficient implementations of the algorithm specifically for both of these classes of states have been added to the open source library `quimb` [5], as well as the general capability to perform TNSLQ for arbitrary tensor networks and quantities.

This chapter is organised as follows: in Sec. 5.3 we introduce the logarithmic negativity and discuss some details of its naive computation. In Sec. 5.4 we introduce stochastic Lanczos quadrature as a method to approximately compute the logarithmic negativity as the spectral sum of a linear operator. In Sec. 5.5 we introduce the basic diagrammatic notation of tensor networks. In Sec. 5.6 we show how to form an efficient partially transposed linear operator for two subsystems of an exactly represented pure state. In Sec. 5.7 we show how to do the same for two block subsystems of matrix product states, which involves a form of 'compression' first. In Sec. 5.8 we present results of using the above methods as applied to relevant physical situations. In Sec. 5.9 we analyse the error of the method and show that it is bounded by the purity of the density operator under consideration. Finally, we discuss the method's future applications and conclude in Sec. 5.10.

5.3 Logarithmic Negativity

The logarithmic negativity [24, 25, 26, 27] is an entanglement monotone and upper bound on the distillable entanglement. For a density matrix, ρ_{AB}, of two subsystems A and B with Hilbert space sizes d_A and d_B respectively, it is defined as

$$\mathcal{E}(\rho_{AB}) = \log_2 \left\| \rho_{AB}^{T_B} \right\|_{\mathrm{Tr}} \tag{5.1}$$

with $\cdot^{T_B}$ denoting the *partial transpose* [275] with respect to subsystem B and $\| \cdot \|_{\mathrm{Tr}}$ the *trace norm*[1]. Unlike the mutual information say, the logarithmic negativity quantifies quantum correlations only - one of the features that mark it out as a refined probe of many-body quantum phenomena.

We note that even if ρ_{AB} is a low-rank operator, the partial transpose operation

[1] Also known as the *nuclear norm*.

generally increases the rank by a factor of $\min\{d_A^2, d_B^2\}$, precluding the use of low-rank methods for the computation of $\mathcal{E}$. Instead, the trace norm of an operator is generally computed as the absolute sum of all eigenvalues, and as such, the full spectrum is required in the exact case. On the other hand, if much of the spectrum can be essentially described as a continuous distribution, then intuition suggests that far less information than every single eigenvalue should be required to approximate its sum. In this case it should also be possible to avoid directly forming the full, partially transposed, density operator $\rho_{AB}^{T_B}$ and instead rely only on its *action* on an arbitrary vector: $\rho_{AB}^{T_B} |\phi\rangle \to |\tilde{\phi}\rangle$. We'll call such an implicit representation simply a *linear operator*, $\hat{X}$.

5.4 Stochastic Lanczos Quadrature

Let's assume we have access to $\rho_{AB}^{T_B}$ as a linear operator, that is, we can use it to evaluate matrix-vector products. We can also recast Eq. (5.1) as the trace of a matrix function where we take the function as the *absolute* function, $|\cdot|$:

$$\mathcal{E}(\rho_{AB}) = \log_2 \mathrm{Tr}\left(\left|\rho_{AB}^{T_B}\right|\right) \ . \tag{5.2}$$

For such a *spectral sum* of a Hermitian linear operator there do indeed exist various methods to estimate the quantity, including polynomial methods [276] and approximate reconstruction of the spectrum [277] We focus here though on *Stochastic Lanczos Quadrature* (SLQ) [274], which is relatively simple to implement but also exhibits excellent performance. It can be thought of as the combination of three separate techniques:

1. **Hutchinson's trace method**[278], which estimates the trace of an operator, $f(\hat{X})$, with N inner product samples of random vectors $\{|\phi_n\rangle\}$:

$$\mathrm{Tr}(f(\hat{X})) \approx \frac{1}{N} \sum_{n=1}^{N} \left\langle \phi_n \middle| f(\hat{X}) \middle| \phi_n \right\rangle \ . \tag{5.3}$$

   This approaches the exact value, $\mathrm{Tr}(f(\hat{X}))$, in a unbiased manner as $N \to \infty$. Practically speaking, we generally need $N \ll d_A d_B$ in order to estimate the

trace to reasonable accuracy. For the purpose of the logarithmic negativity we take $f \to |\cdot|$ and $\hat{X} \to \rho_{AB}^{T_B}$ so that $f(\hat{X}) \to |\rho_{AB}^{T_B}|$.

2. **Gauss Quadrature**, which allows the estimation of the above bi-linear forms, $G_n = \left\langle \phi_n \middle| f(\hat{X}) \middle| \phi_n \right\rangle$, when transformed into a Riemann–Stieltjes integral [274]. Note that the vector $|\tilde{y}\rangle = f(\hat{X})\,|\phi_n\rangle$ is not itself directly computed at any point, which would be expensive.

3. The **Lanczos algorithm** [272], which iteratively constructs a basis for the Krylov space span$\{|\phi_n\rangle, \hat{X}^1\,|\phi_n\rangle, \hat{X}^2\,|\phi_n\rangle, \ldots\}$ using matrix-vector products only, from which the nodes and weights of the Gauss quadrature rule can be directly computed.

Details of each of these three techniques, including error analysis, are extensively addressed in various other publications [273, 274]. Instead, we simply sketch a full implementation of the SLQ method in Algorithm 1.

Algorithm 1 *Inputs:* linear operator $\hat{X}$, scalar function f, target accuracy `tol`, target Lanczos accuracy `ltol`, maximum number of repeats N_{max}, maximum Krylov subspace size K_{max}. *Outputs:* SLQ estimate of $\text{Tr}(f(\hat{X}))$

for $n = 1 : N_{\text{max}}$ **do**
  $|\phi\rangle \leftarrow$ zero mean, unit variance, random vector
  $\beta_1 = \sqrt{\langle\phi|\phi\rangle}$
  $|\phi_0\rangle \leftarrow 0$
  $|\phi_1\rangle \leftarrow |\phi\rangle / \beta_1$
  for $k = 1 : K_{\text{max}}$ **do**
    // perform a Lanczos iteration
    $|\tilde{v}\rangle = \hat{X}\,|\phi_k\rangle - \beta_k\,|\phi_{k-1}\rangle$
    $\alpha_k = \langle\tilde{v}|\phi_k\rangle$
    $|\tilde{v}\rangle = |\tilde{v}\rangle - \alpha_k\,|\phi_k\rangle$
    $\beta_{k+1} = \sqrt{\langle\tilde{v}|\tilde{v}\rangle}$
    $|\phi_{k-1}\rangle \leftarrow |\phi_k\rangle$
    $|\phi_k\rangle = |\tilde{v}\rangle / \beta_{k+1}$
    // compute Gauss quadrature weights
    $T_k = \textbf{tridiag}([\alpha_1, \ldots, \alpha_k], [\beta_2, \ldots, \beta_k])$
    $\{\theta_j\}, \{|w_j\rangle\} = \textbf{eig}(T_k)$
    compute $\tau_j = \langle e_1|w_j\rangle$ for each eigenvector $|w_j\rangle$
    $\tilde{F}_k = \sum_j \tau_j^2 f(\theta_j)$
    if `lanczos_converged`$([\tilde{F}_1, \ldots, \tilde{F}_k]$, `ltol`$)$ **then**
      break
    end if
  end for
  $G_n \leftarrow$ `lanczos_estimate`$([\tilde{F}_1, \ldots, \tilde{F}_k])$
  if `hutchinson_converged(`$[G_1, \ldots, G_n]$`, tol)` **then**
    break
  end if
end for
return $1/n \sum_{j=1}^{n} G_j$

We note that the SLQ algorithm requires storage of 3 vectors of size $d_A d_B$ only – a relatively low memory overhead. For large systems, the computational effort is generally dominated by the matrix-vector product $\hat{X}\,|\phi_k\rangle$. There are multiple options for choosing the Lanczos convergence and estimation functions `lanczos_converged` and `lanczos_estimate` [279]. We find a practical method is to least-squares fit an exponential to the values $[\tilde{F}_1, \ldots, \tilde{F}_k]$. This yields a value for the equilibrium point, as well as an uncertainty, which can respectively be taken as an estimate for the bilinear form and an error to test convergence against. For `hutchinson_converged`

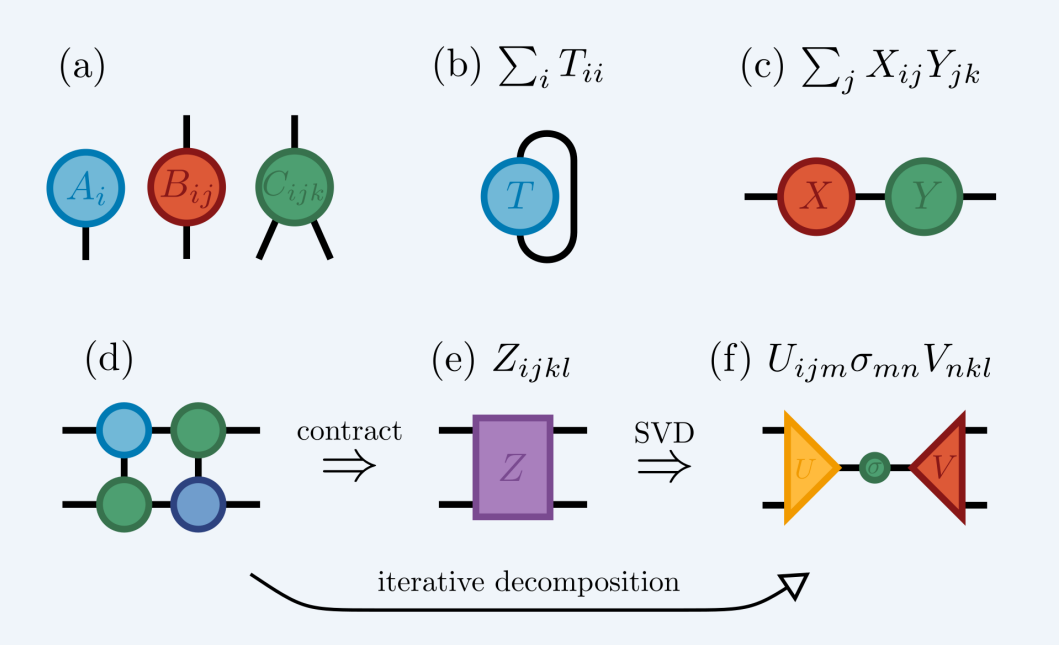


Figure 5.1: *Tensor diagrams:* (a) a vector, matrix and rank-3 tensor in graphical form; (b) the trace of a matrix; (c) tensor contraction, here matrix multiplication; (d) a small network of tensors, which could be contracted into a single tensor (e), itself then decomposed, for example via SVD, into a new network (f). The left and right indices of (d) could also be used to treat it as a linear operator. For here it could be iteratively decomposed, directly into (f).

we simply take the error on the mean. The overall procedure not only allows us to estimate quantities such as the logarithmic negativity, but reliably keep track of the error as well.

5.5 Tensor Networks & Graphical Notation

The remaining task is to find a linear operator representation of $\rho_{AB}^{T_B}$ for our target class of many-body states. We will focus here on using a *tensor network* to represent this implicit operator and perform matrix-vector products. As such we'll first briefly recap the graphical notation associated with tensor networks – more thorough reviews can be found in [93, 94, 95, 280]. The essential idea is to treat all quantum objects as *tensors*, i.e. n-dimensional objects describing linear mappings between spaces, with a labelled index for each dimension. For the purpose of finite quantum mechanics, these tensors are simply numeric arrays, $T_{ij\ldots k}$

The basic graphical notation is shown in Fig. 5.1. We depict tensors as

shapes/nodes, with a leg/edge representing each index. Scalars thus have no legs, vectors one leg, and matrices two legs. A L-body pure quantum state, $|\psi\rangle$, we can view as a rank-L tensor[2], $\psi_{1\ldots L}$. Connecting the legs of tensors implies a combined summation over that shared index – a contraction, see Figs. 5.1(b), (c). In this way networks of tensors can be built up, with the number of free legs indicating the rank of the full, lazily represented object (see Fig. 5.1(d) $\Rightarrow$ (e)). If evaluating a tensor network, it is always most efficient to perform a series of pairwise contractions, the order of which can massively affect performance. Indices can be arbitrarily grouped into new, larger indices (or if their dimension factorizes, ungrouped). This 'vectorization' allows any tensor contraction to be performed as either a vector-vector, matrix-vector or matrix-matrix product. Tensors can also be decomposed, for example via singular value decomposition (SVD), into a new tensor network – see Fig. 5.1(e) $\Rightarrow$ (f).

For any network, or sub-network, we can also mark the open indices as either 'left' or 'right' and treat the resulting object as a linear operator which maps vectorized tensors spanning one set of indices into the other. The key here is that if only the action of such a linear operator on a vector is required, then the full operator does not need to be formed, and instead, the vector can be efficiently contracted into the tensor network. This allows iterative decompositions that directly transform Fig. 5.1(d) into Fig. 5.1(f), for example. One such useful procedure is the interpolative SVD [281, 282, 283], which can be used to estimate the rank of the lazily represented operator to a certain precision, and then perform the decomposition to that target rank. And another possible procedure is of course the SLQ algorithm described above.

The TNSLQ method is thus to take a tensor network, form a lazily represented linear operator, X, by grouping indices into 'left' or 'right' sets, then perform SLQ using the fact the sampling vector (which is really a vectorized tensor) can efficiently be contracted *into* the network to estimate quantities of the form $\mathrm{Tr} f(X)$. We note that in general, such operators do not have a sparse-matrix linear operator representation, and might also be *full-rank*, in the sense that all their singular values

[2]Here we mean 'rank' as the number of indices, or dimensions, of the tensor, rather than number of non-zero singular values.

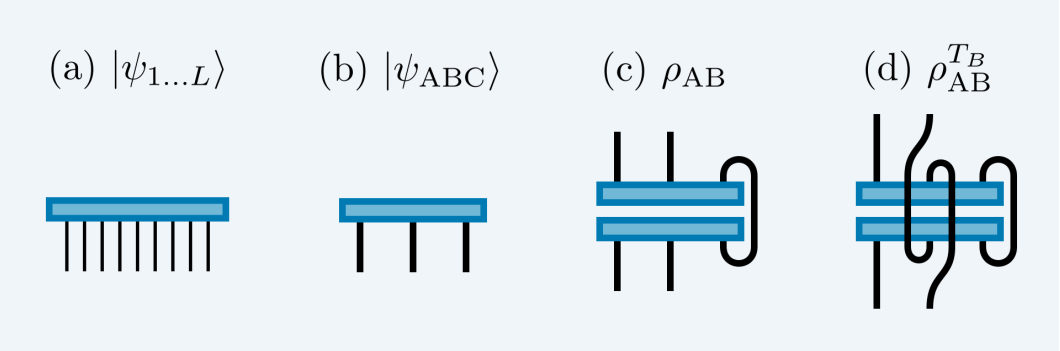


Figure 5.2: Steps to form a tensor network linear operator representation of $\rho_{AB}^{T_B}$: (a) $\rightarrow$ (b) – group indices into subsystems A, B or C; (b) $\rightarrow$ (c) – form the vector outer product and trace out system C; (c) $\rightarrow$ (d) – exchange the 'bra' and 'ket' indices of subsystem B to perform the partial transpose. For both (c) and (d) the linear operator representation of the operator is taken by grouping the upper and lower indices respectively.

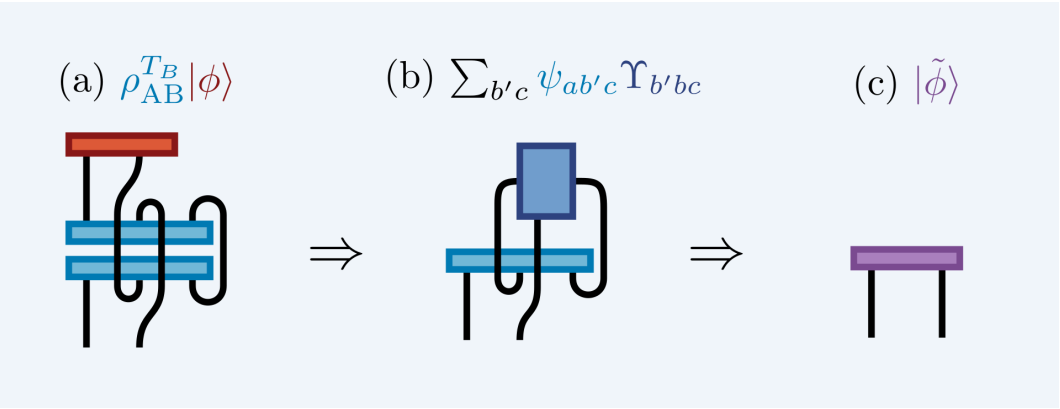


Figure 5.3: Acting on a vector with the linear operator $\rho_{AB}^{T_B}$, formed as in Fig. 5.2: (a) the full tensor network describing $\rho_{AB}^{T_B}|\phi\rangle$; (b) one of three possible intermediate contractions; and (c) the resulting output vector of this contraction. Note that although the tensor network in (a) *represents* a vector, it can not be manipulated as such until it is contracted down to the form (c), the efficiency of which drastically depends on which intermediaries are chosen.

are significant. Nonetheless, the TNSLQ method is applicable.

5.6 Partial Trace States

Having briefly introduced SLQ and tensor networks as linear operators, we now move onto specific instances of many-body quantum states with bipartite density matrix subsystems that can be described in this way. The first such example we'll call *partial trace states* (PTS). These are not genuine tensor networks in the sense

that there is no entanglement induced geometry in the initial state, but the graphical notation is useful nonetheless. The starting point is an exactly represented pure state vector – with no particular requirements on the subsystem structure. Without loss of generality we can take the many-body case of a pure L-body wave-function $|\psi_{1\ldots L}\rangle$ – a rank-L tensor (Fig. 5.2(a)). By grouping indices into either subsystem A, B or C – where we want to trace out C then find the entanglement between A and B – we get a rank-3 tensor, $|\psi_{ABC}\rangle$, of total size $d_A d_B d_C$ (Fig. 5.2(b)). The next step is to form an outer product with the conjugated state and lazily trace out subsystem C to form ρ_{AB} (Fig. 5.2(c)). Finally we partially transpose the operator by swapping the 'bra' and 'ket' indices of subsystem B to form $\rho_{AB}^{T_B}$. The advantage of keeping this operator represented as a tensor network is that is that the total storage remains $\propto d_A d_B d_C$. Whereas clearly any time that $d_C < d_A d_B$ actually performing the partial trace would increase memory usage, potentially drastically, to $d_A^2 d_B^2$.

Since we want to now perform the SLQ procedure on this lazily represented tensor network operator we need to inspect how to act with it on a vector, $|\phi\rangle$, spanning the Hilbert space of subsystems A and B. In standard tensor notation we have:

$$\rho_{AB}^{T_B}|\phi\rangle = \sum_{a',b',c} \psi_{ab'c}\psi^*_{a'bc}\phi_{a'b'} \ , \tag{5.4}$$

for which there are three possible intermediaries: (i) $\sum_{b',c}\psi_{ab'c}\Upsilon_{b'bc}$; (ii) $\sum_{a',c}\tilde{\Upsilon}_{aa'c}\psi^*_{a'bc}$; and (iii) $\sum_{a',b'}\rho_{ab'a'b}\phi_{a'b'}$; the last of which is equivalent to explicitly forming the partially traced, partially transposed density matrix. The dimensions of the subsystems determine which intermediary is best to form - in Fig. 5.3 we demonstrate performing the full contraction using the first intermediary to yield the new vector $|\phi\rangle$. Equipped with this lazy linear operator representation of the $\rho_{AB}^{T_B}$, we can now apply the SLQ procedure as detailed in Algorithm. 1 to compute the logarithmic negativity according to Eq. 5.2.

Clearly we are still limited by needing to explicitly represent the full pure state ψ_{ABC} (to $\lesssim 30$ qubits on a 'standard' desktop computer). However, the need to explicitly represent the full operator $\rho_{AB}^{T_B}$ is lifted, allowing the computation of entanglement for *any* tri-partition of A, B and C. Take for example the scenario

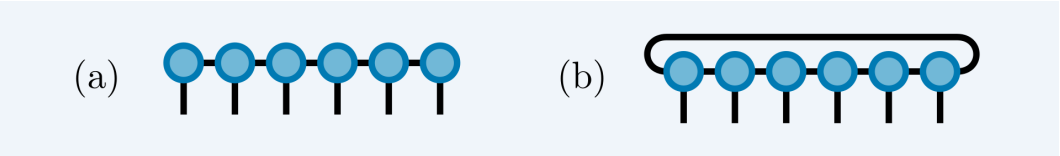


Figure 5.4: A $L = 6$ MPS decomposition of a wavefunction with: (a) open boundary conditions, where the edge tensors are only rank-2; and (b) periodic boundary conditions.

where $L = 30$ and each subsystem consists of 10 qubits. Forming $\rho_{AB}^{T_B}$ would require about 16 terabytes of memory, let alone the time to fully diagonalize it, with the situation becoming even more extreme as we decrease the size of C. On the other hand, with this lazy TNSLQ method, it is an easily tractable computation without a super-computer.

5.7 Matrix Product States

In order to move beyond full Hilbert space representations of many-body states we need a genuine tensor-network decomposition. The most useful and widespread of these is that of the *matrix product state* (MPS), which factorizes the wavefunction into a one-dimensional chain of rank-3 tensors. This ansatz efficiently represents one-dimensional states with area-law entanglement [284] and is the central representation in successful algorithms such as density matrix renormalization group (DMRG) [98, 93] and time evolving block decimation [99]. The form can be explicitly defined as

$$\psi_{abc\ldots z} = \sum_{\alpha,\beta,\gamma,\delta\ldots,\zeta} A_{a\alpha\beta} B_{b\beta\gamma} C_{c\gamma\delta} \ldots Z_{z\zeta\alpha} \tag{5.5}$$

for tensors $A, B, C, \ldots, Z$ with physical indices $a, b, c, \ldots, z$, but it is generally more concise to reason with the graphical notation as depicted in Fig. 5.4. For simplicity we will consider the size of all the physical indices to be p, and the size of all the virtual indices, $\alpha, \beta, \gamma, \delta \ldots, \zeta$, (the *bond dimension*) to be the same value, χ. The index α can be taken as size 1 (and thus ignored) for open boundary conditions - Fig. 5.4(a) - or χ for periodic boundary conditions - Fig. 5.4(b).

Given an MPS with target subsystems A and B to find the entanglement between, is is straightforward to form a tensor network of $\rho_{AB}^{T_B}$. The steps as are follows: (i)

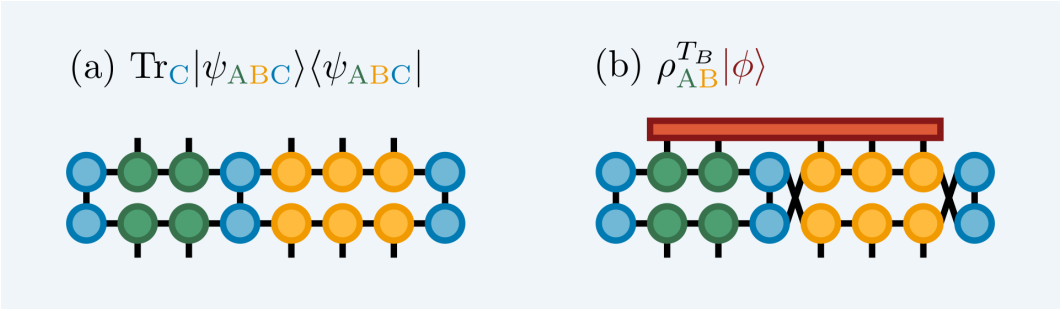


Figure 5.5: (a) Forming the reduced density matrix ρ_{AB} from matrix product state ψ_{ABC}. Here subsystem A is in green, subsystem B is in yellow, and subsystem C is in blue. (b) Tensor network representation of acting on a dense vector, $|\phi\rangle$ (in red), with $\rho_{AB}^{T_B}$ derived from the matrix product state ψ_{ABC}.

form the outer product between a 'ket' and 'bra' of the state; (ii) perform the partial trace of environment C by contracting (joining) all physical indices not contained in subsystems A or B (shown in Fig. 5.5(a)); and (iii) perform the partial transpose by switching the 'ket' indices with the 'bra' indices of all the physical sites in either subsystem A or B. At this point we could directly form a linear operator by grouping all the 'ket' indices and 'bra' indices respectively. In this case, to perform the SLQ procedure we would then need to sample this operator using a dense vector of size $2^{L_{AB}}$ (for qubits), as shown in Fig. 5.5(b). Here, we are now limited $L_{AB} \lesssim 30$ (rather than total length L in the pure state subsystem case).

The above assumes nothing about the geometry of A and B within the MPS. However, if we assume that A and B are contiguous blocks (as is often the case), then we can adapt the method for arbitrarily many sites by *compressing* each block, a process sketched in Fig. 5.6. First, a 'lateral' compression of A, B, and, if necessary, any contiguous blocks of C that form the environment. This is the method derived in [285] in order to efficiently address periodic boundary DMRG in the language of MPS. Secondly, a 'vertical' decomposition of subsystems A and B to reintroduce new effective physical indices to the density matrix [286, 287]. The details of the lateral compression as are follows:

1. Form the transfer matrix of a contiguous section to be compressed – Fig. 5.6(a)⇒(b).

2. Perform a iterative SVD decomposition of the transfer matrix, treating it as

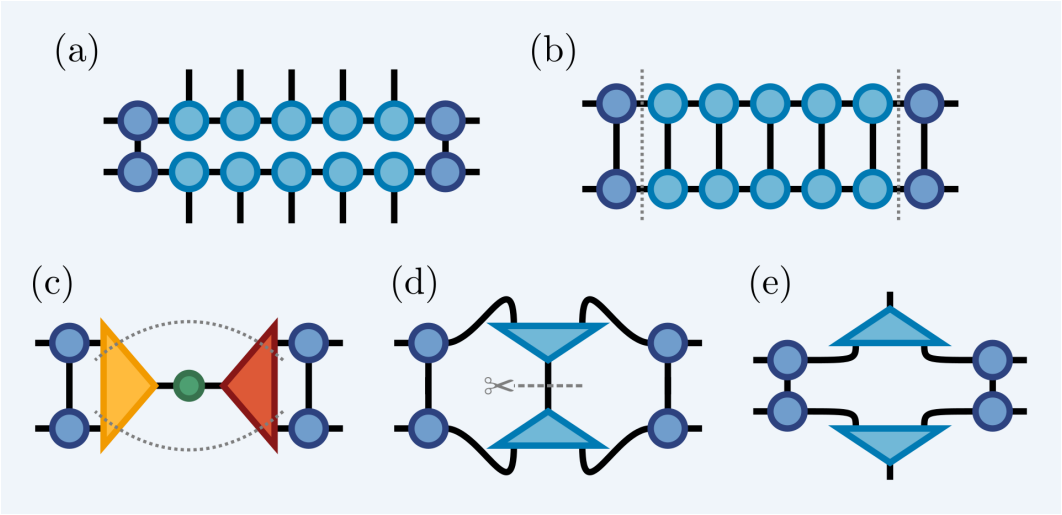


Figure 5.6: Compressing a MPS section in four steps. (a) ⇒ (b): the overlap of the target section, or transfer matrix, is formed. (b) ⇒ (c): this transfer matrix is laterally compressed using, for example, an iterative SVD algorithm (dotted grey lines in (b) denote the 'left' and 'right' groupings of indices). The singular values (central green tensor in (c)), can be absorbed into the left (yellow) or right (red) tensors once small values have been trimmed. (c) ⇒ (d): this newly compressed section is vertically decomposed, for example using a Cholesky decomposition (dotted grey lines in (c) denote the new 'left' and 'right' groupings of indices). (d) ⇒ (e): the bond between the new symmetric factors is 'cut' in order to expose the new physical index.

a linear operator with effective dimensions $\chi^2 \times \chi^2$ by grouping the left and right bonds respectively – Fig. 5.6(b)⇒(c). Note that generally, the longer a section is, the fewer the number of singular values required to represent its transfer matrix to high precision.

The procedure for the 'vertical' compression, which only is performed on subsystems A and B in order to reintroduce physical indices is as follows:

1. Perform a decomposition of the section, which now might be in SVD form, but this time grouping the *upper* and *lower* bonds respectively – Fig. 5.6(c)⇒(d). This operator, with effective dimensions $\chi^2 \times \chi^2$, is generally full-rank, however, we note that it is also positive symmetric, and thus the fast (compared to SVD) Cholesky decomposition can be used.

2. 'Split' the bond connecting the two symmetric factors simply by re-indexing the tensors – Fig. 5.6(d)⇒(e). This re-introduces effective 'ket' and 'bra'

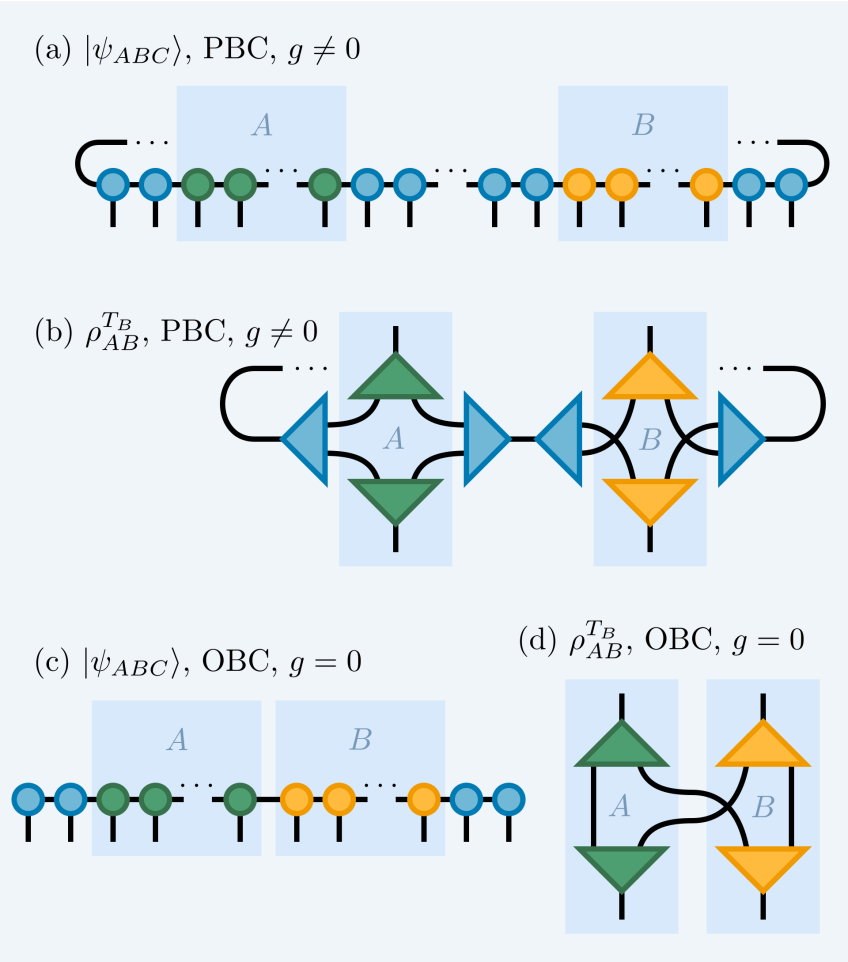


Figure 5.7: Forming a compressed representation of a partially transposed density matrix, $\rho_{AB}^{T_B}$, derived from a MPS, $|\psi_{ABC}\rangle$, namely, lateral compression of all contiguous sections, followed by vertical decomposition of A and B. (a) and (b): state and derived operator for PBC and non-zero gap, g, between contiguous subsystems A and B – the most general geometry. (c) and (d): state and derived operator for OBC with zero gap between subsystems A and B – a common geometry. The effective linear operator of $\rho_{AB}^{T_B}$ is formed by grouping the upper and lower indices of (b) and (d).

physical indices to the section, with size $\leq \chi^2$ rather than exponential in the number of sites.

With these two steps we have a method to derive a 'compressed' representation of the partially traced, partially transposed density operator $\rho_{AB}^{T_B}$, from a MPS with A and B contiguous blocks separated with gap g, as shown in Fig. 5.7(a). First partition the state into A, B, and potentially several C sections, then perform lateral compression on any of these that are long enough for it to make sense (e.g. $\chi^2 < p^{L_A}$

for section A). Next, perform vertical decompositions on subsystems A and B, and finally swap the 'ket' and 'bra' indices on subsystem B to effect the partial transpose. This resulting tensor network, in the most general geometry, is shown in Fig. 5.7(b). The largest tensor it contains is always of size $\leq \chi^4$.

We note simplifications can be made to the network in several common scenarios. If open boundary conditions (OBC) are used, the gauge freedom can be utilized to eliminate both the left and right environments completely. Similarly, if either subsystem A or B contains the end of the chain, they can be represented as two identity tensors. For periodic boundary conditions (PBC), the left and right environments are the *same* section, and an effective gauge to eliminate them can only be introduced if the section's transfer matrix has a single dominant singular value – i.e. it is separable. Finally, clearly if A and B are adjacent ($g = 0$), no environment is needed separate them. Given this MPS-derived, compressed, tensor network representation of $\rho_{AB}^{T_B}$, we can as before apply the SLQ method to this operator using a sample vector $|\phi\rangle$, also of size $\leq \chi^4$, to compute the logarithmic negativity of arbitrary contiguous sections. As before, being able to contract the sample vector into the network to yield a new vector, rather than first contracting the full operator, yields the key efficiency saving. The best contraction order depends on the various index dimensions, and in practice, we choose the order automatically using a greedy approach [6]. Bond dimension now becomes the limiting factor of the algorithm, and if we translate the memory requirement of densely representing 30 qubits into this language, we find that $\chi \lesssim 180$ is the equivalent limit.

5.8 Results

We now move on to demonstrating the TNSLQ method in three different scenarios. The first two results involve 'partial trace states' – random pure states and a many-body quench – for which $L \leq 30$ and the entanglement varies from zero to highly-entangled. The third studies the scaling of entanglement in a large matrix product state, namely, the ground-state of the Heisenberg Hamiltonian acquired using DMRG, for which analytic results are available. All computations were performed using the

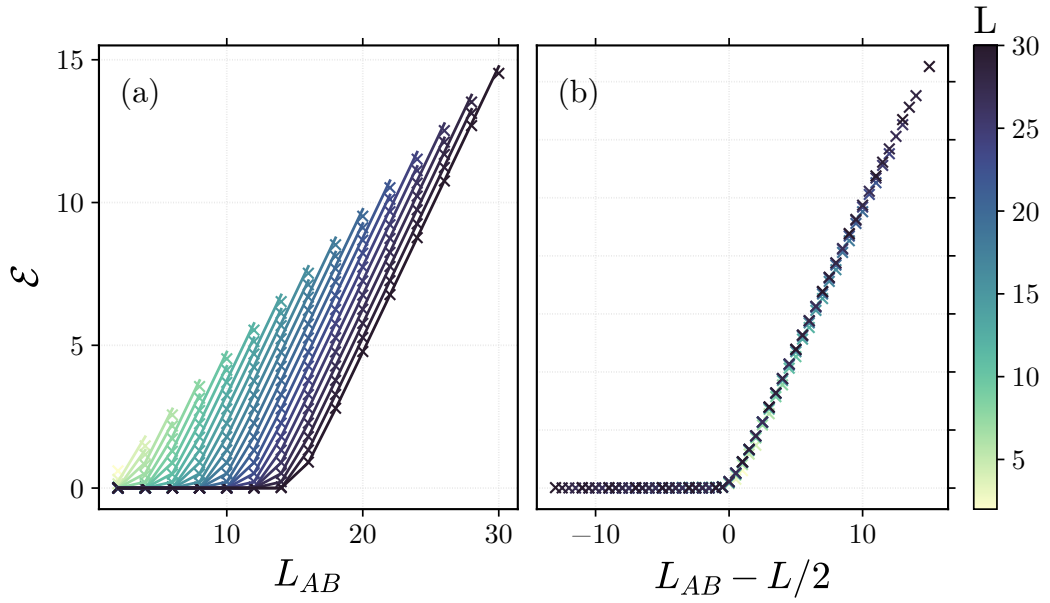


Figure 5.8: (a) Logarithmic negativity, $\mathcal{E}(\rho_{AB})$, between two equal subsystems of combined size L_{AB} in a random pure state of total size L. The lines denote the analytical result, Eq. (5.6), derived in [288] while the crosses show results computed with TNSLQ. (b) The same but offset by $L/2$, where we see a clear collapse to universal behaviour across all lengths L.

open-source library `quimb` [5], which has implementations of tensor network linear operators, the SLQ algorithm, and two-site DMRG.

5.8.1 Random pure states

First, we benchmark the TNSLQ method for density matrices derived from random pure states, in full dense representation, of length up to $L = 30$ (total Hilbert space size $2^{30} = 1073741824$). Since these states are completely permutationally symmetric, the only variables are the size of L_A, L_B and L. We simply take $L_A = L_B = L_{AB}/2$, then compute $\mathcal{E}$ for varying L_{AB} and L. For each configuration we average over 10 different random realizations, though there is very little variance between them. The analytic result for these states is known [288]:

$$\mathcal{E}(\rho_{AB}) = \log_2 \left[\frac{2}{\pi} \sin^{-1}\left(\frac{1}{\tilde{R}}\right) + \frac{2\left(1 + 2\tilde{R}^2\right)}{3\pi\tilde{R}} \sqrt{1 - \frac{1}{\tilde{R}^2}}\,, \right] \tag{5.6}$$

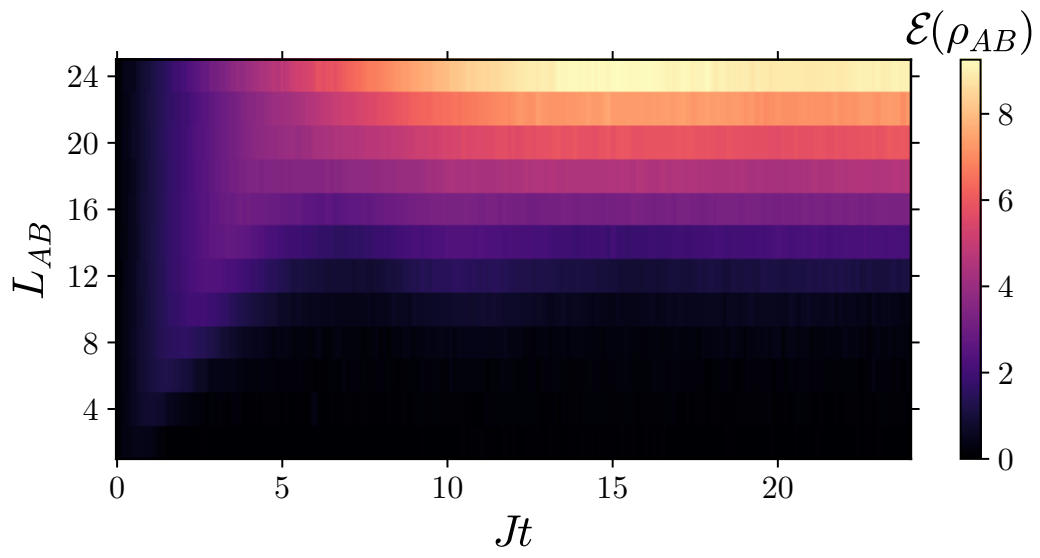


Figure 5.9: The entanglement, $\mathcal{E}(\rho_{AB})$, as computed with the TNSLQ method, across a central cut for a quench in the Heisenberg model as a function of time Jt and reduced subsystem size L_{AB} The total chain length is $L = 24$, and we take the reduced density operator ρ_{AB} to be centered around the half way point such that either side there is environement of length $L - L_{AB}/2$.

where $\tilde{R} = 2\sqrt{d_A d_B / d_C}$, which we compare to in Fig. 5.8(a). We find very good accordance with the analytic prediction, and confirm a universal behaviour whereby the entanglement is zero for $L_{AB} < L/2$, and rises linear afterwards – see Fig. 5.8(b). We note that with the standard method of computing $\mathcal{E}$, approximately the right half of Fig. 5.8(a) ($L_{AB} \geq 15$) would not be available.

5.8.2 Scrambling in a Quench

We next move on to applying the TNSLQ method to a more physical example - the time evolution of a state after a quench with an interacting many-body Hamiltonian. We take a system of L spin-1/2 particles with nearest neighbour Heisenberg Hamiltonian:

$$\hat{H} = J \sum_{i=1}^{L-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} \,, \tag{5.7}$$

where J is the interaction strength and $\boldsymbol{\sigma}_i = (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ the vector of spin operators matrices acting on site i. The system is initialized in the (separable) Neel-state $|\Psi(0)\rangle = |\uparrow\downarrow\uparrow \ldots\rangle$ and evolved using integration according to the equation $\frac{\partial}{\partial_t} |\psi(t)\rangle = -i\hat{H} |\psi(t)\rangle$, where we have set the Planck constant $\hbar = 1$. In terms of

geometry, we choose A and B as neighbouring blocks of equal length, either side of a central cut in the chain, of total length L_{AB}. In Fig. 5.9 we plot the logarithmic negativity, $\mathcal{E}(\rho_{AB})$, computed using TNSLQ, for this set-up as a function of time t, for a chain of total length $L = 24$. Again, approximately the upper half of this figure would not be computable using the exact method of calculating $\mathcal{E}$, but for those sizes that are, we find very good accordance (not shown) within the target precision of 1% for the TNSLQ method.

In relation to scrambling [289], we expect information describing the initial system to quickly ($Jt \lesssim L$) de-localize. More specifically, consider a local operator M_0 encoding some initial information such as 'spin x is up'. Theoretically, at *any* later time, the initial expectation value of this might be retrieved by measuring $M_t = U(t)M_0U(-t)$. However, if M_t after some time has significant support on more sites than is feasible for us to measure, then this information has effectively become lost. The fact that local measurements can no longer reveal such information implies the eventual build up of many-body entanglement in *large* degrees of freedom only. This also means that for sufficiently short length-scales, the entanglement should grow and *then* decrease, as the combined subsystem ρ_{AB} becomes increasingly entangled with C, precluding entanglement between A and B. This is exactly what we see in Fig. 5.9, entanglement growing and then dying at increasing length-scales, such that a subsystem of size L_{AB} is 'scrambled' after time $L_{AB}/2$. Eventually the system 'equilibrates' with an entanglement structure roughly similar to that of a random state – zero entanglement if $L_{AB} < L/2$, then rising roughly linearly as $L_{AB} \to L$.

Note that the model of Eq. (5.7) is in fact integrable which has several important implications. Firstly, the dynamics here are unusual - we do not expect normal thermalization and certainly not *fast* scrambling. Nonetheless we do see the de-localization of information associated with normal scrambling. Secondly, in thermodynamic limit, the dynamics of the logarithmic negativity have in fact been solved analytically [290] including for the initial state chosen here - the Neel-state. In this case, for large subsystems (which should nonetheless still much smaller than the full system), ones finds linear growth of the entanglement, followed abruptly by

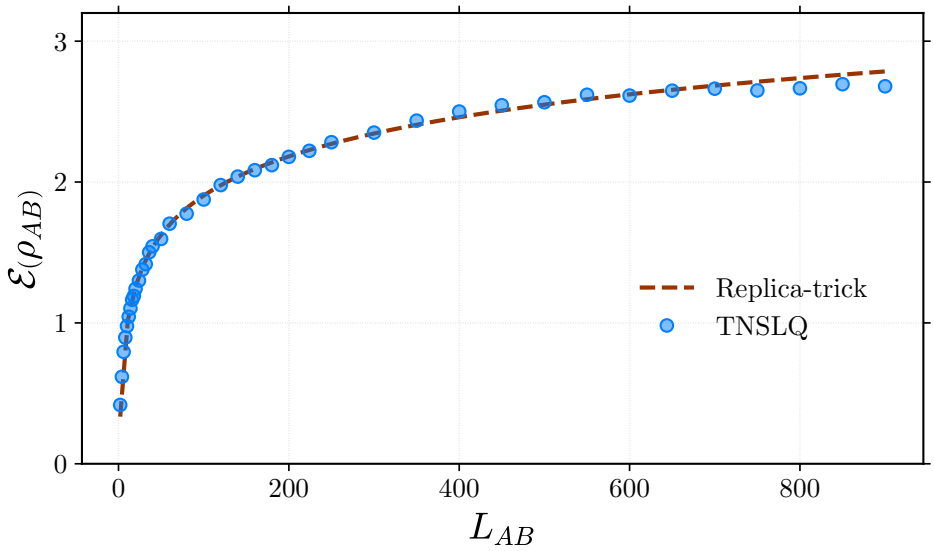


Figure 5.10: Logarithmic negativity, $\mathcal{E}$, between two neighbouring contiguous blocks of total size, L_{AB}, at the center of system of total length $L = 1000$. Blue markers show the entanglement computed using the TNSLQ method. The red dashed line shows a fit (using data where $L_{AB} \leq L/2$) to the conformal field theory prediction as given in Eq. (5.8).

linear decay to zero. A 'rounded' version of this behaviour is exactly what we find in Fig. 5.9 for $L_{AB} \lesssim 10$.

5.8.3 Heisenberg ground-state

Finally, we demonstrate the TNSLQ method for MPS subsystems by studying the logarithmic negativity for two adjacent blocks in the ground-state of the Heisenberg model, with Hamiltonian as defined in Eq. (5.7). The analytic form of this has been derived using conformal field theory and the 'replica trick' [37, 38, 39]. The behaviour, which is universal, is logarithmic scaling of the entanglement with block size, as given by

$$\mathcal{E}(\rho_{AB}) = \frac{c}{4} \log_2 (L_{AB}) + K \tag{5.8}$$

for central charge c and constant K. We take $L = 1000$ then generate a MPS description of the ground-state of $\hat{H}$ using two-site DMRG [98, 93]. The state has a maximum bond dimension of $\chi = 116$ at the center, with χ also remaining above 90 for $\sim 80\%$ of the chain. Since we are using OBC and neighbouring blocks, so that $g = 0$, the form of the compressed version of $\rho_{AB}^{T_B}$ is that of Fig. 5.7(d). The

TNSLQ computed entanglement, $\mathcal{E}(\rho_{AB})$, is shown as a function of L_{AB} in Fig. 5.10. Also plotted is a fit of Eq. (5.8) – found using $10 \leq L_{AB} \leq 500$ to avoid as much as possible the effect of finite size. From this fit of we find $c = 1.11 \pm 0.01$ and $K = 0.05 \pm 0.01$. Notably, the analytic result [38] in the thermodynamic limit is in fact $c = 1$. While the results in Fig. 5.10 closely match the expected logarithmic form, there is a small but significant discrepancy in this coefficient. This likely originates from the fact that for open boundary conditions, there is no exact generalization of the result in Eq. (5.8) to finite sizes, unlike in the periodic case where one can define a modified arc length. As such, even for $L = 1000$, we might be quite far from the asymptotic behaviour where ideally both $L \to \infty$ and $L_{AB} \ll L$. In Ref. [286], a similar discrepancy was also found specifically for the case of finite open chains.

5.9 Error Analysis

The TNSLQ method is fundamentally a stochastic process and thus comes with a certain limitation on achieving very high precision estimates. In fact, the effort scales exponentially with the number of decimal places required [291]. Crucially however, and as shown by our results above, a constant level of precision of 0.1 - 1% is easily achievable, and for many simulation purposes, completely sufficient. Moreover, the error on the estimate is easy to keep track of. To put this on more concrete terms, consider that the variance of a single estimate, G_n, of operator $\hat{Y} = f(\hat{X})$, using Hutchinson's trace method is bounded by [278, 291]:

$$\text{Var}(G_n) = 2\text{Tr}(\hat{Y}^\dagger \hat{Y}) - 2\sum_i \hat{Y}_{ii}^2 \,. \tag{5.9}$$

Ignoring the second term, which is strictly negative and thus beneficial, we can assess the first for $\hat{Y} \to |\rho_{AB}^{T_B}|$. Since $\rho_{AB}^{T_B}$ is Hermitian it follows that:

$$\begin{aligned} \text{Tr}(\hat{Y}^\dagger \hat{Y}) &= \text{Tr}(|\rho_{AB}^{T_B}||\rho_{AB}^{T_B}|) \\ &= \text{Tr}((\rho_{AB}^{T_B})^2) \\ &= \text{Tr}(\rho_{AB}^2) \end{aligned} \tag{5.10}$$

which is simply the *purity* of the joint density matrix, whose value lies between $(d_A d_B)^{-1}$ and 1. Interestingly, this implies that the entanglement will be easier to compute the *more* mixed a state ρ_{AB} is. Even more importantly, the upper limit on the variance is constant. By substituting Eq. (5.10) into Eq. (5.9) we find $\mathrm{Var}(G_n) \leq 2$, and thus there are no hidden costs of scaling to larger system sizes. Additionally, as one estimates a quantity with TNSLQ, the error on the estimate can be tracked simply as the standard error on the mean of the N actual estimates computed so far, $\{G_1, G_2, \ldots, G_N\}$, yielding estimated error $\sqrt{\mathrm{Var}(\{G_i\})/N}$. As such, the TNSLQ method for quantities based on ρ_{AB} yield errors which have the desirable properties of being both well-controlled and readily accessible.

5.10 Discussion

We have seen that the TNSLQ method enables the fast computation of many-body entanglement for various states. It involves treating a tensor network representation of the partially traced, partially transposed density matrix, $\rho_{AB}^{T_B}$ as a linear operator. The action of this operator can be efficiently evaluated by contracting the sample vector into it, allowing one to use the SLQ procedure to compute any quantity of the form $\mathrm{Tr} f(A)$. The entanglement negativity is one such quantity when we set $A = \rho_{AB}^{T_B}$ and $f = \mathrm{abs}$. We note that since generally these operators do not have equivalent efficient representations as either sparse matrices or low-rank operators, both the tensor network description and SLQ procedure seem necessary components. We have focussed particularly on the entanglement of bipartite density matrices derived as subsystems of larger, pure states, but we note that the method should be just as applicable to the situation where one begins with an efficient tensor network representation of a mixed state.

For pure states represented as vectors in their full Hilbert space, the TNSLQ method enables the efficient computation of the logarithmic negativity between subsystems A and B for *any* tri-partition of $|\psi\rangle$, that is, with any choice of environment C. Roughly speaking, in many simulations this doubles the size L_{AB} for which it is tractable to compute $\mathcal{E}(\rho_{AB})$. For MPS, we combined a method of efficiently finding

a compressed form of the bipartite reduced density matrix, ρ_{AB}, with the SLQ procedure used to then find the logarithmic negativity. The method is tractable for OBC or PBC as well as disjoint blocks separated by length g. We note that unlike some previous studies, at no point do we have to arbitrarily curtail the number of states kept (which likely introduces a systematic error), as long as the initial bond size, $\chi \lesssim 180$. To move beyond this limit, it might be worth exploring the actual effect of limiting certain bond sizes, for example in the vertical decomposition of the A and B subsystem sections into symmetric Cholesky factors. Another interesting avenue is whether one could store the Lanczos sampling vectors $\{|\phi_n\rangle\}$ in an efficient form - an obvious choice being as MPS. In this case, during the SLQ procedure the bond dimension would steadily rise, probably requiring the restriction to a fixed bond size manifold – how this might bias the estimate is not clear. A relevant approach was taken recently in [292], where both the target operator X is a *matrix product operator* (MPO), as well as sampling unitaries used to perform a block Lanczos procedure. While it is simple to form a MPO representation of $\rho_{AB}^{T_B}$ from a MPS, this has an increased storage cost $Lp^2\chi^4$, and the bond dimension of the sampling unitaries must also be artificially restricted, making the MPO approach potentially unsuitable in this particular instance.

The TNSLQ method is easily capable of estimating quantities to the level of 0.1-1%, but, as a fundamentally stochastic process, it might not be suitable for computing quantities to many digits of precision. On the other hand, we have shown that the error for the density operator $|\rho_{AB}^{T_B}|$, in terms of the variance of individual estimates, is both well controlled - being bounded by a system size independent constant - and easy to keep track of. Moreover, this feature of being an average over many low-precision estimates, as well as having a low memory-overhead, makes the TNSLQ method easy to accelerate. Firstly, it is trivial to parallelize the algorithm over independent random estimates. Secondly, single precision arithmetic can be used, for which graphical processing units (GPUs) are particularly suited. Implementations of both of these accelerations have been incorporated into the open-source library `quimb` [5].

Finally, we note that although the logarithmic negativity of 'partial trace states' and matrix product states are particular instances of TNSLQ, there are plenty of other potential candidates for the both the tensor network operator and computed function $\mathrm{Tr} f(\cdot)$. For instance, the Von Neumann entropy is given by $-\mathrm{Tr}\rho\log_2\rho$. It is easy to simplify the PTS and MPS procedures presented above, by removing the partial transpose and considering a single subsystem only, to compute this and hence, for example, the mutual information. Other suitable functions include the partition function, $Z = \mathrm{Tr}(e^{-\beta\hat{H}})$ for tensor network Hamiltonian $\hat{H}$, and the Frobenius norm $||\hat{X}||_F = \sqrt{\mathrm{Tr}(\hat{X}^2)}$ for Hermitian tensor network $\hat{X}$. In terms of other many-body quantum states, the TNSLQ method should be trivially applicable to tree tensor networks [293, 294] and multi-scale entanglement renormalization ansatz states [295, 296]. Furthermore, state of the art classical simulations of quantum computation have also recently relied on tensor network descriptions of the full circuit[297, 298, 299, 300, 301]. Even without developing any compression schemes specific to these structures, the computation of $\mathcal{E}(\rho_{AB})$ for L_{AB} up to ~ 30 should now be possible. Needless to say, in all the above cases there are many interesting questions that might be probed with a genuine, many-body entanglement measure such as the logarithmic negativity.

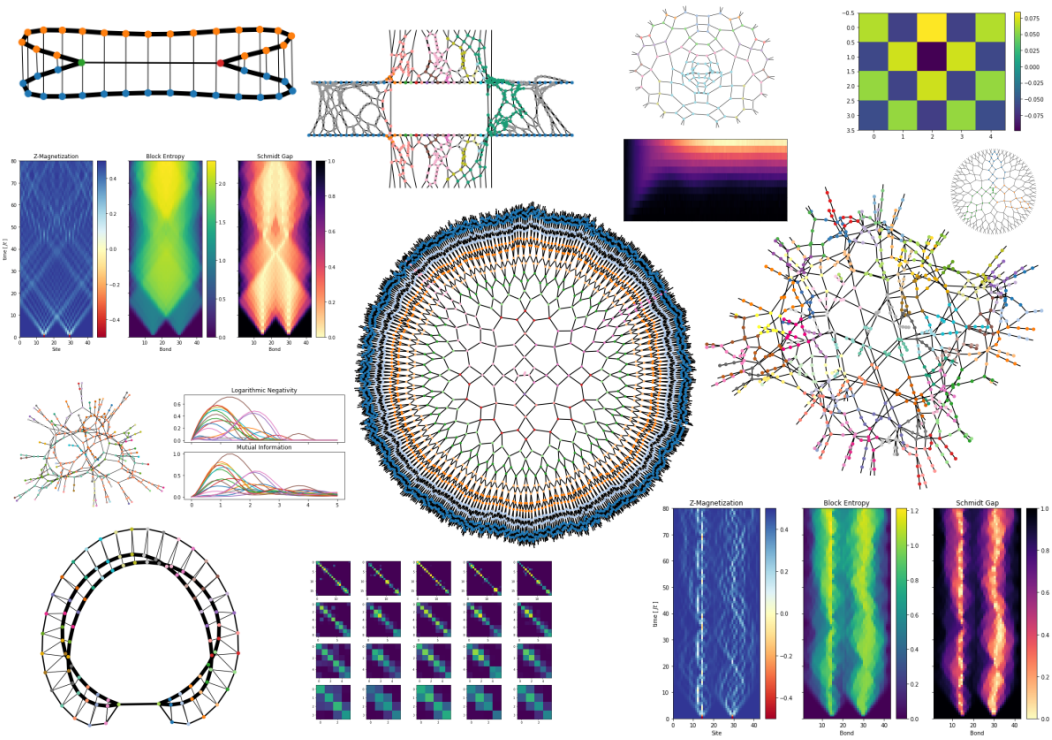


Figure 6.1: Sample output from the open-source python library `quimb` for quantum information and many-body calculations including tensor networks.

Chapter 6

`quimb` - A library for quantum many-body calculations

In this chapter we introduce `quimb` , an open-source python library for quantumm information and many-body calculations written during the course of this thesis. It contains practical implementations of many parts of the previous chapters and was used to produce the majority if not all the numerical results. Additionally, it contains

various novel features and algorithms not found anywhere else, including the TNSLQ method from chapter 5. Being python, it is very easy to use `quimb` interactively in a range of environments. However it can also leverage powerful linear algebra libraries under the hood that making its performance generally excellent. Moreover, as a open source project it conforms to many modern scientific software development best practices – including unit testing, continuous integration and automatic documentation. The following resources on `quimb` are available online:

- Main documentation – https://quimb.readthedocs.io.
- Source code – https://github.com/jcmgray/quimb.
- Short published paper – [5].

A selection of graphs produced by `quimb` are shown in Fig. 6.1.

Discrete quantum systems are in some ways straightforward to begin to simulate classically since the essential language of linear algebra has very strong support across many programming languages. However the curse of dimensionality means that much care must be taken in terms of performance when trying to scale simulations to many-body systems. `quimb` approaches this is two ways. Firstly, a core, 'exact', module with support for dense and sparse representations of quantum objects as matrices that can scale to distributed systems. Secondly, a tensor-network module for simulations of large many-body systems with moderate amounts of entanglement. In what follows we demonstrate basic usage of both parts, noting design choices where appropriate.

6.1 The core module

6.1.1 Basic representation and operations

For dense numerical linear algebra python has one stand-out library – `numpy` [302] and we use its n-dimensional array as our core object, with column vectors representing kets, row vectors representing bras, and matrices representing operators:

```
>>> import quimb as qu  # import the quimb 'namespace'

>>> data = [1 / 2**0.5, -1j / 2**0.5]  # data we want in 'quantum' form
>>> qu.ket(data)  # cast the data as a ket
qarray([[ 0.707107+0.j       ],
        [-0.       -0.707107j]])

>>> qu.bra(data)  # cast the data as a bra
qarray([[ 0.707107-0.j       , -0.       +0.707107j]])

>>> qu.dop(data)  # cast the data as a density operator
qarray([[ 0.5+0.j , -0. +0.5j],
        [ 0. -0.5j,  0.5+0.j ]])
```

Quantum objects can act on each other with the dot product using the `@` symbol (with the `.H` attribute performing the conjugate transpose):

```
>>> psi_1 = qu.rand_ket(7)
>>> psi_2 = qu.rand_ket(7)
>>> (psi_1.H @ psi_2)  # state overlap ⟨ψ1|ψ2⟩
qarray([[0.064982-0.209516j]])

>>> A = qu.rand_herm(7)  # a random hermitian matrix
>>> (psi_1.H @ A @ psi_2)  # operator expectation ⟨ψ1|A|ψ2⟩
qarray([[-0.082228+0.150639j]])
```

States and operators can be tensored together using the & symbol:

```
>>> X = qu.pauli('X')
>>> Z = qu.pauli('Z')
>>> XZ = X & Z  # like X ⊗ Z
>>> XZ.shape
(4, 4)
```

There are many other built-in states and operators, most of which can be constructed optionally in sparse form. The underlying sparse functionality is provided by `scipy` [303]. For performance reasons it is also possible to specify the data-type of the underlying array:

```
>>> # construct a sparse, single precision, random positive matrix
>>> P = qu.rand_pos(100, sparse=True, dtype='float32')
>>> P
<100x100 sparse matrix of type '<class 'numpy.float32'>'
  with 307 stored elements in Compressed Sparse Row format>
```

Even though `numpy` and `scipy` are themselves largely written in C and offer very good performance within their own routines, `quimb` further accelerates and parallelizes many functions using `numba` [304] – including for example random number generation, a key part of the TNSLQ algorithm.

6.1.2 Multipartite states and partial traces

Partially tracing and various other quantities make used of specified `dims` – a list of integers representing the size of each subsystem:

```
>>> # construct a 12 qubit w-state
>>> psi = qu.w_state(12)
>>> dims = [2] * 12  # [2, 2, 2, ...]

>>> # trace out all but sites (3, 4, 5, 6, 7, 8)
>>> rho_ab = qu.partial_trace(psi, dims, keep=range(3, 
>>> rho_ab.shape
(64, 64)

>>> # compute the logarithmic negativity for the remain
>>> dims_ab = [2] * 6
>>> qu.logarithmic_negativity(rho_ab, dims_ab, sysa=(0,
0.2715533031636124

>>> # OR perform both steps at once
>>> qu.logneg_subsys(psi, dims, sysa=[3, 5, 7], sysb=[4
0.2715533031636124
```

6.1.3 Time Evolutions

There is also a convenient interface for managing time evolutions, which can be performed in a number of ways (full diagonalization, integration, or Krylov exponentia-

```
>>> # set up an initial state and hamiltonian
>>> H = qu.ham_heis(18, sparse=True)
>>> psi0 = qu.rand_product_state(18)

>>> # set and and perform an evolution
>>> evo = qu.Evolution(psi0, H, progbar=True)
>>> evo.update_to(5)
100%|###################################| 100/100 [00:04<00:00, 27.55%/s]

>>> # find the overlap between the evolved and initial state
>>> psit = evo.pt  # retrieve |ψ(t)⟩ = e^{-iHt}|ψ(0)⟩
>>> qu.expec(psi0, psit)
1.7917475317189076e-05
```

This `Evolution` class can also be used as a *generator* to yield the time-evolved state at many times in a memory-efficient manner, or functions can be supplied which will be automatically computed at each dynamic integration step.

6.1.4 Distributed eigensolving

In terms of distributed calculations, `quimb` offers two key features. The first is the ability to use `slepc` [305, 306] as an eigensolver with the same interface as `scipy`. `slepc` is a massively parallel library that makes use of the fast message passing interface (MPI) [307]. If it is installed it is straightforward to begin using its more advanced algorithms:

```
>>> A = qu.ham_hubbard_hardcore(10, sparse=True)
>>> qu.groundenergy(A, backend='scipy')
-7.015446354492036

>>> qu.groundenergy(A, backend='slepc')
-7.015446354492048
```

The second is the ability to only construct certain rows from a sequence of kronecker products. Take, for example, the case of forming an operator which acts with the Pauli-Z on two of many spins. The usual way of doing this is the `ikron` (identity

kronecker) function:

```
>>> Z = qu.pauli('Z', sparse=True)
>>> dims = [2] * 10
>>> inds = (4, 5)  # act on the 5th and 6th spins only
>>> IIIIZZIIII = qu.ikron(Z, dims, inds)
>>> IIIIZZIIII.shape
(1024, 1024)
```

As the total Hilbert space becomes much larger, in general we only want each local MPI process to construct a small selection of rows. `quimb` allows this by intelligently slicing the input operators:

```
>>> IIIIZZIIII_local = qu.ikron(Z, dims, inds, ownership=(300, 350))
>>> IIIIZZIIII_local.shape
(50, 1024)
```

In this way any operators, such as Hamiltonians, which are generally constructed as the sum of many kronecker product sequences can be simultaneously distributed and parallelized – ready for input to `slepc`. `quimb` can be launched in one of three MPI modes and generally handles all the slicing and communicating of operators automatically.

These two features combined are crucial when, for example, examining the mid-spectrum eigenstates of the MBL Hamiltonian, as in chapter 3.

6.1.5 Stochastic Lanczos Quadrature

In terms of the SLQ functionality discussed in chapter 5, this is provided by the function `approx_spectral_function`. Here we compute $\mathrm{Tr}(\sqrt{A})$ for a large, random,

```
>>> A = qu.rand_pos(2**18, sparse=True)
>>> A
<262144x262144 sparse matrix of type '<class 'numpy.complex128
  with 862615 stored elements in Compressed Sparse Row format>

>>> qu.approx_spectral_function(A, f=qu.sqrt, verbosity=1, tol
LANCZOS f(A) CALC: tol=0.001, tau=0.00025, R=1024, bsz=1
k=40: Returning estimate 185713.95980018922.
Repeat 1: estimate is 185713.95980018922
k=40: Returning estimate 185237.39196082807.
Repeat 2: estimate is 185237.39196082807
k=40: Returning estimate 185212.01899937398.
Repeat 3: estimate is 185212.01899937398
Total estimate = 185387.79025346375 +- 133.29239096525717
Repeat 3: converged to tol 0.001
ESTIMATE is 185387.79025346375 +- 133.29239096525717
185387.79025346375
```

sparse and positive matrix A:

function accepts *any* operator defined as having a linear action – not just a dense or sparse matrix.

Other features of the core module not demonstrated here include: two or more dimensional Hilbert space construction, automatic exploiting of symmetries when eigensolving, fast randomized SVD for linear operators, open system time evolutions, and numerous non-trivial pure and mixed state quantities.

6.2 The tensor network module

Since, as discussed extensively before, tensor networks have become one of the most successful approaches to efficiently simulating quantum systems, a tensor network module has been developed alongside `quimb` with the same principles - easy, interactive usage whilst retaining high performance. At its core at two generic classes, the `Tensor` and the `TensorNetwork`, which together allow a concise framework for any type of tensor network algorithm. From this point special one-dimensional algorithms in particular are developed, but the core library itself is totally agnostic to geometry.

6.2.1 Tensors

The key difference between a n-dimensional array and a tensor in `quimb` is that tensors have *named* dimensions which are propagated through any operations. On top of this, `quimb` associates an arbitrary number of 'tags' with each tensor, allow it to be uniquely identified or grouped with other tensors. In this way, shared *indices* specifying unperformed contractions completely define the geometry of a network, while the tags allow you succinctly address particular sections.

We'll begin by illustrating basic instantiation and contraction of tensors. Imagine we want to compute the dot product between two operators A and B:

```
>>> # first we import the tensor network 'namespace'
>>> import quimb.tensor as qtn

>>> # create two dummy arrays and label them as tensors
>>> A, B = qu.randn((3, 4)), qu.randn((5, 3))
>>> A
array([[-0.19225872,  0.74364341,  0.2376952 , -1.15387431],
       [-0.38995519,  1.4374562 , -0.36928995,  2.24068108],
       [-0.72356227, -0.09718222, -0.54441159,  0.87804104]])

>>> B
array([[ 1.61546384, -0.44581857, -1.05435023],
       [-0.98493191,  0.20710894, -1.28524228],
       [ 1.27386758, -0.03764016, -1.02159698],
       [-0.69690787, -1.1528959 , -0.91709318],
       [-1.2999978 ,  0.43790411, -0.63433899]])

>>> TA = qtn.Tensor(A, inds=['bond', 'left'], tags={'A'})
>>> TB = qtn.Tensor(B, inds=['right', 'bond'], tags={'B'})
```

Note that the dimension labelled `'bond'` occurs on both tensors and has the same size on each, but the raw arrays are otherwise not aligned. If we dot these together using the `@` symbol this bond is both automatically aligned and contracted:

```
>>> TA @ TB
Tensor(shape=(4, 5), inds=('left', 'right'), tags={'A', 'B'})
```

Note that the tags have been merged as well. This is the fundamental behaviour that

underpins the rest of the tensor network module.

6.2.2 Tensor networks

If instead of eagerly contracting the two tensors together we merely lazily 'combine' them with the & symbol, we get a `TensorNetwork`:

```
>>> TN_AB = TA & TB
>>> print(TN_AB)
TensorNetwork([
    Tensor(shape=(3, 4), inds=('bond', 'le
    Tensor(shape=(5, 3), inds=('right', 'b
])
```

This `TensorNetwork` can efficiently address and retrieve tensors based on their indices and tags. At this point we could just contract the whole network again using the shorthand ^ symbol:

```
>>> TN_AB ^ all
Tensor(shape=(4, 5), inds=('left', 'right'), tags={'A', 'B'})
```

produces the same result as before. We could also at this point *draw* the network:


```
>>> TN_AB.graph(color=['A', 'B'])
```

where we can see for the first time the utility of adding tags to the tensors.

6.2.3 One-dimensional tensor networks

`quimb` has various tensor network classes representing states on a one-dimensional lattice, the most important of these is probably the `MatrixProductState`. We'll start just be creating a random one:

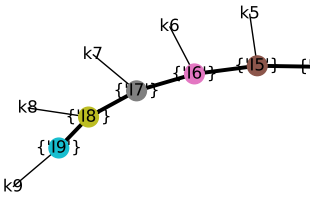
```
>>> mps = qtn.MPS_rand_state(10, bond_dim=7)
>>> print(mps)
MatrixProductState([
    Tensor(shape=(7, 2), inds=('_2fe8120000002', 'k0'), tags={'I0'}),
    Tensor(shape=(7, 7, 2), inds=('_2fe8120000002', '_2fe8120000004', '
    Tensor(shape=(7, 7, 2), inds=('_2fe8120000004', '_2fe8120000006', '
    Tensor(shape=(7, 7, 2), inds=('_2fe8120000006', '_2fe8120000008', '
    Tensor(shape=(7, 7, 2), inds=('_2fe8120000008', '_2fe812000000a', '
    Tensor(shape=(7, 7, 2), inds=('_2fe812000000a', '_2fe812000000c', '
    Tensor(shape=(7, 7, 2), inds=('_2fe812000000c', '_2fe812000000e', '
    Tensor(shape=(7, 7, 2), inds=('_2fe812000000e', '_2fe812000000A', '
    Tensor(shape=(7, 7, 2), inds=('_2fe812000000A', '_2fe812000000C', '
    Tensor(shape=(7, 2), inds=('_2fe812000000C', 'k9'), tags={'I9'}),
], structure='I{}', nsites=10)
```

things are notable here: (i) random indices have been generated for the internal bonds – their exact value is generally never needed; and (ii) the

tensors have been tagged with a 'structure' 'I0', 'I1', 'I2', ... denoting which site is which. This structure allows the tensors to be selected:

```
>>> mps[5]
Tensor(shape=(7, 7, 2), inds=('_2fe812000000a', '_2fe812000000c', 'k5'), tags={'I5'})
>>> mps[3:6]
<TensorNetwork(tensors=3, structure='I{}', nsites=10)>
```

Or we can visualize the indices and tags of the whole matrix product state:




```
>>> mps.graph([f'I{i}' for i in range(mps.nsites)])
```

Two copies of a tensor network naturally share the same external indices and thus contracting them automatically computes the frobenius norm:

```
>>> mps.H @ mps  # ⟨ψ|ψ⟩
1.0
```

nicely mirrors the syntax for dense states. When many tensors are contracted at once like this, `quimb` utilizes the library `opt_einsum` [308] to quickly find an efficient contraction path and then perform the contraction automatically. Identifying the *optimal* path is an NP-Hard problem, however in many cases a heuristic approach works just as well, and there is on-going research into improving the path quality for trickier cases [297, 301, 309]. As part of this thesis, work on `opt_einsum` has significantly improved this path-finding time and the quality of contractions found, though we skip the details of these advancements here.

To demonstrate the geometry agnostic design of `quimb` , we can also here gener-

ate and graph a MERA [295, 296] state:

```
>>> mera = qtn.MERA.rand_invar(64)
>>> mera.graph(color=[f'_LAYER{i}' for i in range(6)])
```

Tensors in this MERA are now tagged by which 'layer' of unitaries and isometries they are in, as well which sites they are in the 'light-cone' of:

```
>>> mera.select(['_LAYER0' , 'I32'], which='all')
<TensorNetwork(tensors=3, structure='I{}', nsites=64)>
```

6.2.4 DMRG

The DMRG implementations are encapsulated in the classes `DMRG1` and `DMRG2`

which are instantiated using a matrix product operator:

```
>>> # generate a matrix product operator Hamiltoni
>>> H = qtn.MPO_ham_heis(100, S=1 / 2)

>>> # make a two-site DMRG solver and solve to one
>>> dmrg = qtn.DMRG2(H)
>>> dmrg.solve(tol=1e-4, verbosity=1)
SWEEP-1, direction=R, max_bond=8, cutoff:1e-08
100%|###################################| 99/99 [0
Energy: -43.937109025586565 ... not converged.
SWEEP-2, direction=R, max_bond=16, cutoff:1e-08
100%|###################################| 99/99 [00
Energy: -44.1028154398287 ... not converged.
SWEEP-3, direction=R, max_bond=32, cutoff:1e-08
100%|###################################| 99/99 [00
Energy: -44.12190473574812 ... not converged.
SWEEP-4, direction=R, max_bond=64, cutoff:1e-08
100%|###################################| 99/99 [00
Energy: -44.127003485897276 ... not converged.
SWEEP-5, direction=R, max_bond=128, cutoff:1e-08
100%|###################################| 99/99 [00
Energy: -44.12770555415954 ... not converged.
SWEEP-6, direction=R, max_bond=256, cutoff:1e-08
100%|###################################| 99/99 [00
Energy: -44.1277352910854 ... converged!
True
```

can see for this 100 site, OBC, spin-half Heisenberg chain finding the groundstate to a reasonable degree of accuracy takes only a few seconds. The optimized state can be inspected with a schematic of bond sizes shown for the whole chain:

```
>>> gs = dmrg.state
>>> gs.show()
     2 4 8 16 19 22 24 27 32 29 36 33 38 36 40 38 40 39 42 42 44 4
    >->->->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->- ...
    | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
                              ...
    4 45 45 46 46 47 47 48 48 49 49 49 50 50 51 50 52 51 52 52 53
... ->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-> ...
     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
                              ...
    53 53 53 53 53 54 54 54 54 54 54 54 54 54 53 53 53 53 53 53 52
... -->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-- ...
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
                              ...
     52 51 51 50 50 49 49 49 47 48 47 47 46 46 44 44 42 43 39 40 3
... >-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->-->- ...
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
                              ...
    8 40 36 38 33 36 29 32 27 24 21 19 16 8 4 2
... ->-->-->-->-->-->-->-->-->-->-->-->-->->->->-o
     |  |  |  |  |  |  |  |  |  |  |  | | | | |
```

Or we can compute some basic quantities with it:

```
>>> gs.entropy(41)  # entropy of the left 41 spi
1.4112329037142548
```

non-isotropic, non-translationally invariant Hamiltonians can also easily be built using the `SpinHam` class.

`quimb` is also notable for being able to efficiently handle periodic boundary conditions using transfer matrix compression [285], which is often useful to suppress boundary effects. It can also perform DMRG-X [310, 179] to find highly excited states of matrix product operators with area law eigenstates throughout their spectrum.

6.2.5 TEBD

In order to enable the time evolution of matrix product states `quimb` implements time-evolving-block decimation, encapsulated in the class `TEBD`. Unlike DMRG, `TEBD` must be instantiated with a nearest neighbour interaction hamiltonian – a `NNI`:

```
>>> # use a MBL hamiltonian near the critical point
>>> H = qtn.NNI_ham_mbl(100, dh=5.0)

>>> # start with a computational state
>>> psi0 = qtn.MPS_rand_computational_state(100)

>>> # instantiate the TEBD and evolve
>>> tebd = qtn.TEBD(psi0, H)
>>> tebd.update_to(3, tol=1e-3)
t=3, max-bond=21: 100%|##################| 100/100 [00:08<00:00,  9.46%/s]
```

We can then investigate how much entanglement has been generated in the time-evolved state:

```
>>> psit = tebd.pt
>>> psit.show()
     2 4 7 8 10 10 8 8 8 8 8 7 9 8 11 9 9 10 13 10 9 6 7 5 6 5 4 4
    >->->->->->-->-->->->->->->->->->->->->->-->->->-->-->->->->->->->->->->- ...
    | | | | |  |  | | | | | | | | |  | | |  |  |  | | | | | | | | |
                                ...
     7 11 14 16 8 9 8 7 8 7 6 6 6 7 9 10 8 10 12 11 12 11 11 12 8
... >->-->-->-->->->->->->->->->->->->->->->->->->->->-->-->-->-->-->-->-->-> ...
    | |  |  |  | | | | | | | | | | | | | | |  | |  |  |  |  |  |  | | |
                                ...
    11 8 11 9 8 8 8 7 8 10 11 10 13 12 10 11 11 10 10 9 9 8 10 10
... -->->-->->->->->->->->-->-->-->-->-->-->-->-->-->-->->->->->->->-->--> ...
      | |  | | | | | | | |  |  |  |  |  |  |  |  |  |  | | | |  |  |
                                ...
    12 11 13 18 18 21 15 14 16 12 11 10 9 10 9 10 10 11 12 8 4 2
... -->-->-->-->-->-->-->-->-->-->-->-->-->->->->->->->->->->->->->->->->->->->-o
      |  |  |  |  |  |  |  |  |  |  |  | |  | |  |  |  |  | | | |
```

The same machinery that allows TEBD can also be used to simulate full quantum circuits, whether in MPS or generic tensor network form. While this allows the simulation of, for example, 49 qubits to a gate depth of 20 or more (in terms of single amplitudes), we do not show this here.

6.2.6 TNSLQ

In order to demonstrate the TNSLQ method, we show here how to cast a tensor network describing the subsystem of a periodic MPS as a linear operator. From that point

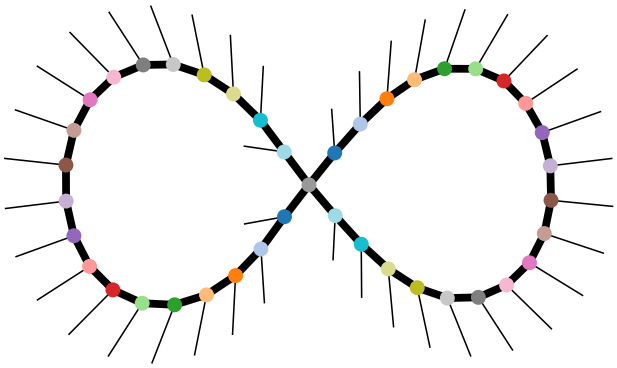
the SLQ method can be used to estimate spectral quantities such as its entropy easily.

```
>>> psi = qtn.MPS_rand_state(60, bond_dim=16, cyclic=True)

>>> # for the density operator by only 'untracing' the first 20 sites
>>> rho_a = psi.H & psi.reindex_sites('b{}', range(0, 20))

>>> # perform an inplace contraction of the remaing, traced sites
>>> rho_a ^= slice(20, 60)
>>> rho_a
<TensorNetwork(tensors=41, structure='I{}', nsites=60)>
```

At this point our tensor network density operator looks like so:



with the central rank-4 tensor containing all information regarding the effective environment. In this picture, the operator we want to define maps states spanning all of the left hand side indices to states spanning the all of the right hand side indices.

```
>>> # define 'left' and 'right' hand side indices of a linear operator
>>> left_inds = [f'k{i}' for i in range(20)]
>>> right_inds = [f'b{i}' for i in range(20)]
>>> rho_a_lo = rho_a.aslinearoperator(left_inds, right_inds)
>>> rho_a_lo
<1048576x1048576 TNLinearOperator with dtype=float64>

>>> # now use SLQ procedure to estimate the entropy: -Tr(ρ_A log2 ρ_A)
>>> S_a = - qu.approx_spectral_function(rho_a_lo, f=qu.xlogx, tol=0.05)
>>> S_a
5.787216528009696
```

Which concludes our demonstration of the capabilities of `quimb` and it tensor network module, though there are many more features not shown. One notable such

feature is the ability to back the tensors with many types of array other than those of `numpy`, for example, 'out-of-core' arrays or GPU arrays. In particular, the use of machine learning frameworks such as `pytorch` [311] mean that tensor-network derived scalar quantities can be auto-differentiated with respect to every constituent tensor. This allows the use of modern stochastic gradient descent algorithms to effectively *globally* optimize tensor-network states, avoiding reliance on hand-crafted optimization schemes such as DMRG (which does however likely perform better). Continuing to develop `quimb` in a open manner that interfaces with other such libraries in a modular way should hopefully bring many further rewards into the field of classical simulation of quantum many-body systems.

Chapter 7

General Conclusions

Quantum simulation of many-body quantum systems is a very promising prospect for quantum technology – both in the near future and indeed into the far future. Many important fields such as those involved with the development of advanced materials and chemicals are significantly constrained by the essential impossibility of fully simulating large quantum systems classically. In parallel, entanglement is a key concept for dealing with such many-body systems. This is not only as an essential 'ingredient' for quantum simulations, but as a perspective for understanding a huge variety of phenomena. The two topics together span a very wide range of fields, which lends itself to a profitable interplay of ideas. In this thesis we have made advancements pertinent across several of these fields, including quantum dot arrays, the many-body localization (MBL) transition, and the practical quantum measurement and classical computation of many-body entanglement. To conclude, we briefly recap the the achievements of each chapter and offer some perspective on possible limitations and extensions.

In Chapter 2, 'Unravelling Quantum Dot Array Simulators', we examined the near-term prospects for quantum dot arrays as simulators given both the limited availability of measurements, and the noisy environment. Developing solid state simulators such as quantum dot arrays is potentially important due to fact that their fundamental components – hopping electrons – naturally correspond to the most important many-body lattice models, namely, the Fermi-Hubbard model and its derivatives. On the other hand, quantum dot arrays are quite a fledgling technology

in comparison to cold atom optical lattice set-ups, for example, and so there is a need to quickly establish tools to verify their operation. We found that singlet-triplet measurements, available through the parallel capacitance of the quantum dot system, were sufficient to identify the groundstate of the expected Heisenberg model with high certainty. Moreover, due to the symmetry of the underlying model, we showed that a partial singlet-triplet measurement was sufficient to localize entanglement to distant ends of a chain, in a heralded fashion. Additionally, we showed that the effective noise model governing a quantum dot system displays a transition to an MBL phase, opening up the possibility of using quantum dot arrays as direct simulators of non-thermalizing models. A limitation to our singlet-triplet approach was the restriction to a single output parameter – the number of triplets found – and the discrimination of states based on the classical probability distribution this parameter produced. This was in part due to the fact that the singlet-triplet measurement is not a complete measurement and cannot thus be used for traditional tomography. Recent advances make use of machine learning techniques [239] to train an efficient state representation that recreates a measured probability distribution. This approach would seem a natural approach to explore for singlet-triplet measurements in quantum dot arrays in the future.

Picking up the thread of MBL, in Chapter 3 we study the transition between ergodic and non-thermalizing phases using two entanglement related quantities. The highly unusual nature of MBL and its full-spectrum transition make it highly desirable to develop a better understanding of it. We advance this area in two ways. Firstly, by performing a rigorous finite size scaling analysis of the Schmidt gap across the transition we find a critical exponent in line with theoretical predictions – unlike all previous studies. Moreover, we rationalize this result by investigating the scaling of fluctuations across samples with system size, from which we conclude the Schmidt gap is far less sensitive to certain finite size effects. To match the scaling behaviour we also probe an entanglement length, and find strong evidence that it diverges at the critical point. An obvious limitation of this research is the restriction to fairly short lengths in the context of usual phase transition scaling studies. However, part

of what makes MBL so interesting is the precise fact there is no established and effective model of the transition point – meaning that exact diagonalisation must be used. Moving forward, an interesting avenue of research is whether enough can be revealed about the entanglement structure at the MBL transition to inform a tensor network approach beyond matrix product states.

In Chapter 4 we move onto the question of actually measuring many-body entanglement in a quantum simulator. Although schemes exist for measuring the entanglement entropy (which is only for pure bipartitions), or for tomographically reconstructing lowly entangled, many-body states, there has not been a way so far to measure many-body entanglement in generic, many-body mixed states. Being able to perform such measurements would give us a much more versatile probe of entanglement in quantum simulators. We show that this is possible using a combination of choice measurements on three or more copies of a state along with a neural network estimator with which to extract the logarithmic negativity. Notably, the method does not require tomographic reconstruction of the state, and is also particularly effective for *highly* entangled states. An essential aspect of this work was that half of the sample states used to train the network were generated with a controlled amount of entanglement by utilising a random MPS ansatz. The neural network otherwise received no 'physical' data and yet was able to predict the many-body entanglement in physical models. A limitation of this method however is the need to classically train the neural network, with it not being clear how well the neural network could extrapolate its knowledge beyond the system sizes it has already seen – though feature scaling may well alleviate this in the future.

Motivated in part by the need to compute the many-body entanglement in the two earlier chapters, in Chapter 5, we focus on developing efficient ways of computing the logarithmic negativity in generic many-body states as well as matrix product states. The first ingredient is a lazy representation of the partially traced, partially transposed density operator as an uncontracted tensor network. The second is the stochastic Lanczos quadrature procedure, which the allows the estimation of spectral quantities of linear operators *without* assuming said operator is low-rank.

The combination of these two features – dubbed TNSLQ – allows the tractable computation of logarithmic negativity in many-body states roughly double the size (Hilbert space dimension *squared*) of what was previously possible. TNSLQ is a very general method that indeed can compute a very wide variety of quantities alongside the logarithmic negativity and we expect this to be a fruitful line of research in the future.

Finally, in Chapter 6, we have introduced `quimb` [5], a open source python library for quantum information and many-body calculations including general purpose tensor network support. `quimb` contains implementations of many of the ideas discussed in the thesis so far including the TNSLQ algorithm. More than that however, `quimb` has been designed with modern scientific software best-practices in mind, including extensive documentation, unit testing, and usage of other high quality open-source libraries as its foundation. In part due to these efforts, `quimb` has already seen moderate uptake and received minor contributions. It has also been designed with interactivity and performance in mind, allowing one to use it easily as a conceptual aid but then also scale calculations to large distributed systems. For all these reasons we hope and expect that it will be a useful tool for rapid, research orientated quantum calculations in the future.

Appendix A

Appendix for 'Entanglement & the many-body localization transition'

A.1 Quality of Collapse

To find the quality of data collapse of scaled Schmidt data for a given h_c and ν we use the `pyfssa` program [312]. The underlying procedure, which is based on the method as refined in [203] is as follows. Assume we have data points y_{ij} (e.g the Schmidt gap, Δ) and their standard errors dy_{ij}, where i indexes the lengths L_i and j the disorder strengths h_j. Since we assume that there is only a correlation length exponent, ν, we scale only the disorder strength as such: $x_{ij} = L_i^{1/\nu}(h_j - h_c)$. A curve is then fitted through the data using least squares which yields the fitted points Y_{ij} and their estimated errors dY_{ij}. The quality can then be defined as a χ^2 statistic based on the relative deviation from this fitted curve:

$$Q = \frac{1}{\mathcal{N}} \sum_{i,j} \frac{(y_{ij} - Y_{ij})^2}{dy_{ij}^2 + dY_{ij}^2} \tag{A.1}$$

with normalization $\mathcal{N}$ accounting for the number of terms where the fitted curve is defined. This quantity is minimized when all the actual data lies close to the curve of best fit. In particular, when the deviations from the fit are approximately equal to the uncertainty in the fit, $Q \sim 1$, which is generally what we find with the Schmidt gap Δ.

Since finite size scaling is only strictly relevant close to the transition, and finite

size effects may be too strong at very small lengths, there is also some freedom in how one selects the data to scale. Similarly to [79], we vary the size of the window around h_c to select, the minimum L of which to include data for, and also whether to include odd as well as even L. The advantage of using Eq. (A.1) is that the the combinations of the above which give statistically reasonable collapse can be objectively identified. The minimum L is chosen such that increasing it does not significantly change the values of ν and h_c found. One can then average over all parametrizations that achieve a 'good' value of Q (determined visually to be ~ 10) – yielding the plot in Fig. 3(b).

The different parametrizations above, as well as bootstrap sampling over disorder realizations, then yields a spread in the locations of h_c and ν where the best value of Q is found. This allows an estimation of the error in h_c and ν which takes into account both the collapse method and the random error.

A.2 Entanglement Length with Larger Block Size

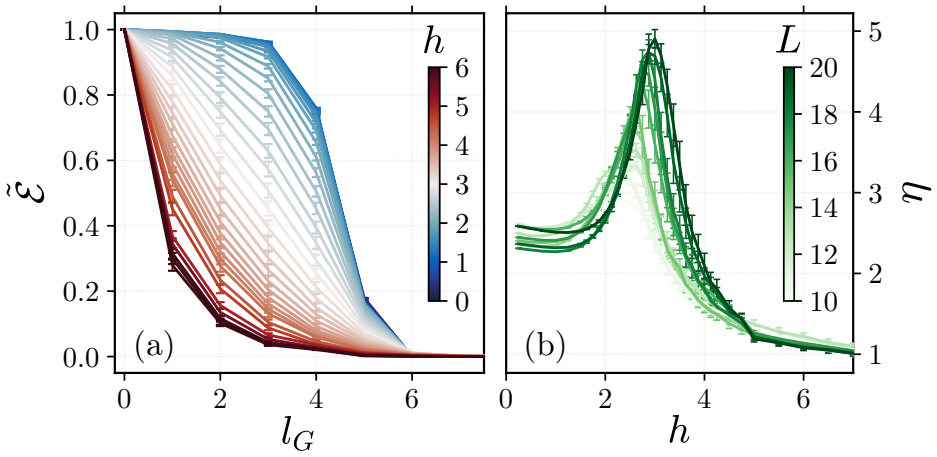


Figure A.1: (a) Average logarithmic negativity as a function of the gap between two disjoint blocks, the first block being two spins, the second the rest of the system, for $L = 20$, which has a pseudo-critical point at $h \sim 3$. (b) Bipartite entanglement length, as computed with Eq. (3), across the MBLT for varying chain length L. Error bars shown where visible.

For completeness, we show here the equivalent of Fig. 5 taking instead the size of the block A to be 2. As can be seen in Fig. A.1, this yields almost exactly the

same shapes and plots, including a divergence of the length ν at the pseudo-critical points $\tilde{h_c}$, which shift right with length L.

A.3 Detailed Behaviour of the Disjoint Entanglement vs. L

Finally, in Fig. A.2, we show in detail the behaviour of the normalized disjoint entanglement as a function of L, which sheds some light on the physical picture of the entanglement length. In the ergodic phase – left column, $h = 0.5$ – one can see that the point of decay for $\tilde{\mathcal{E}}$ shifts linearly to the right with system size L, but the length scale of the decay (which is what η captures) remains constant. In fact, the curves collapse onto each other with a shift of $-L/4$ (not shown). In the localized phase – right column, $h = 6$ – the decay of entanglement, and thus η, is practically identical for all the lengths. At the transition point however – central column, h taken as pseudo-critical points $\tilde{h}_c$ – the scale of decay stretches with system size. Lastly, we note that the difference between the two block sizes tested (upper row 1, lower row 2) is very minimal.

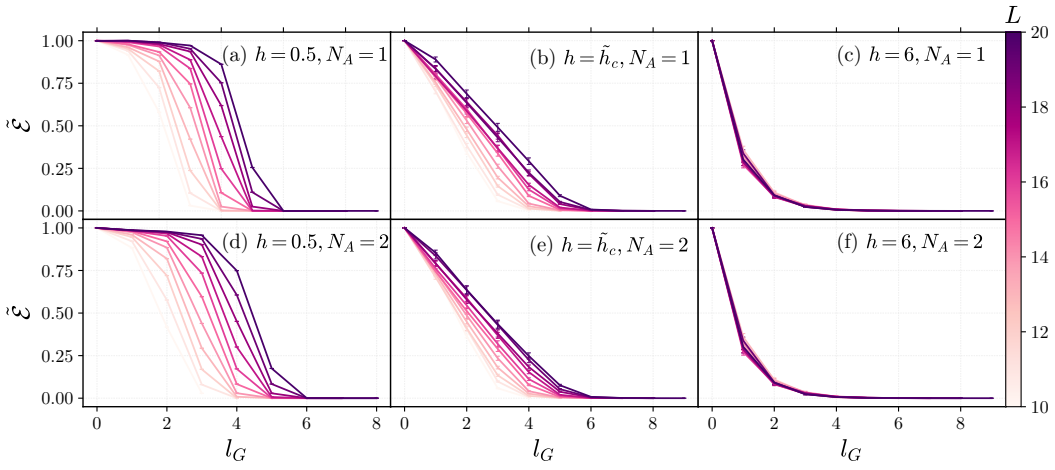


Figure A.2: Behaviour of the normalized disjoin block entanglement, $\tilde{\mathcal{E}}$, as a function of the gap size l_G for varying chain lengths L. The top line of panels corresponds to an initial blocksize of $N_A = 1$, and the bottom 2. The left column is shows behaviour deep in the ergodic phase, the middle behaviour at the pseudo-critical points, and the right behaviour deep in the localized phase. Error bars shown where visible.

Appendix B

Appendix for 'Machine-learning-assisted many-body entanglement measurement'

B.1 Measuring Moments in spin-$1/2$ Systems

Here we show that the operational procedure described in the main text allows us to measure the moments μ_m for spin-1/2 systems. Let $S_X^{c,d}$ be the operator that swaps copies c and d on subsystem X, which can be written as $S_X^{c,d} = \prod_{j\in X} \Xi_j^{c,d}$ where $\Xi_j^{c,d} = (\mathbb{1} + \boldsymbol{\sigma}^{j,c}\cdot\boldsymbol{\sigma}^{j,d})/2$. The projective measurement of $\Xi_j^{c,d}$ corresponds to a singlet triplet measurement (ST-measurement) between spins sitting at the same site j, but different copies c and d. Indeed, $\Xi_j^{c,d}$ has an outcome -1 for the singlet state and 1 for the triplet states. In view of this we write $S_X^{c,d}=\Pi_+^{c,d}-\Pi_-^{c,d}$ where $\Pi_\pm^{c,d}$ correspond to the eigen-projections with corresponding eigenvalues ± 1.

We now consider the case m=3 and then generalise to arbitrary m. We first perform a sequential set of ST-measurements on copies $(1,2)$, with outcome β_1 and then do the same measurement on copies $(2,3)$, with outcome β_2. We introduce the notation $S_X^{2,3}\circ S_X^{1,2}$ to describe this process. After the first measurement, the (non-normalized) state of the system will be $\Pi_{\beta_1}^{1,2}\rho^{\otimes 3}\Pi_{\beta_1}^{1,2}$, while after the two sets of

measurements it is $\Pi^{2,3}_{\beta_2}\Pi^{1,2}_{\beta_1}\rho^{\otimes 3}\Pi^{1,2}_{\beta_1}\Pi^{2,3}_{\beta_2}$. Therefore,

$$\begin{aligned}\left\langle S_X^{2,3}\circ S_X^{1,2}\right\rangle &= \sum_{\beta_2}\sum_{\beta_1}\beta_1\beta_2\mathrm{Tr}\left[\Pi^{2,3}_{\beta_2}\Pi^{1,2}_{\beta_1}\rho^{\otimes 3}\Pi^{1,2}_{\beta_1}\Pi^{2,3}_{\beta_2}\right]\\ &= \sum_{\beta_1}\beta_1\mathrm{Tr}\left[\Pi^{1,2}_{\beta_1}S_X^{2,3}\Pi^{1,2}_{\beta_1}\rho^{\otimes 3}\right]\\ &= \frac{1}{2}\left(\mathrm{Tr}[S_X^{1,2}S_X^{2,3}\rho^{\otimes 3}]+\mathrm{Tr}[S_X^{2,3}S_X^{1,2}\rho^{\otimes 3}]\right)\\ &= \mathrm{Tr}(\rho^3), \end{aligned} \tag{B.1}$$

where we used the identity $\Pi^{c,d}_{\pm}=(\mathbb{1}\pm S_X^{c,d})/2$.

We now generalize the above argument for higher values of m. We apply sequential ST-measurements on neighbouring copies, using the notation $S_X^{m,m-1}\circ\cdots\circ S_X^{2,3}\circ S_X^{1,2}$, meaning that we first perform $S_X^{1,2}$ and so forth. Taking the averages one then finds that

$$\left\langle S_X^{m-1,m}\circ\cdots\circ S_X^{1,2}\right\rangle = \mathrm{Tr}[\mathcal{P}^{m-1,m}[\cdots\mathcal{P}^{23}[\mathcal{P}^{12}[\rho^{\otimes m}]]]\cdots],$$

where $\mathcal{P}^{j,j+1}[\rho]=\sum_{\beta_j}\beta_j\Pi^{j,j+1}_{\beta_j}\rho\Pi^{j,j+1}_{\beta_j}$. We define the operators $S_X^{a,b,c,\cdots}$ recursively as: $\mathcal{P}^{j+1,j}[S_X^{j,a,b,\cdots}]=[S_X^{j+1,j,a,\cdots}+S_X^{j,j+1,a,\cdots}]/2$. Then, using the cyclic property of the trace, one finds that $\left\langle S_X^{m-1,m}\circ\cdots\circ S_A^{1,2}\right\rangle = 2^{2-m}\sum_\kappa\left\langle S_X^\kappa\right\rangle$ where the κ are 2^{m-2} different cyclic permutation of the elements $1,\ldots,m$. For instance, for $m=3$ one has $\kappa=\{123,132\}$. In view of the above, this sequential set of ST-measurements corresponds to the measurement of the operator $\mathbb{P}_X^m = 2^{2-m}\sum_\kappa S_X^\kappa$.

In summary, when X is equal to both A and B, one obtains the operator, $\mathbb{P}^m$ defined in Eq. (B.2). As proven above, such an operator is defined according to the following recursion relations [218]: $\mathbb{P}^m = \left(S_A^{m,m-1}S_B^{m,m-1}\mathbb{P}^{m-1}+\mathrm{h.c.}\right)/2$. Moreover, the effect of the partial transpose on the recursion relation is as follows: $(\mathbb{P}^m)^{T_B} = \left(S_A^{m,m-1}(\mathbb{P}^{m-1})^{T_B}S_B^{m,m-1}+\mathrm{h.c.}\right)/2$. We take $m=3$ as an example:

$$(\mathbb{P}^3)^{T_B} = \left(S_A^{3,2}S_A^{2,1}\otimes S_B^{2,1}S_B^{3,2}+\mathrm{h.c.}\right)/2. \tag{B.2}$$

From which it can be seen, as described in the main text, that the order of measure-

ments on A and B is reversed. Indeed, for A, the ST-measurement is performed between copies 1 and 2, then 2 and 3, whereas for B, the ST-measurements is performed for copies 2 and 3, then 1 and 2. Because of the non-commutative nature of these measurements, this ordering is crucial in order to yield moments of the partially transposed state as in Eq. (B.2).

From the above derivation we find that $\mu_m = \mathrm{Tr}\left[\rho_{AB}^{\otimes m}\mathbb{P}_m\right]$. The variance of the above measurement can be found using the procedure of (B.1), noting that $\mathrm{Tr}\left[\Pi_{\beta_1}^{1,2}S_X^{2,3}\Pi_{\beta_1}^{1,2}\rho^{\otimes 3}\right]$ is the provability of getting the outcome sequence β_1, β_2. Therefore the variance is

$$\begin{aligned}(\Delta\mu_m)^2 &= \left(\sum_{\beta_m}\cdots\sum_{\beta_1}\beta_1^2\cdots\beta_m^2\,\mathrm{Tr}\left[\Pi_{\beta_{m-1}}^{m-1,m}\cdots\Pi_{\beta_1}^{1,2}\rho^{\otimes m}\Pi_{\beta_1}^{1,2}\cdots\Pi_{\beta_{m-1}}^{m-1,m}\right]\right)-\mu_m^2 = \\ &= \left(\sum_{\beta_m}\cdots\sum_{\beta_1}\mathrm{Tr}\left[\Pi_{\beta_{m-1}}^{m-1,m}\cdots\Pi_{\beta_1}^{1,2}\rho^{\otimes m}\Pi_{\beta_1}^{1,2}\cdots\Pi_{\beta_{m-1}}^{m-1,m}\right]\right)-\mu_m^2 = \\ &= 1-\mu_m^2,\end{aligned} \tag{B.3}$$

where we used the fact that $\beta_i^2 = 1$ and $\sum_{\beta_i}(\Pi_{\beta_i}^{i,i+1})^2 = \sum_{\beta_i}\Pi_{\beta_i}^{i,i+1} = \mathbb{1}$. Repeating the experiment R times we find the standard deviation

$$\Delta\mu_m = \sqrt{\frac{1-\mu_m^2}{R}}\,. \tag{B.4}$$

B.2 Measuring Moments in Bosonic Systems

We show here how to measure the moments as given in Eq. 4.2 for a bosonic system. Unlike for the case of spin systems, we directly choose the operator $\mathbb{P}_m$ as a product of specific, non-hermitian, permutations, π, such that

$$\mathbb{P}_m^{T_B} = \bigotimes_{j\in A} V_{j,\pi}\bigotimes_{j\in B}V_{j,\pi}^T \tag{B.5}$$

where $V_{j,\pi} = \sum_{\{n_{j,c}\}} |n_{j,1}, \ldots, n_{j,m}\rangle \langle \pi(n_{j,1}), \ldots, \pi(n_{j,m})|$, and $n_{j,c} = 0, \ldots, \infty$ labels the number of bosons in copy c and physical site j. We can write this operator in second quantized form as

$$V_{j,\pi} =: e^{\sum_c a^\dagger_{j,c} a_{j,\pi(c)} - a^\dagger_{j,c} a_{j,c}} : \tag{B.6}$$

where $: O :$ denotes the normal ordering of the operator O, and $a_{j,c}$ denotes the annihilation operator acting on site j and copy c. We choose π as the shift permutation such that $\pi(c) = c + 1$. Note that $V^T_{j,\pi} = V_{j,\pi^{-1}}$. In order to diagonalise $V_{j,\pi}$ we introduce the Fourier transform, which acts independently on each site j as

$$\begin{aligned} \tilde{a}_{j,c} &= \frac{1}{\sqrt{m}} \sum_{c'=0}^{m-1} e^{+\frac{i2\pi}{m} cc'} a_{j,c'}, && \text{for } j \in A, \\ \tilde{a}_{j,c} &= \frac{1}{\sqrt{m}} \sum_{c'=0}^{m-1} e^{-\frac{i2\pi}{m} cc'} a_{j,c'}, && \text{for } j \in B. \end{aligned} \tag{B.7}$$

After such a transformation, both the operators $V_{j,\pi}$ for $j \in A$ and $V_{j,\pi^{-1}}$ for $j \in B$, take the form $:e^{\sum_c (e^{\frac{i2\pi}{m} c} - 1) \tilde{a}^\dagger_{j,c} \tilde{a}_{j,c}}:$. The normal ordering can be removed by using the identity [313] $:e^{(e^\lambda - 1) a^\dagger a}: = e^{\lambda a^\dagger a}$ bringing Eq. (B.5) to the form:

$$\mathbb{P}^{T_B}_m = \prod_{j \in \{A,B\}, c} e^{\frac{i2\pi c}{m} \tilde{a}^\dagger_{j,c} \tilde{a}_{j,c}}. \tag{B.8}$$

The expectation value in Eq. (4.2) can thus be measured in three steps. First, perform the Fourier (inverse Fourier) transform between copies at the sites belonging to A (B), as written in Eq. (B.7). Second, measure the bosonic occupation number with outcome $n_{j,c}$ at every site and compute the outcome of the permutation operator as $\phi = e^{\sum_{j \in \{A,B\}, c} \frac{i2\pi c}{m} n_{j,c}}$. Finally, compute the expectation value as the average over many repeatations of the above steps: $\mu_m = \langle \phi \rangle$.

The standard deviation can obtained from the fact that $\phi^* = \phi$. As such $(\Delta \mu_m)^2 = \langle \phi \phi^* \rangle - \langle \phi \rangle^2 = 1 - \mu_m^2$ and after R repetitions we find the same result of Eq. (B.4).

B.3 Experimental Feasibility.

Our proposal for measuring logarithmic negativity can be realised in different physical set-ups. In solid state quantum technologies the possibility of obtaining multiple copies is rapidly emerging. For example, in [124] a system of two parallel quantum dot arrays, each with 7 sites, was realised in GaAs. Alternatively, in silicon quantum technologies, recent proposals have put forward schemes for large two-dimensional arrays of quantum dots exploiting the well-established CMOS technology [314]. Preparing multiple copies of strongly correlated many-body systems in such structures would be a natural capability. In these quantum dot arrays, singlet-triplet measurements are already well developed [127], using either charge detection through quantum point contacts, or capacitance measurement through radio-frequency reflectometry. Since the singlet and triplet states correspond to the anti-symmetric and symmetric subspaces respectively, they are the eigenstates of the swap operator and thus the singlet-triplet measurement outcome can be mapped to the swap measurement outcome . The full details of the above, and its generalization to multiple consecutive swaps, are available elsewhere in Appendix B.

In optical lattices the situation is currently even more advanced, and multiple copies of many-body systems can be easily isolated from each other in adjacent lattice rows. However, the counterpart to the solid state swap measurement in such systems is destructive. This means that consecutive swaps cannot be measured, limiting the number of copies to only two. Nonetheless, as explained above, our procedure is valid for whenever forward-backward measurements that are equivalent to permutations are available. In the case of optical lattices, the Fourier transform provides such a capability that has also been experimentally implemented [220]. This is performed by tilting alternating rows of the lattice such that the bosons undergo a series of effective beam-splitters [216]. In our case, this operation would need to be performed on two subsystems in forwards and backwards fashion – the formal equivalence of this to two permutations is also detailed elsewhere in Appendix B. Combining these above experimental techniques for measuring the partially transposed moments, with our machine learning method for extracting

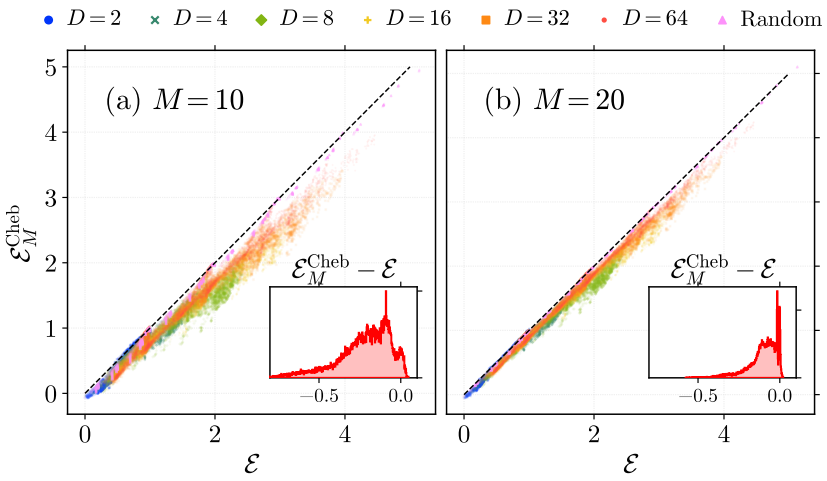


Figure B.1: Estimated logarithmic negativity $\mathcal{E}_M^{\mathrm{Cheb}}$, using the Chebyshev approximation vs. actual logarithmic negativity $\mathcal{E}$, for a wide range of random states and partitions. We use both R-GPS and R-MPS states with varying bond-dimension D, and various system sizes N_A, N_B and N_C, such that $N_A + N_B \leq 12$ and the total length $N \leq 24$. The Chebyshev approximation is calculated using the moments μ_m generated from: (a) $M = 10$ copies; (b) $M = 20$ copies. The respective insets show the distribution of error, $\mathcal{E}_M^{\mathrm{Cheb}} - \mathcal{E}$.

the logarithmic negativity, provides a realistic and complete scheme for accurately estimating entanglement in general many-body systems.

One potential source of error experimentally is that the copies may not be perfectly identical. However, it can be shown that small deviations in the fidelity of copies leads to only small changes in the moments. Since both the Chebyshev and neural network approaches yield entanglement estimators that are smooth functions of the moments, the overall error is therefore well controlled.

B.4 Chebyshev Expansion

In this section, we demonstrate an analytical method, based on functional approximation, for estimating the logarithmic negativity from the information contained in the moments, μ_m.

Since logarithmic negativity, $\mathcal{E}$, is a function of the eigenvalues $\{\lambda_k\}$, one could try to directly reconstruct the main features of the spectrum $\{\lambda_k\}$ using a few measured moments $\{\mu_m\}$ – an approach closely related to general Hausdorff moments problem in statistics [315]. Nonetheless, it is known from a numerical

perspective that the Hausdorff problem is unstable [250]. To avoid such instabilities, we might try an alternative approach based on functional approximation. Considering that $\mathcal{E} = \log_2 \mathrm{Tr} f(\rho_{AB}^{T_B})$ with $f(x) = |x|$, if we can find a polynomial expansion $f(x) \approx \sum_{m=0}^{M} \alpha_m x^m$, then by linearity of the trace, $\mathcal{E} = \log_2 \sum_{m=0}^{M} \alpha_m \mu_m$, with μ_m as given in Eq. (4.2). In other words, given a polynomial expansion of the absolute function, $f(x)$, – i.e. the coefficients α_m – one can approximate the entanglement using a finite number of moments. A naive choice for this would be a Taylor expansion, but the non-analyticity of $f(x)$ at $x = 0$ prevents convergence. On the other hand, a nearly optimal choice for approximating a function throughout an interval rather than around a point, is a Chebyshev expansion [276]. On the interval $[-1, 1]$, this yields $f(x) \approx \sum_{m=0}^{M} t_m T_m(x)$ where the $M + 1$ Chebyshev polynomials $T_m(x)$ are known m-th order polynomials. The coefficients t_m are given, via the orthogonality of $T_m(x)$, as $t_m = \frac{2-\delta_{m0}}{n+1} \sum_{j=0}^{M} f(x_j) T_m(x_j)$, where $x_j = \cos[\pi(j+1/2)/(M+1)]$ are the Chebyshev nodes. When f is defined on a different interval, one can simply linearly transform the Chebyshev points and polynomials. Although in principle the spectrum of a generic state $\rho_{AB}^{T_B}$ lies between $-1/2$ and 1, in practice it is often much more tightly clustered. Decreasing the window size significantly improves the approximation for a fixed M. Therefore, we need to find the minimal sized window such that all of the spectrum of a given $\rho_{AB}^{T_B}$ is contained. A tight guess for such a window can be found since $\mu_m = \sum_k \lambda^m \geq \lambda_{\max}^m$ when m is even, with $\lambda_{\max}$ being the eigenvalue with largest absolute value. Thus in our numerics, we define the window as $[-a, a]$, with $a = \mu_M^{1/M}$. The quality of the Chebyshev approximation rapidly increases with M and becomes exact in the limit $M \to \infty$.

In Fig. B.1(a) we plot the relationship between the real logarithmic negativity, $\mathcal{E}$, and the approximated value, $\mathcal{E}_M^{\mathrm{Cheb}}$, calculated using the Chebyshev expansion for $M = 10$ copies. The random states tested are the same as those considered in the main numerical results text. As the figure shows, $\mathcal{E}_M^{\mathrm{Cheb}}$ typically underestimates $\mathcal{E}$, especially for MPS states. As we discuss in the next section, we attribute this to the fact that the distribution of λ_k is peaked around zero, particularly for MPS,

where the Chebyshev approximation error is always negative. In the inset we plot a histogram of the errors $\mathcal{E}_M^{\rm Cheb} - \mathcal{E}$ which clearly shows this negative bias. As shown in Fig. B.1(b) and its inset, by doubling the number of copies to $M = 20$, the accuracy of this method is significantly improved. While $M = 20$ is a significant experimental requirement, we emphasize that the Chebyshev approach provides an analytical tool for estimating $\mathcal{E}$ from a known expression of the moments. In particular, when the moments can be expressed in compact form, such as for free fermionic systems [316, 194], then the Chebyshev expansion can provide an analytic formula for $\mathcal{E}$.

B.5 Generation of random states

Random generic pure states (R-GPS) have been obtained by generating random vectors with complex elements distributed according to the normal distribution. R-GPS obtained by sampling from the Haar measure have also been considered, though they are numerically more demanding. Nonetheless, they provide the same results. Random Matrix Product States, R-MPS, have been obtained by writing $|\psi\rangle = \sum_{\{i_j\}} \mathrm{Tr}[A^{(1),i_1} A^{(2),i_2} \dots A^{(N),i_N}] \, |i_1, i_2, \dots, i_N\rangle$ for random tensors $A_{kl}^{(j),i}$, where $j = 1, \dots, N$, $i = 0, 1$, and $k, l = 1, \dots, D$, being D the bond dimension, with complex elements drawn from a normal distribution.

B.6 Chebyshev Approximation Error Analysis

In this section we study in detail where the error in the Chebyshev approximation stems from and how it is related to the spectral properties of $\rho_{AB}^{T_B}$. It is convenient to consider the thermodynamic limit ($N_A, N_B \gg 1$), where one can describe the spectrum of $\rho_{AB}^{T_B}$ as a probability density function, $\omega(\lambda)$, and write the logarithmic negativity of Eq. (4.1) as

$$\mathcal{E} = \log_2 \int |\lambda| \omega(\lambda) d\lambda. \tag{B.9}$$

Therefore, this underlying function $\omega(\lambda)$ determines $\mathcal{E}$, though it is not directly accessible from measurements of the moments μ_m. Nonetheless, from a theoretical

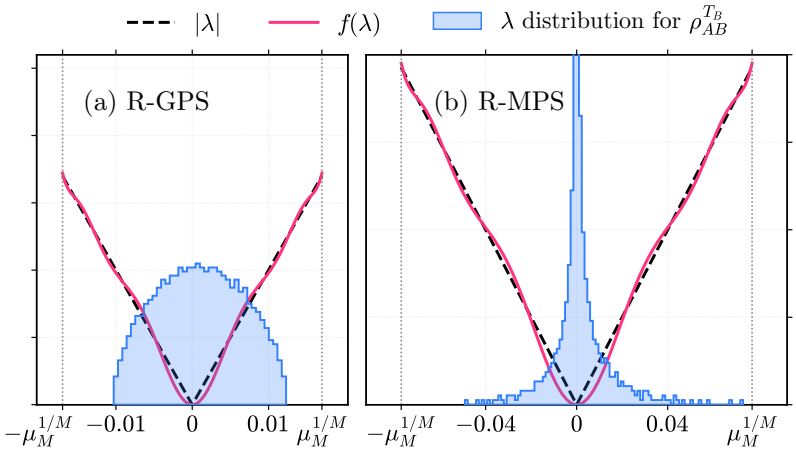


Figure B.2: Accuracy of the Chebyshev approximation. Distribution of the eigenvalues $\{\lambda_k\}$ of $\rho_{AB}^{T_B}$ for (a) R-GPS and (b) R-MPS with $D = 32$ with $N_A = 5$, $N_B = 5$ and $N_C = 5$. The dashed lines show $|\lambda|$ while the solid lines show its Chebyshev approximation for $M = 10$. Vertical dotted lines denote the estimated bounds on the spectrum calculated as $\pm\mu_M^{1/M}$.

perspective, the study of $\omega(\lambda)$ for certain classes of states yields insight into the performance of the Chebyshev approximation and machine learning approaches.

For R-GPS, $|\Psi_{ABC}\rangle$, as proven in [317, 318], the spectral distribution $\omega(\lambda)$ tends towards a shifted Wigner semi-circle law in the limit of large N_A, N_B and N_C. This is described by the continuous distribution $\omega_{\text{SC}}(\lambda) = \frac{d^2}{2\pi\sigma^2}\sqrt{4\sigma^2 - (d\lambda - 1)^2}$ where $d = 2^{N_A+N_B}$ and $\sigma^2 = 2^{N_A+N_B-N_C}$, and $|d\lambda - 1| < 2\sigma$. An instance of the spectral distribution $\omega(\lambda)$ is shown in Fig. B.2(a) for finite values of N_A, N_B and N_C, where one can already see a clear semi-circular shape. As is also evident from the figure, the support of the distribution is far smaller than the theoretical interval $\lambda_k \in [-1/2, 1]$, yet the bound established above as $|\lambda_k| \leq \mu_M^{1/M}$ quite tightly captures the real interval.

On the other hand, random states constructed using a MPS ansatz [93] with fixed bond dimension $D \ll 2^{N_A+N_B+N_C}$ – which inherently obey an area-law – show a significantly different distribution $\omega(\lambda)$, with a high concentration around 0 but long tails on either side. This can be seen in Fig. B.2(b) for a single R-MPS instance of ψ_{ABC}. Nonetheless, the support of this type of distribution is even more tightly bounded by $|\lambda_k| \leq \mu_M^{1/M}$.

In this thermodynamic limit the Chebyshev approximation can be understood from Eq. (B.9), as $|\lambda| \to f(\lambda)$. The support of $\omega(\lambda)$ for the two classes of states discussed above is typically very different, being much wider for R-MPS. The effect of this wider range can be seen if we consider the error in the Chebyshev approximation, $f(\lambda) - |\lambda|$, as a function of λ. By construction this error is spread roughly throughout the interval, with alternating sign. In the case of $\omega(\lambda)$ for R-MPS, the peak at zero, together with the large support, concentrate a large number of eigenvalues into a small region with the same signed error. This gives a negative bias that does not exist when $\omega(\lambda)$ is more even throughout the interval, as for random pure states. Therefore, Chebyshev methods are expected to have a larger error for area law states.

B.7 Neural Network Details

In deep neural networks, the unknown function f mapping inputs to outputs is approximated with a directed graph organized in layers, where the first layer is the input data and the last one is the output. In our case, the input data consists of the numbers $(N_A, N_B, \mu_2, \ldots, \mu_M)$, namely, the number of spins in each subsystem and the non-trivial moments. The value $s_k^{(\ell)}$ of the k-th node in layer ℓ is updated via the equation $s_k^{(\ell)}{=}A^\ell \left[\sum_j w_{kj}^{(\ell-1)} s_j^{(\ell-1)}\right]$, where A^ℓ is an appropriate (typically non-linear) activation function and $w_{kj}^{(\ell-1)}$ is the weight between node k in layer ℓ and node j in $\ell{-}1$. The training procedure consists in finding the optimal weights w by minimizing a suitable *cost function*.

In our numerical investigations, we use the Hyperopt [319] and Keras [320] packages to find the optimal network structure, including the number of hidden layers. For example, for $M = 3$, the resulting network consists of two hidden layers, both with rectified linear unit (ReLU) activation functions, with 100 and 56 neurons respectively. For $M = 10$ however, the resulting network consists of three hidden layers, with exponential linear unit (ELU), ReLU and linear activation functions, with 61, 87 and 47 neurons respectively.

B.8 Neural Network Error Analysis

We have seen that neural networks can provide very accurate predictions for the logarithmic negativity given only a few moments. However, in a experimental setting it is also important to be aware of the effective size of any error bars. We note first of all, that the question of neural networks themselves providing error estimates is an important yet open problem. On the other hand, broad values for the error can be inferred statistically from the training process. For example, the standard deviation in the error of the neural networks predictions for unseen test data is a good guide — and this is exactly the standard deviation of the distributions plotted in the insets of Fig. 2. One could also calculate the standard deviation as a function of $\mathcal{E}$ (i.e. windowed between $\mathcal{E}-\delta, \mathcal{E}+\delta$, for a small value δ) to give a more dynamic estimate. As this data has never been used to optimise the neural network, there is every reason to associate similar error bounds for any new experimental data use to predict $\mathcal{E}$, as long it is qualitatively similar. This similarity, in the example of the quantum quench shown in Fig. 3, is achieved by training with both area- and volume-law entangled random states, yielding errors that roughly match those of the test data.

Another area of concern might be the effect of supplying slightly inaccurate moments, $\{\mu_m\}$, (due to imperfect copies or measurement error for example) to the neural network. We note however that since the neural networks produce smooth functions of their variables, any errors are well controlled – small errors will only produce small changes in the estimate.

B.9 Neural Network Sensitivity

Although the neural network produces a smooth output, there is the question of how much the predicted logarithmic negativity changes if the measured moments are imprecise. We have shown very few moments are required to estimate the entanglement accurately, and these can be measured with an effort that scales linearly with the subsystem sizes. However, to be scalable, it is necessary that the number of repeat measurements, R, which sets the standard error $\Delta\mu_m$ via (B.4), also scales efficiently, i.e., sub-exponentially.

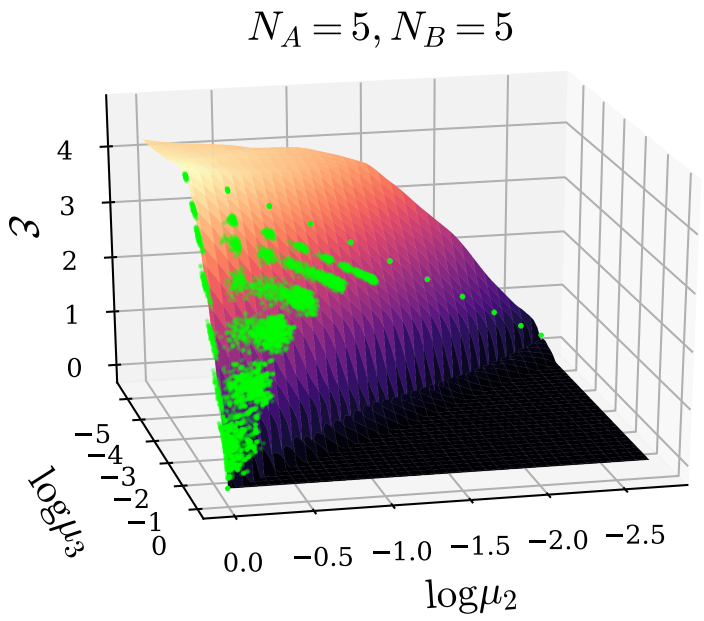


Figure B.3: Entanglement as a function of μ_2 and μ_3 for the random set of density matrices with $N_A = N_B = 5$. The green points represent the real logarithmic negativities, while the surface shows the neural network's predictions for the whole space.

To give a feeling for the dependence of the logarithmic negativity estimator on the the moments μ_m, in Fig. B.3 we show a plot of $\mathcal{E}^{\mathrm{ML}}_{\mathrm{M}=3}$ as a function of μ_2 and μ_3 (for fixed N_A and N_B which are also inputs to the neural network). In this plot, the scatter points are real data from the random training set, while the surface is effective function of the neural network. As can be seen, the neural network is smooth (apart from the enforced lower bound of 0), with the vast majority of data overlaying areas with reasonable gradient.

The error in the entanglement estimation is given by the gradient of the neural network estimator with respect to the set of input moments $\{\mu_m\}$, multiplied by the measurement error, namely:

$$\Delta\mathcal{E} \approx \sum_m \Delta\mu_m \partial\mathcal{E}/\partial\mu_m, \tag{B.10}$$

where the partial derivative denotes the neural network sensitivity. Since $\Delta\mu_m \propto R^{-1/2}$, as N_{A+B} increases, it is required that the ratio of the sensitivity to R remains

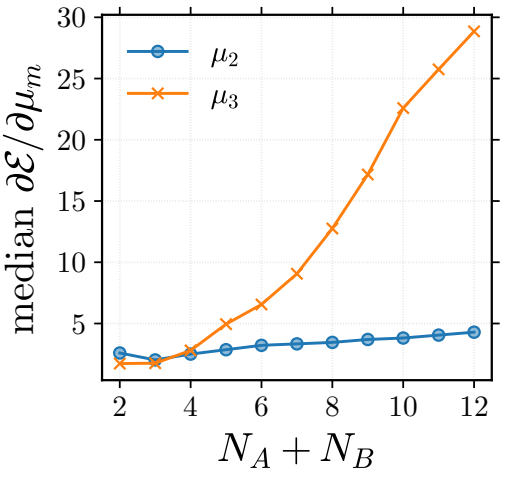


Figure B.4: Typical gradient of the neural network estimator with respect to the measured moments μ_2 and μ_3.

manageable. Specifically, this requires that the sensitivity does not diverge as the subsystem sizes grow. To probe this, we take the set of random data and compute the gradient of $\mathcal{E}^{\mathrm{ML}}_{\mathrm{M}=3}$ with respect to mu_2 and mu_3 for each data point in the set. In figure B.4 we plot the typical gradient, the median, as a function of total system size N_{A+B}. As is shown, there is a roughly linear increase in sensitivity with respect to both moments, with the absolute value in the manageable range < 30, and more accuracy required for μ_3. This confirms that the number of required measurements, R does not scale exponentially with the system sizes.

B.10 Comparison with approximate state reconstruction methods

Unlike our method, approximate polynomial state reconstruction schemes, e.g. based tensor networks [139, 238] or neural network states [239, 240], are normally focused on finding an approximate representation of a state that accurately reproduces experimental observables. Let ρ_r be the experimentally reconstructed state from a polynomial number of measurements and suppose that the expectation values predicted by this state are "close enough" to those predicted by the true state ρ. This means that $\mathrm{Tr}[\rho_r A] \approx \mathrm{Tr}[\rho A]$, for any observable A. Alternatively we may ask that $\mathrm{Tr}[\rho_r M_i] \approx \mathrm{Tr}[\rho M_i]$, for any positive operator valued measurement (POVM) M_i.

Because of Helstrom's theorem [241] the maximal distance between the distributions obtained from the true and reconstructed state is given by the trace distance

$$D(\rho_r, \rho) = \frac{1}{2}\|\rho - \rho_r\|_1 = \frac{1}{2}\max_{\{M\}} \sum_i \mathrm{Tr}[M_i(\rho - \rho_r)] \,. \tag{B.11}$$

Therefore, if the states ρ and ρ_r are close with respect to the trace distance, any expectation value obtained from the reconstructed state is close to the true value. We may therefore assume that a "good reconstruction" has $D(\rho, \rho_r) \approx 0$. Although this is true for any expectation value, entanglement measures are non-analytic functions of expectation values and, as such, may display a higher sensitivity to small errors in the reconstruction. Indeed, many entanglement measures are not continuous, so

$$D(\rho, \sigma) \to 0, \qquad \text{does not imply} \qquad |\mathcal{E}(\rho) - \mathcal{E}(\sigma)| \to 0. \tag{B.12}$$

In fact, the entanglement negativity does not even satisfy the requirement of asymptotic continuity (see table 15.2 in [241]). Consider two N qubit states ρ and σ such that $\|\rho - \sigma\| \to 0$ for large N (many particles). A measure if asymptotically continuous if $\|\rho - \sigma\|_1 \to 0$ implies $|\mathcal{E}(\rho) - \mathcal{E}(\sigma)|/\log(d_N) \to 0$ where $d_N = 2^N$ is the Hilbert space dimension. Since the negativity does not satisfy asymptotic continuity the difference between the entanglement of two 'close' many-body state can diverge faster than $\mathcal{O}(\log(d_N)) \simeq \mathcal{O}(N)$ for large N. Because of this important point, the negativity obtained from states reconstructed with a polynomial number of measurements can not be considered a valid approximation of the true negativity.

In contrast, our method avoids the intermediate step of reconstructing the quantum state and uses a polynomial number of measurements for the sole purpose of reconstructing the negativity. As shown in the previous section, this has the advantage that the prediction is accurate and the error is well controlled.

B.11 Additional Numerical Results

In this section we further demonstrate the ability of the neural network, trained only with random states, to accurately predict the logarithmic negativity in several different

classes of physical states. These include states likely to be the most challenging for the neural network due to their high degree of symmetry, which is not enforced in the random training set. We note that it is highly likely that these results could be improved even further with the adoption of specialized training sets, chosen to match the model under study. However, for the sake of generality, we focus here on training only with a set of random states with no underlying physical assumptions.

B.12 Ground-states through a quantum phase transition

Quantum phase transitions are a topic of interest for various fields of physics, noted for their display of various entanglement structures. Here we study the XX-model with transverse field:

$$H_{\mathrm{XX}} = \sum_{i=1}^{L-1} \left(\sigma_i^X \sigma_{i+1}^X + \sigma_i^Y \sigma_{i+1}^Y\right) + B_Z \sum_{i=1}^{L} \sigma_i^Z, \tag{B.13}$$

where B_Z denotes the magnetic field. This system undergoes a quantum phase transition at $B_Z = 1$, above which the system enters a phase with a separable ground-state.

In Fig. B.5 we show the logarithmic negativity, exact and predicted from moments using the neural network, for a variety of sizes of two adjacent blocks of spins within the groundstate of H_{XX} across the transition. The spin chain is divided into three blocks, of size N_A, N_B and N_C. System C is traced out, then the entanglement is found between A and B. This state is generally mixed as long as $N_C > 0$. The neural network is exactly as defined in the main text — trained solely with random states — and we show results for using 3 and 6 moments (copies). As can be seen from the figure, the key features of the transition are well reproduced in both cases, with the critical point ($B_Z = 1$) clearly defined. While there are some fluctuations in the quantitative estimate of $\mathcal{E}$ for 3 copies, these are significantly suppressed by raising the number of moments used to 6 copies. It is remarkable that the neural network can capture the entanglement properties of these highly symmetric ground-states

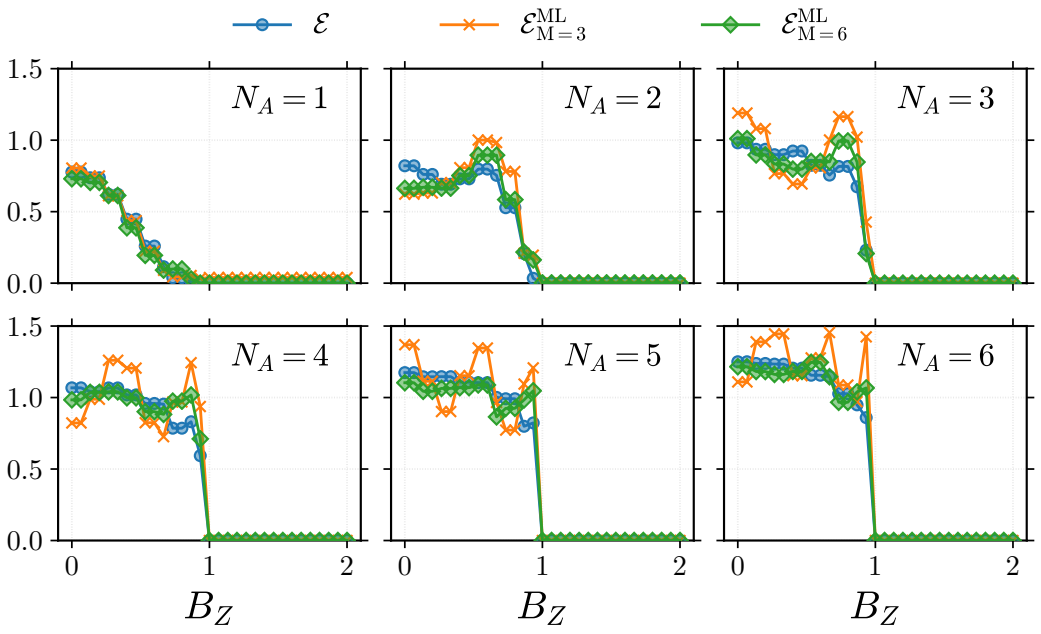


Figure B.5: Entanglement estimation in the ground-state of the XX-model phase across the phase transition driven by transverse field, B_Z. The groundstate, of total length $L = 20$ is tri-partitioned, with two adjacent subsystems of size $N_A = N_B$ and environment size $N_C = L - N_A + N_B$. The entanglement between A and B is then computed and estimated from ρ_{AB}. The blue, orange and green lines show the true logarithmic negativity, the neural network estimated quantity with 3 copies, and the neural network estimated quantity with 6 copies respectively.

despite having only been trained with random states.

B.13 W-state

Our training set is composed of random states with no imposed symmetry, in which highly entangled states represent the largest class. As such, low-entangled, highly symmetric states should generally pose the greatest challenge. Here we study the paradigmatic example of the W-state, which is defined as:

$$|W_L\rangle = \frac{1}{\sqrt{L}}\left(|100\cdots0\rangle + |010\cdots0\rangle + \cdots + |000\cdots1\rangle\right) \tag{B.14}$$

for L qubits. This state has a MPS representation with bond-dimension 2, making it one of the lowest possible entangled many-body states. Among other symmetries, it is also fully permutationally symmetric, making it highly distinct from the random training set. As in the previous sub-section, the W-state is divided into three blocks,

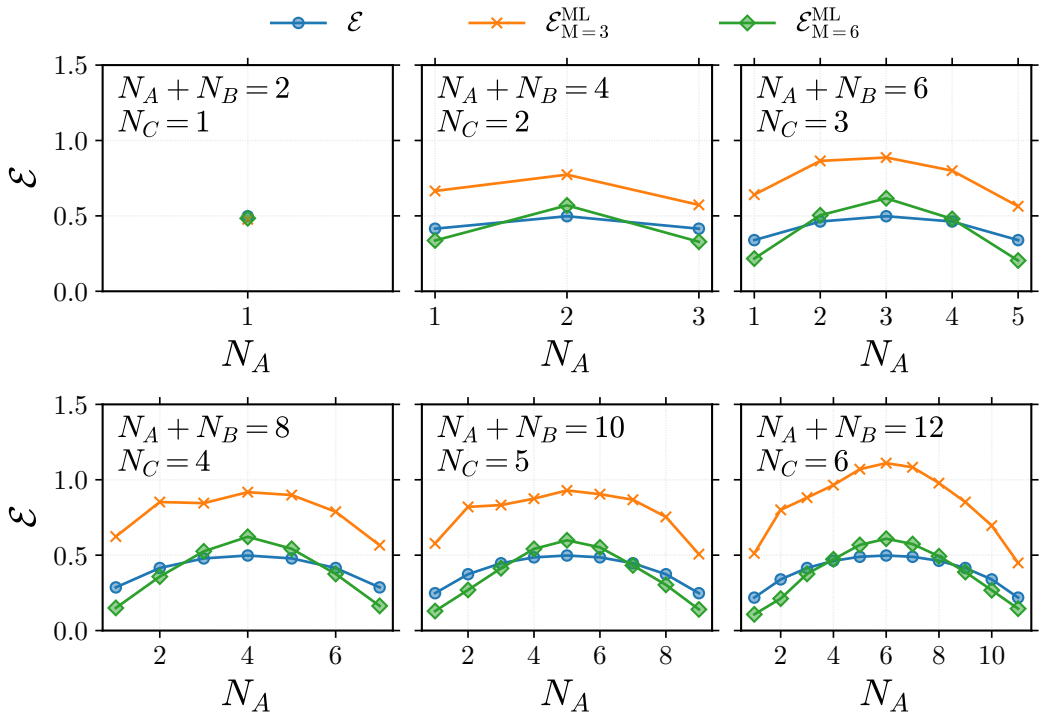


Figure B.6: Entanglement estimation in the W-state. We show here a representative sample of the entanglement estimation in the W-state for various lengths ($L = N_{A+B} + N_C$) and partitions, as a function of subsystem A size, N_A. The blue, orange and green lines show the true logarithmic negativity, the neural network estimated quantity with 3 copies, and the neural network estimated quantity with 6 copies respectively.

of size N_A, N_B and N_C. System C is traced out, then the entanglement is found between A and B. This state is generally mixed as long as $N_C > 0$. The choice of partitions is irrelevant due to the permutation symmetry. In Fig. B.6 we plot the real logarithmic negativity and the neural network predictions using 3 and 6 copies respectively, trained as before solely with random states. We show a few representative combinations of N_A, N_B and N_C. Although the overall trend is well captured by both the 3-copy and 6-copy neural network, the absolute accuracy is only reasonable for the 6-copy scheme. Given the atypicality of the W-state with respect to the training set, it is expected that this requirement for extra resources could be alleviated by training with more specialized states such as those with high degree of symmetry. We also note our protocol works best for highly entangled state (> 1 ebit of entanglement unlike the W-state here). Nonetheless neural network is

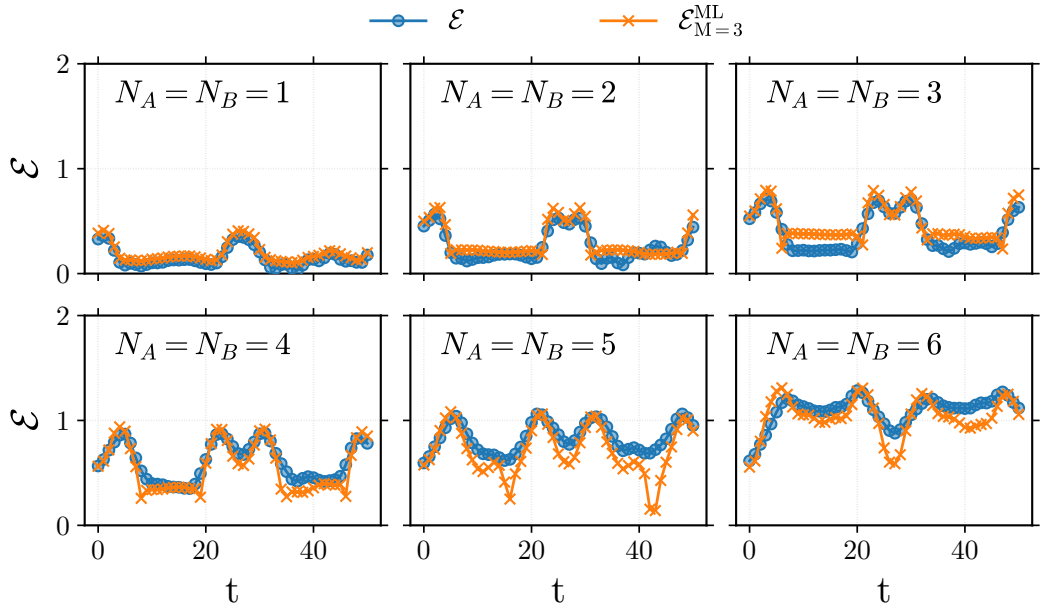


Figure B.7: Estimated entanglement when quenching across the Ising phase transition at $B_X = 0.5$. The initial state is the groundstate at $B_X = 0.5 + \Delta$, dynamics are generated by quenching with the Hamiltonian at $0.5 - \Delta$, taking $\Delta = 0.1$. The total size is $L = 20$ and the tri-partition is chosen so that subsystems A and B are adjacent and of equal size. The blue and orange lines show the true logarithmic negativity and the neural network estimated quantity with 3 copies respectively.

easily capable of identifying when the entanglement is relatively low (< 1 ebit of entanglement). In these cases, instead of resorting to more copies, it would also be feasible to switch to MPS-tomography [238], which is efficient for low levels of entanglement.

B.14 Quench across a phase transition

In the main text we quench from the Neel-state (which can be thought of as the ground-states of the XXZ-model with infinite anisotropy) to the isotropic Heisenberg point, which is the location of a Kosterlitz-Thouless phase transition. This type of quench quickly generates volume-law entanglement, and this is accurately captured by our neural network approach, as show in Fig. 3 of the main text. One might also consider a different quench, across a different type of quantum phase transition.

Here, we take the paradigmatic example of the transverse field Ising model:

$$H_{\text{Ising}}(B_X) = \sum_{i=1}^{L-1} \sigma_i^Z \sigma_{i+1}^Z + B_X \sum_{i=1}^{L} \sigma_i^X, \tag{B.15}$$

where B_X is the magnetic field, which induces a second order phase at the critical point $B_X = 0.5$.

In Fig. B.7 we show the logarithmic negativity as a function of time, exact and predicted from moments using the neural network, for a variety of sizes of two adjacent blocks of spins during a quench across this phase transition. Specifically, we take the initial state as the ground-state of $H_{\text{Ising}}(1+\delta)$, and evolve it with $H_{\text{Ising}}(1-\delta)$. As before, the spin chain is divided into three blocks, of size N_A, N_B and N_C. System C is traced out, then the entanglement is found between A and B. This state is generally mixed as long as $N_C > 0$. The neural network is also the same as before, and we only show results for 3 moments (copies) since this is already sufficient for good accuracy. This further confirms that the training set chosen is particularly suitable for highly entangled states such as those generated in quenches.

Bibliography

[1] Johnnie Gray, Abolfazl Bayat, Reuben K Puddy, Charles G Smith, and Sougato Bose. Unravelling quantum dot array simulators via singlet-triplet measurements. *Phys. Rev. B*, 94(19):195136, 2016.

[2] Johnnie Gray, Sougato Bose, and Abolfazl Bayat. Many-body localization transition: Schmidt gap, entanglement length, and scaling. *Phys. Rev. B*, 97(20):201105, 2018.

[3] Johnnie Gray, Leonardo Banchi, Abolfazl Bayat, and Sougato Bose. Machine-learning-assisted many-body entanglement measurement. *Phys. Rev. Lett.*, 121(15):150503, 2018.

[4] Johnnie Gray. Fast computation of many-body entanglement. *arXiv preprint arXiv:1809.01685*, 2018.

[5] Johnnie Gray. `quimb`: a python library for quantum information and many-body calculations. *Journal of Open Source Software*, 3(29):819, 2018.

[6] Daniel GA Smith and Johnnie Gray. `opt_einsum` - a python package for optimizing contraction order for einsum-like expressions. *Journal of Open Source Software*, 3(26):753, 2018.

[7] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[8] David Deutsch. Quantum theory, the church-turing principle and the universal quantum computer. In *Proceedings of the Royal Society of London A: Mathe-*

matical, Physical and Engineering Sciences, volume 400, pages 97–117. The Royal Society, 1985.

[9] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.*, 41(2):303–332, 1999.

[10] Iulia Buluta and Franco Nori. Quantum Simulators. *Science*, 326(5949):108–111, October 2009.

[11] Vivien M. Kendon, Kae Nemoto, and William J. Munro. Quantum analogue computing. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1924):3609–3620, August 2010.

[12] Philipp Hauke, Fernando M. Cucchietti, Luca Tagliacozzo, Ivan Deutsch, and Maciej Lewenstein. Can one trust quantum simulators? *Reports on Progress in Physics*, 75(8):082401, 2012.

[13] J. Ignacio Cirac and Peter Zoller. Goals and opportunities in quantum simulation. *Nature Physics*, 8(4):264–266, April 2012.

[14] Tobias Schaetz, Chris R. Monroe, and Tilman Esslinger. Focus on quantum simulation. *New Journal of Physics*, 15(8):085009, 2013.

[15] M. Georgescu, I. S. Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153–185, March 2014. 00172.

[16] T. H. Johnson, S. R. Clark, and D. Jaksch. What is a quantum simulator? *arXiv:1405.2831 [quant-ph]*, May 2014. 00010.

[17] Anastasios Vergis, Kenneth Steiglitz, and Bradley Dickinson. The complexity of analog computation. *Mathematics and Computers in Simulation*, 28(2):91–113, April 1986.

[18] Seth Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, August 1996.

[19] David P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik*, 48(9-11):771–783, 2000.

[20] Jens Eisert and Martin B. Plenio. Focus on Quantum Information and Many-Body Theory. *New Journal of Physics*, 12(2):025001, 2010.

[21] Bei Zeng, Xie Chen, Duan-Lu Zhou, and Xiao-Gang Wen. Quantum Information Meets Quantum Matter – From Quantum Entanglement to Topological Phase in Many-Body Systems. *arXiv:1508.02595 [cond-mat, physics:quant-ph]*, August 2015.

[22] Nicolas Laflorencie. Quantum entanglement in condensed matter systems. *Physics Reports*, 646:1–59, 2016.

[23] Martin B Plenio and Shashank Virmani. An introduction to entanglement measures. *Quantum Inf. Comput.*, 7(1):1–51, 2007.

[24] Karol Życzkowski, Paweł Horodecki, Anna Sanpera, and Maciej Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58(2):883, 1998.

[25] Jinhyoung Lee, MS Kim, YJ Park, and S Lee. Partial teleportation of entanglement in a noisy environment. *J. Mod. Opt.*, 47(12):2151–2164, October 2000.

[26] Guifré Vidal and Reinhard F Werner. Computable measure of entanglement. *Phys. Rev. A*, 65(3):032314, February 2002.

[27] Martin B Plenio. Logarithmic negativity: a full entanglement monotone that is not convex. *Phys. Rev. Lett.*, 95(9):090503, 2005.

[28] Yichen Huang. Computing quantum discord is np-complete. *New J. Phys.*, 16(3):033027, 2014.

[29] Lluís Masanes. All bipartite entangled states are useful for information processing. *Physical Review Letters*, 96(15):150501, 2006.

[30] Gerardo Adesso and Fabrizio Illuminati. Strong monogamy of bipartite and genuine multipartite entanglement: the gaussian case. *Physical review letters*, 99(15):150501, 2007.

[31] Hui Li and F. D. M. Haldane. Entanglement spectrum as a generalization of entanglement entropy: Identification of topological order in non-abelian fractional quantum hall effect states. *Phys. Rev. Lett.*, 101:010504, Jul 2008.

[32] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner. Entanglement properties of the harmonic chain. *Phys. Rev. A*, 66:042327, Oct 2002.

[33] Tobias J. Osborne and Michael A. Nielsen. Entanglement in a simple quantum phase transition. *Phys. Rev. A*, 66:032110, Sep 2002.

[34] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90:227902, Jun 2003.

[35] Raoul Dillenschneider. Quantum discord and quantum phase transition in spin chains. *Phys. Rev. B*, 78:224413, Dec 2008.

[36] Abolfazl Bayat, Henrik Johannesson, Sougato Bose, and Pasquale Sodano. An order parameter for impurity systems at quantum criticality. *Nat. Commun.*, 5, May 2014.

[37] Pasquale Calabrese and John Cardy. Entanglement entropy and conformal field theory. *J. Phys. A*, 42(50):504005, 2009.

[38] Pasquale Calabrese, John Cardy, and Erik Tonni. Entanglement negativity in quantum field theory. *Phys. Rev. Lett.*, 109(13):130502, 2012.

[39] Pasquale Calabrese, John Cardy, and Erik Tonni. Entanglement negativity in extended systems: a field theoretical approach. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(02):P02008, 2013.

[40] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2 – 30, 2003.

[41] Fabian HL Essler, Holger Frahm, Frank Göhmann, Andreas Klümper, and Vladimir E Korepin. *The one-dimensional Hubbard model*. Cambridge University Press, 2005.

[42] Alexander Altland and Ben D Simons. *Condensed matter field theory*. Cambridge University Press, 2010.

[43] Qijin Chen, Jelena Stajic, Shina Tan, and Kathryn Levin. Bcs–bec crossover: From high temperature superconductors to ultracold superfluids. *Physics Reports*, 412(1):1–88, 2005.

[44] Antoine Georges and Thierry Giamarchi. Strongly correlated bosons and fermions in optical lattices. *arXiv preprint arXiv:1308.2684*, 2013.

[45] Fabian HL Essler, Vladimir E Korepin, and Kareljan Schoutens. Complete solution of the one-dimensional hubbard model. *Physical review letters*, 67(27):3848, 1991.

[46] Fabian HL Essler, Vladimir E Korepin, and Kareljan Schoutens. New exactly solvable model of strongly correlated electrons motivated by high-t c superconductivity. *Physical review letters*, 68(19):2960, 1992.

[47] Patrick A. Lee, Naoto Nagaosa, and Xiao-Gang Wen. Doping a mott insulator: Physics of high-temperature superconductivity. *Rev. Mod. Phys.*, 78:17–85, Jan 2006.

[48] James D Whitfield, Jacob Biamonte, and Alán Aspuru-Guzik. Simulation of electronic structure hamiltonians using quantum computers. *Molecular Physics*, 109(5):735–750, 2011.

[49] Ulrich Schollwöck, Johannes Richter, Damian JJ Farnell, and Raymond F Bishop. *Quantum magnetism*, volume 645. Springer, 2008.

[50] Hans Bethe. Zur theorie der metalle. *Zeitschrift für Physik*, 71(3-4):205–226, 1931.

[51] Pascual Jordan and Eugene Paul Wigner. über das paulische äquivalenzverbot. In *The Collected Works of Eugene Paul Wigner*, pages 109–129. Springer, 1993.

[52] Rolando Somma, Gerardo Ortiz, James E Gubernatis, Emanuel Knill, and Raymond Laflamme. Simulating physical phenomena by quantum networks. *Physical Review A*, 65(4):042323, 2002.

[53] J. Eisert, M. Friesdorf, and C. Gogolin. Quantum many-body systems out of equilibrium. *Nature Physics*, 11(2):124–130, February 2015. 00054.

[54] Christian Gogolin and Jens Eisert. Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems. *Reports on Progress in Physics*, 79(5):056001, 2016.

[55] Subir Sachdev. *Quantum phase transitions*. Cambridge University Press, 2011.

[56] Simon Braun, Mathis Friesdorf, Sean S. Hodgman, Michael Schreiber, Jens Philipp Ronzheimer, Arnau Riera, Marco del Rey, Immanuel Bloch, Jens Eisert, and Ulrich Schneider. Emergence of coherence and the dynamics of quantum phase transitions. *Proceedings of the National Academy of Sciences*, 112(12):3641–3646, March 2015.

[57] Anthony J. Short and Terence C. Farrelly. Quantum equilibration in finite time. *New Journal of Physics*, 14(1):013063, 2012.

[58] Anthony J. Short. Equilibration of quantum systems and subsystems. *New Journal of Physics*, 13(5):053009, 2011.

[59] Noah Linden, Sandu Popescu, Anthony J. Short, and Andreas Winter. Quantum mechanical evolution towards thermal equilibrium. *Physical Review E*, 79(6):061103, June 2009.

[60] Peter Reimann. Foundation of Statistical Mechanics under Experimentally Realistic Conditions. *Phys. Rev. Lett.*, 101(19):190403, November 2008.

[61] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, 452(7189):854, 2008.

[62] Markus P Müller, Emily Adlam, Lluís Masanes, and Nathan Wiebe. Thermalization and canonical typicality in translation-invariant quantum lattice systems. *Communications in Mathematical Physics*, 340(2):499–561, 2015.

[63] Josh M Deutsch. Quantum statistical mechanics in a closed system. *Physical Review A*, 43(4):2046, 1991.

[64] Mark Srednicki. Chaos and quantum thermalization. *Physical Review E*, 50(2):888, 1994.

[65] Mark Srednicki. The approach to thermal equilibrium in quantized chaotic systems. *Journal of Physics A: Mathematical and General*, 32(7):1163, 1999.

[66] Anatoli Polkovnikov, Krishnendu Sengupta, Alessandro Silva, and Mukund Vengalattore. Nonequilibrium dynamics of closed interacting quantum systems. *Reviews of Modern Physics*, 83(3):863–883, August 2011. 00026.

[67] Jean-Sébastien Caux and Fabian HL Essler. Time evolution of local observables after quenching to an integrable model. *Physical review letters*, 110(25):257203, 2013.

[68] Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, 2016.

[69] Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, May 2016.

[70] Amy C Cassidy, Charles W Clark, and Marcos Rigol. Generalized thermalization in an integrable lattice system. *Physical review letters*, 106(14):140405, 2011.

[71] Marcus Kollar, F Alexander Wolf, and Martin Eckstein. Generalized gibbs ensemble prediction of prethermalization plateaus and their relation to nonthermal steady states in integrable systems. *Physical Review B*, 84(5):054304, 2011.

[72] Bruno Bertini, Fabian HL Essler, Stefan Groha, and Neil J Robinson. Prethermalization and thermalization in models with weak integrability breaking. *Physical review letters*, 115(18):180601, 2015.

[73] P. W. Anderson. Absence of Diffusion in Certain Random Lattices. *Phys. Rev.*, 109(5):1492–1505, March 1958.

[74] D. M. Basko, I. L. Aleiner, and B. L. Altshuler. Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states. *Ann. Phys.*, 321(5):1126–1205, May 2006.

[75] Arijeet Pal and David A. Huse. Many-body localization phase transition. *Phys. Rev. B*, 82(17):174411, November 2010.

[76] John Z. Imbrie. Diagonalization and Many-Body Localization for a Disordered Quantum Spin Chain. *Phys. Rev. Lett.*, 117(2):027201, July 2016.

[77] John Z Imbrie. On many-body localization for quantum spin chains. *Journal of Statistical Physics*, 163(5):998–1048, 2016.

[78] Dmitry A Abanin and Zlatko Papić. Recent progress in many-body localization. *Annalen der Physik*, 529(7):1700169, 2017.

[79] David J. Luitz, Nicolas Laflorencie, and Fabien Alet. Many-body localization edge in the random-field Heisenberg chain. *Phys. Rev. B*, 91(8):081103, February 2015.

[80] Maksym Serbyn, Z. Papić, and Dmitry A. Abanin. Local Conservation Laws and the Structure of the Many-Body Localized States. *Phys. Rev. Lett.*, 111(12):127201, September 2013.

[81] David A. Huse, Rahul Nandkishore, and Vadim Oganesyan. Phenomenology of fully many-body-localized systems. *Phys. Rev. B*, 90(17):174202, November 2014.

[82] Rahul Nandkishore, Sarang Gopalakrishnan, and David A. Huse. Spectral features of a many-body-localized system weakly coupled to a bath. *Phys. Rev. B*, 90(6):064203, August 2014.

[83] Anushya Chandran, Isaac H. Kim, Guifre Vidal, and Dmitry A. Abanin. Constructing local integrals of motion in the many-body localized phase. *Phys. Rev. B*, 91(8):085425, February 2015.

[84] V. Ros, M. Müller, and A. Scardicchio. Integrals of motion in the many-body localized phase. *Nucl. Phys. B*, 891:420–465, February 2015.

[85] Cécile Monthus. Many-body localization: Construction of the emergent local conserved operators via block real-space renormalization. *J. Stat. Mech. Theor. Exp*, 2016(3):033101, 2016.

[86] M. B. Hastings. Lieb-Schultz-Mattis in higher dimensions. *Physical Review B*, 69(10):104431, March 2004.

[87] Bruno Nachtergaele and Robert Sims. Lieb-Robinson Bounds and the Exponential Clustering Theorem. *Communications in Mathematical Physics*, 265(1):119–130, July 2006.

[88] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

[89] Michael M. Wolf, Frank Verstraete, Matthew B. Hastings, and J. Ignacio Cirac. Area Laws in Quantum Systems: Mutual Information and Correlations. *Phys. Rev. Lett.*, 100(7):070502, February 2008. 00239.

[90] Lluís Masanes. Area law for the entropy of low-energy states. *Physical Review A*, 80(5):052104, 2009.

[91] J. Eisert, M. Cramer, and M. B. Plenio. \textit{Colloquium} : Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277–306, February 2010.

[92] Fernando G. S. L. Brandão and Michał Horodecki. Exponential Decay of Correlations Implies Area Law. *Communications in Mathematical Physics*, 333(2):761–798, January 2015.

[93] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.*, 326(1):96–192, 2011.

[94] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014.

[95] Jacob C Bridgeman and Christopher T Chubb. Hand-waving and interpretive dance: an introductory course on tensor networks. *J. Phys. A*, 50(22):223001, 2017.

[96] Stellan Östlund and Stefan Rommer. Thermodynamic Limit of Density Matrix Renormalization. *Phys. Rev. Lett.*, 75(19):3537–3540, November 1995.

[97] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Physical Review B*, 48(14):10345–10356, October 1993.

[98] Steven R White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863, 1992.

[99] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91(14):147902, 2003.

[100] Guifré Vidal. Efficient Simulation of One-Dimensional Quantum Many-Body Systems. *Phys. Rev. Lett.*, 93(4):040502, July 2004.

[101] Norbert Schuch, David Pérez-García, and Ignacio Cirac. Classifying quantum phases using matrix product states and projected entangled pair states. *Physical Review B*, 84(16):165139, October 2011.

[102] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. Many-body physics with ultracold gases. *Rev. Mod. Phys.*, 80(3):885, 2008.

[103] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbène. Quantum simulations with ultracold quantum gases. *Nat. Phys.*, 8(4):267–276, April 2012.

[104] Markus Greiner, Olaf Mandel, Tilman Esslinger, Theodor W. Hänsch, and Immanuel Bloch. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature*, 415(6867):39–44, January 2002.

[105] Jonathan Simon, Waseem S. Bakr, Ruichao Ma, M. Eric Tai, Philipp M. Preiss, and Markus Greiner. Quantum simulation of antiferromagnetic spin chains in an optical lattice. *Nature*, 472(7343):307–312, April 2011. 00475.

[106] Michael Schreiber, Sean S. Hodgman, Pranjal Bordia, Henrik P. Lüschen, Mark H. Fischer, Ronen Vosk, Ehud Altman, Ulrich Schneider, and Immanuel Bloch. Observation of many-body localization of interacting fermions in a quasirandom optical lattice. *Science*, 349(6250):842–845, August 2015.

[107] T. Schaetz, D. Leibfried, J. Chiaverini, M. D. Barrett, J. Britton, B. DeMarco, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland. Towards a scalable quantum computer/simulator based on trapped ions. *Applied Physics B*, 79(8):979–986, December 2004.

[108] R. Blatt and C. F. Roos. Quantum simulations with trapped ions. *Nat. Phys.*, 8(4):277–284, April 2012.

[109] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos. Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature*, 511(7508):202–205, July 2014. 00085.

[110] J. Q. You and Franco Nori. Superconducting circuits and quantum information. *Physics Today*, 58(11):42–47, November 2005.

[111] J. Q. You and Franco Nori. Atomic physics and quantum optics using superconducting circuits. *Nature*, 474(7353):589–597, June 2011.

[112] John Clarke and Frank K. Wilhelm. Superconducting quantum bits. *Nature*, 453(7198):1031–1042, June 2008.

[113] Xinhua Peng, Jingfu Zhang, Jiangfeng Du, and Dieter Suter. Quantum Simulation of a System with Competing Two- and Three-Body Interactions. *Phys. Rev. Lett.*, 103(14):140501, September 2009.

[114] Zhaokai Li, Man-Hong Yung, Hongwei Chen, Dawei Lu, James D. Whitfield, Xinhua Peng, Alán Aspuru-Guzik, and Jiangfeng Du. Solving Quantum Ground-State Problems with Nuclear Magnetic Resonance. *Scientific Reports*, 1:88, September 2011.

[115] Jingfu Zhang, Man-Hong Yung, Raymond Laflamme, Alán Aspuru-Guzik, and Jonathan Baugh. Digital quantum simulation of the statistical mechanics of a frustrated magnet. *Nature Communications*, 3:880, June 2012.

[116] Alán Aspuru-Guzik and Philip Walther. Photonic quantum simulators. *Nature physics*, 8(4):285, 2012.

[117] Iulia Buluta and Franco Nori. Quantum simulators. *Science*, 326(5949):108–111, 2009.

[118] Collective Coulomb blockade in an array of quantum dots: A Mott-Hubbard approach. 72.

[119] Pierre Barthelemy and Lieven MK Vandersypen. Quantum dot systems: a versatile platform for quantum simulations. *Annalen der Physik*, 525(10-11):808–826, 2013.

[120] Song Yang, Abolfazl Bayat, and Sougato Bose. Spin-state transfer in laterally coupled quantum-dot chains with disorders. *Phys. Rev A*, 82(2):022336, August 2010.

[121] Umer Farooq, Abolfazl Bayat, Stefano Mancini, and Sougato Bose. Adiabatic many-body state preparation and information transfer in quantum dot arrays. *Phys. Rev. B*, 91(13):134303, April 2015.

[122] J. Salfi, J. A. Mol, R. Rahman, G. Klimeck, M. Y. Simmons, L. C. L. Hollenberg, and S. Rogge. Quantum simulation of the Hubbard model with dopant atoms in silicon. *Nat. Commun.*, 7:11342, April 2016.

[123] Enrico Prati, Masahiro Hori, Filippo Guagliardo, Giorgio Ferrari, and Takahiro Shinada. Anderson-mott transition in arrays of a few dopant atoms in a silicon transistor. *Nat. Nano.*, 7(7):443–447, 2012.

[124] R. K. Puddy, L. W. Smith, H. Al-Taie, C. H. Chong, I. Farrer, J. P. Griffiths, D. A. Ritchie, M. J. Kelly, M. Pepper, and C. G. Smith. Multiplexed charge-locking device for large arrays of quantum devices. *Appl. Phys. Lett.*, 107(14):143501, October 2015.

[125] Daniel Loss and David P. DiVincenzo. Quantum computation with quantum dots. *Phys. Rev A*, 57(1):120–126, January 1998.

[126] R Hanson, LP Kouwenhoven, JR Petta, Seigo Tarucha, and LMK Vandersypen. Spins in few-electron quantum dots. *Rev. Mod. Phys*, 79(4):1217, 2007.

[127] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard. Coherent Manipulation of Coupled Electron Spins in Semiconductor Quantum Dots. *Science*, 309(5744):2180–2184, September 2005.

[128] Jeremy Levy. Universal Quantum Computation with Spin-$1/2$ Pairs and Heisenberg Exchange. *Phys. Rev. Lett.*, 89(14):147902, September 2002.

[129] KD Petersson, CG Smith, D Anderson, P Atkinson, GAC Jones, and DA Ritchie. Charge and spin state readout of a double quantum dot coupled to a resonator. *Nano Lett.*, 10(8):2789–2793, 2010.

[130] M Jung, MD Schroer, KD Petersson, and JR Petta. Radio frequency charge sensing in inas nanowire double quantum dots. *App. Phys. Lett.*, 100(25):253508, 2012.

[131] Julien Basset, D-D Jarausch, Anna Stockklauser, Tobias Frey, Christian Reichl, Werner Wegscheider, Thomas Markus Ihn, Klaus Ensslin, and Andreas Wallraff. Single-electron double quantum dot dipole-coupled to a single photonic mode. *Phys. Rev. B*, 88(12):125312, 2013.

[132] SJ Chorley, J Wabnig, ZV Penfold-Fitch, KD Petersson, J Frake, CG Smith, and MR Buitelaar. Measuring the complex admittance of a carbon nanotube double quantum dot. *Phys. Rev. Lett.*, 108(3):036802, 2012.

[133] T Frey, PJ Leek, M Beck, Alexandre Blais, Thomas Ihn, Klaus Ensslin, and Andreas Wallraff. Dipole coupling of a double quantum dot to a microwave resonator. *Phys. Rev. Lett.*, 108(4):046807, 2012.

[134] JI Colless, AC Mahoney, JM Hornibrook, AC Doherty, H Lu, AC Gossard, and DJ Reilly. Dispersive readout of a few-electron double quantum dot with fast rf gate sensors. *Phys. Rev. Lett,*, 110(4):046805, 2013.

[135] MG House, Takashi Kobayashi, Bent Weber, SJ Hile, TF Watson, J van der Heijden, Sven Rogge, and MY Simmons. Radio frequency measurements of tunnel couplings and singlet-triplet spin states in si: P quantum dots. *Nat. Comm.*, 6, 2015.

[136] Terry Rudolph and Shashank Soyuz Virmani. A relational quantum computer using only two-qubit total spin measurement and an initial supply of highly mixed single-qubit states. *New J. Phys*, 7(1):228, 2005.

[137] G Mauro D'Ariano, Matteo GA Paris, and Massimiliano F Sacchi. Quantum tomography. *Adv. Imag. Electr. Phys.*, 128:206–309, 2003.

[138] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105(15), October 2010.

[139] Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin,

and Yi-Kai Liu. Efficient quantum state tomography. *Nat. Commun.*, 1:149, December 2010.

[140] C. Barthel, D. J. Reilly, C. M. Marcus, M. P. Hanson, and A. C. Gossard. Rapid Single-Shot Measurement of a Singlet-Triplet Qubit. *Phys. Rev. Lett.*, 103(16):160503, October 2009.

[141] M. D. Shulman, O. E. Dial, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby. Demonstration of Entanglement of Electrostatically Coupled Singlet-Triplet Qubits. *Science*, 336(6078):202–205, April 2012.

[142] Stephen Blundell. *Magnetism in condensed matter*. Oxford University Press, 2001.

[143] Sougato Bose. Quantum communication through an unmodulated spin chain. *Phys. Rev. Lett.*, 91(20):207901, 2003.

[144] Georgios M Nikolopoulos and Igor Jex. *Quantum State Transfer and Network Engineering*. Springer, 2014.

[145] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[146] Kevin M. O'Connor and William K. Wootters. Entangled rings. *Phys. Rev A*, 63(5):052302, April 2001.

[147] Thomas Baigneres, Pascal Junod, and Serge Vaudenay. How far can we go beyond linear cryptanalysis? In *Advances in Cryptology-Asiacrypt 2004*, pages 432–450. Springer, 2004.

[148] Carl W. Helstrom. Quantum detection and estimation theory. *J. Stat. Phys.*, 1(2):231–252, June 1969.

[149] Mischa Thesberg and Erik S Sørensen. General quantum fidelity susceptibilities for the j 1-j 2 chain. *Phys. Rev. B*, 84(22):224435, 2011.

[150] Shu Chen, Li Wang, Shi-Jian Gu, and Yupeng Wang. Fidelity and quantum phase transition for the Heisenberg chain with next-nearest-neighbor interaction. *Phys. Rev. E*, 76(6):061108, December 2007.

[151] F. Verstraete, M. Popp, and J. I. Cirac. Entanglement versus Correlations in Spin Systems. *Phys. Rev. Lett.*, 92(2):027901, January 2004.

[152] Markus Popp, Frank Verstraete, Miguel Angel Martín-Delgado, and J Ignacio Cirac. Localizable entanglement. *Phys. Rev. A*, 71(4):042306, 2005.

[153] L. Campos Venuti, C. Degli Esposti Boschi, and M. Roncaglia. Qubit Teleportation and Transfer across Antiferromagnetic Spin Chains. *Phys. Rev. Lett.*, 99(6):060401, August 2007.

[154] J. P. Barjaktarevic, R. H. McKenzie, J. Links, and G. J. Milburn. Measurement-Based Teleportation along Quantum Spin Chains. *Phys. Rev. Lett.*, 95(23):230501, November 2005.

[155] J. M. Taylor, J. R. Petta, A. C. Johnson, A. Yacoby, C. M. Marcus, and M. D. Lukin. Relaxation, dephasing, and quantum control of electron spins in double quantum dots. *Phys. Rev. B*, 76(3):035315, July 2007.

[156] Filip K Malinowski, Frederico Martins, Peter D Nissen, Edwin Barnes, Mark S Rudner, Saeed Fallahi, Geoffrey C Gardner, Michael J Manfra, Charles M Marcus, and Ferdinand Kuemmeth. Notch filtering the nuclear environment of a spin qubit. *arXiv:1601.06677 [Nat. Nanotechnol. (to be published)]*, 2016.

[157] Erika Kawakami, P Scarlino, DR Ward, FR Braakman, DE Savage, MG Lagally, Mark Friesen, SN Coppersmith, MA Eriksson, and LMK Vandersypen. Electrical control of a long-lived spin qubit in a si/sige quantum dot. *Nat. Nanotechnol.*, 9(9):666–670, 2014.

[158] Scott Hill and William K. Wootters. Entanglement of a Pair of Quantum Bits. *Phys. Rev. Lett.*, 78(26):5022–5025, June 1997.

[159] Matthew P Wardrop and Andrew C Doherty. Characterization of an exchange-based two-qubit gate for resonant exchange qubits. *Phys. Rev. B*, 93(7):075436, 2016.

[160] David Petrosyan and P Lambropoulos. Coherent population transfer in a chain of tunnel coupled quantum dots. *Opt. Commun*, 264(2):419–425, 2006.

[161] David Petrosyan, Georgios M Nikolopoulos, and P Lambropoulos. State transfer in static and dynamic spin chains with disorder. *Phys. Rev. A*, 81(4):042307, 2010.

[162] N Ares, FJ Schupp, A Mavalankar, G Rogers, J Griffiths, GAC Jones, I Farrer, DA Ritchie, CG Smith, A Cottet, et al. Sensitive radio-frequency measurements of a quantum dot by tuning to perfect impedance matching. *Phys. Rev. Applied*, 5(3):034011, 2016.

[163] T. A. Baart, N. Jovanovic, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen. Nanosecond-timescale spin transfer using individual electrons in a quadruple-quantum-dot device. *Appl. Phys. Lett.*, 109(4), 2016.

[164] KD Petersson, CG Smith, D Anderson, P Atkinson, GAC Jones, and DA Ritchie. Microwave-driven transitions in two coupled semiconductor charge qubits. *Phys. Rev. Lett.*, 103(1):016805, 2009.

[165] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux. Distribution of the Ratio of Consecutive Level Spacings in Random Matrix Ensembles. *Phys. Rev. Lett.*, 110(8):084101, February 2013.

[166] David J. Luitz and Yevgeny Bar Lev. The Ergodic Side of the Many-Body Localization Transition. *arXiv:1610.08993*, October 2016.

[167] Rahul Nandkishore and David A. Huse. Many-Body Localization and Thermalization in Quantum Statistical Mechanics. *Annu. Rev. Condens. Matter Phys.*, 6(1):15–38, 2015.

[168] Robert Sims and Gunter Stolz. Many-Body Localization: Concepts and Simple Models. *arXiv:1312.0577*, December 2013.

[169] Jens H. Bardarson, Frank Pollmann, and Joel E. Moore. Unbounded Growth of Entanglement in Models of Many-Body Localization. *Phys. Rev. Lett.*, 109(1):017202, July 2012.

[170] Norman Y. Yao, Chris R. Laumann, and Ashvin Vishwanath. Many-body localization protected quantum state transfer. *arXiv:1508.06995*, August 2015.

[171] R. Vasseur, S. A. Parameswaran, and J. E. Moore. Quantum revivals and many-body localization. *Phys. Rev. B*, 91(14):140202, April 2015.

[172] M. Serbyn, M. Knap, S. Gopalakrishnan, Z. Papić, N. Y. Yao, C. R. Laumann, D. A. Abanin, M. D. Lukin, and E. A. Demler. Interferometric probes of many-body localization. *Phys. Rev. Lett.*, 113(14):147204, October 2014.

[173] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe. Many-body localization in a quantum simulator with programmable random disorder. *Nature Phys.*, 12(10):907–911, October 2016.

[174] M. Friesdorf, A. H. Werner, W. Brown, V. B. Scholz, and J. Eisert. Many-Body Localization Implies that Eigenvectors are Matrix-Product States. *Phys. Rev. Lett.*, 114(17):170505, May 2015.

[175] Maksym Serbyn, Alexios A. Michailidis, Dmitry A. Abanin, and Z. Papić. Power-Law Entanglement Spectrum in Many-Body Localized Phases. *Phys. Rev. Lett.*, 117(16):160601, October 2016.

[176] Thorsten B. Wahl, Arijeet Pal, and Steven H. Simon. Entire spectrum of fully many-body localized systems using tensor networks. *arXiv:1609.01552*, September 2016.

[177] Carolyn Zhang, Frank Pollmann, S. L. Sondhi, and Roderich Moessner. Density-Matrix Renormalization Group study of Many-Body Localization in Floquet Eigenstates. *arXiv:1608.06411*, August 2016.

[178] Marko Žnidarič. Diffusive and Subdiffusive Spin Transport in the Ergodic Phase of a Many-Body Localizable System. *Phys. Rev. Lett.*, 117(4), 2016.

[179] Trithep Devakul, Vedika Khemani, Frank Pollmann, David Huse, and Shivaji Sondhi. Obtaining highly excited eigenstates of the localized XX chain via DMRG-X. *arXiv:1702.07721*, February 2017.

[180] Pedro Ponte, Z. Papić, François Huveneers, and Dmitry A. Abanin. Many-Body Localization in Periodically Driven Systems. *Phys. Rev. Lett.*, 114(14):140401, April 2015.

[181] Vedika Khemani, S. P. Lim, D. N. Sheng, and David A. Huse. Critical Properties of the Many-Body Localization Transition. *arXiv:1607.05756*, July 2016.

[182] Vedika Khemani, D. N. Sheng, and David A. Huse. Two universality classes for the many-body localization transition. *arXiv:1702.03932*, February 2017.

[183] Shao-Kai Jian and Hong Yao. Solvable SYK models in higher dimensions: A new type of many-body localization transition. *arXiv:1703.02051*, March 2017.

[184] J. T. Chayes, L. Chayes, D. S. Fisher, and T. Spencer. Finite-Size Scaling and Correlation Lengths for Disordered Systems. *Phys. Rev. Lett.*, 57(24):2999–3002, 1986.

[185] A. Chandran, C. R. Laumann, and V. Oganesyan. Finite size scaling bounds on many-body localized phase transitions. *arXiv:1509.04285*, September 2015.

[186] G. De Chiara, L. Lepori, M. Lewenstein, and A. Sanpera. Entanglement Spectrum, Critical Exponents, and Order Parameters in Quantum Spin Chains. *Phys. Rev. Lett.*, 109(23):237208, December 2012.

[187] L. Lepori, G. De Chiara, and A. Sanpera. Scaling of the entanglement spectrum near quantum phase transitions. *Phys. Rev. B*, 87(23):235107, June 2013.

[188] Asher Peres. Separability Criterion for Density Matrices. *Phys. Rev. Lett.*, 77(8):1413–1415, August 1996.

[189] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223(1):1–8, November 1996.

[190] M. B. Plenio. Logarithmic Negativity: A Full Entanglement Monotone That is not Convex. *Phys. Rev. Lett.*, 95(9), 2005.

[191] Hannu Wichterich, Javier Molina-Vilaplana, and Sougato Bose. Scaling of entanglement between separated blocks in spin chains at criticality. *Phys. Rev. A*, 80(1):010304, July 2009.

[192] Hannu Wichterich, Julien Vidal, and Sougato Bose. Universality of the negativity in the Lipkin-Meshkov-Glick model. *Phys. Rev. A*, 81(3):032311, March 2010.

[193] Cristiano De Nobili, Andrea Coser, and Erik Tonni. Entanglement entropy and negativity of disjoint intervals in CFT: Some numerical extrapolations. *J. Stat. Mech. Theor. Exp.*, 2015(6):P06021, 2015.

[194] Andrea Coser, Erik Tonni, and Pasquale Calabrese. Towards the entanglement negativity of two disjoint intervals for a one dimensional free fermion. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(3):033116, 2016.

[195] Andrea Coser, Erik Tonni, and Pasquale Calabrese. Spin structures and entanglement of two disjoint intervals in conformal field theories. *J. Stat. Mech. Theor. Exp.*, 2016(5):053109, 2016.

[196] Abolfazl Bayat, Pasquale Sodano, and Sougato Bose. Negativity as the entanglement measure to probe the kondo regime in the spin-chain kondo model. *Phys. Rev. B*, 81(6):064429, February 2010.

[197] Abolfazl Bayat, Sougato Bose, Pasquale Sodano, and Henrik Johannesson. Entanglement Probe of Two-Impurity Kondo Physics in a Spin Chain. *Phys. Rev. Lett.*, 109(6):066403, August 2012.

[198] Lisandro D. Dalcin, Rodrigo R. Paz, Pablo A. Kler, and Alejandro Cosimo. Parallel distributed computing using Python. *Adv. Water Resour.*, 34(9):1124–1139, September 2011.

[199] Vicente Hernandez, Jose E. Roman, and Vicente Vidal. SLEPc: A Scalable and Flexible Toolkit for the Solution of Eigenvalue Problems. *ACM Trans. Math. Softw.*, 31(3):351–362, September 2005.

[200] Don N. Page. Average entropy of a subsystem. *Phys. Rev. Lett.*, 71(9):1291–1294, August 1993.

[201] Michael E. Fisher and Michael N. Barber. Scaling Theory for Finite-Size Effects in the Critical Region. *Phys. Rev. Lett.*, 28(23):1516–1519, June 1972.

[202] Jonas A. Kjäll, Jens H. Bardarson, and Frank Pollmann. Many-Body Localization in a Disordered Quantum Ising Chain. *Phys. Rev. Lett.*, 113(10):107204, September 2014.

[203] Jérôme Houdayer and Alexander K. Hartmann. Low-temperature behavior of two-dimensional Gaussian Ising spin glasses. *Phys. Rev. B*, 70(1):014418, July 2004.

[204] Andrew C. Potter, Romain Vasseur, and A. Parameswaran, S. Universal Properties of Many-Body Delocalization Transitions. *Phys. Rev. X*, 5(3):031033, September 2015.

[205] Soumya Bera and Arul Lakshminarayan. Local entanglement structure across a many-body localization transition. *Physical Review B*, 93(13):134204, 2016.

[206] Fernando Iemini, Angelo Russomanno, Davide Rossini, Antonello Scardicchio, and Rosario Fazio. Signatures of many-body localization in the dynamics of two-site entanglement. *Physical Review B*, 94(21):214206, 2016.

[207] Giuseppe De Tomasi, Soumya Bera, Jens H. Bardarson, and Frank Pollmann. Quantum Mutual Information as a Probe for Many-Body Localization. *Phys. Rev. Lett.*, 118(1):016804, January 2017.

[208] Francesca Pietracaprina, Giorgio Parisi, Angelo Mariano, Saverio Pascazio, and Antonello Scardicchio. Entanglement critical length at the many-body localization transition. *arXiv:1610.09316*, October 2016.

[209] C Radhakrishna Rao. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc*, 37(3):81–91, 1945.

[210] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103(15):150502, 2009.

[211] Charles H Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70(13):1895, 1993.

[212] Charles H Bennett and Stephen J Wiesner. Communication via one-and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69(20):2881, 1992.

[213] Artur K Ekert. Quantum cryptography based on bells theorem. *Phys. Rev. Lett.*, 67(6):661, 1991.

[214] Luigi Amico, Rosario Fazio, Andreas Osterloh, and Vlatko Vedral. Entanglement in many-body systems. *Rev. Mod. Phys.*, 80(2):517, 2008.

[215] Paweł Horodecki and Artur Ekert. Method for direct detection of quantum entanglement. *Phys. Rev. Lett.*, 89(12):127902, 2002.

[216] AJ Daley, H Pichler, J Schachenmayer, and P Zoller. Measuring entanglement growth in quench dynamics of bosons in an optical lattice. *Phys. Rev. Lett.*, 109(2):020505, 2012.

[217] C Moura Alves and D Jaksch. Multipartite entanglement detection in bosons. *Phys. Rev. Lett.*, 93(11):110501, 2004.

[218] Leonardo Banchi, Abolfazl Bayat, and Sougato Bose. Entanglement entropy scaling in solid-state spin arrays via capacitance measurements. *Phys. Rev. B*, 94(24):241117, 2016.

[219] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J Cerf, Timothy C Ralph, Jeffrey H Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84(2):621, 2012.

[220] Rajibul Islam, Ruichao Ma, Philipp M Preiss, M Eric Tai, Alexander Lukin, Matthew Rispoli, and Markus Greiner. Measuring entanglement entropy in a quantum many-body system. *Nature*, 528(7580):77–83, 2015.

[221] Ioannis Pitsios, Leonardo Banchi, Adil S Rab, Marco Bentivegna, Debora Caprara, Andrea Crespi, Nicolò Spagnolo, Sougato Bose, Paolo Mataloni, Roberto Osellame, et al. Photonic simulation of entanglement growth after a spin chain quench. *arXiv:1603.02669*, 2016.

[222] Benni Reznik. Entanglement from the vacuum. *Found. Phys.*, 33(1):167–176, 2003.

[223] Bedoor Alkurtass, Abolfazl Bayat, Ian Affleck, Sougato Bose, Henrik Johannesson, Pasquale Sodano, Erik S Sørensen, and Karyn Le Hur. Entanglement structure of the two-channel kondo model. *Phys. Rev. B*, 93(8):081106, 2016.

[224] Leonardo Banchi, Filippo Colomo, and Paola Verrucchi. When finite-size corrections vanish: The s= 1 2 x x z model and the razumov-stroganov state. *Phys. Rev. A*, 80(2):022341, 2009.

[225] Tommaso Roscilde, Paola Verrucchi, Andrea Fubini, Stephan Haas, and Valerio Tognetti. Studying quantum spin systems through entanglement estimators. *Phys. Rev. Lett.*, 93(16):167203, 2004.

[226] Barbara M Terhal. Bell inequalities and the separability criterion. *Phys. Lett. A*, 271(5):319–326, 2000.

[227] Maciej Lewenstein, B Kraus, JI Cirac, and P Horodecki. Optimization of entanglement witnesses. *Phys. Rev. A*, 62(5):052310, 2000.

[228] Fernando GSL Brandao. Quantifying entanglement with witness operators. *Phys. Rev. A*, 72(2):022310, 2005.

[229] KMR Audenaert and MB Plenio. When are correlations quantum?verification and quantification of entanglement by simple measurements. *New J. Phys.*, 8(11):266, 2006.

[230] Daniel Cavalcanti and Marcelo O Terra Cunha. Estimating entanglement of unknown states. *Applied physics letters*, 89(8):084102, 2006.

[231] Jens Eisert, Fernando GSL Brandao, and Koenraad MR Audenaert. Quantitative entanglement witnesses. *New J. Phys.*, 9(3):46, 2007.

[232] O Gühne, M Reimpell, and RF Werner. Estimating entanglement measures in experiments. *Phys. Rev. Lett.*, 98(11):110502, 2007.

[233] O Gühne, M Reimpell, and RF Werner. Lower bounds on entanglement measures from incomplete information. *Phys. Rev. A*, 77(5):052317, 2008.

[234] Hilary A Carteret. Estimating the entanglement negativity from low-order moments of the partially transposed density matrix. *arXiv preprint arXiv:1605.08751*, 2016.

[235] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81(2):865, 2009.

[236] William K Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80(10):2245, 1998.

[237] Michael M Wolf, Geza Giedke, O Krüger, RF Werner, and J Ignacio Cirac. Gaussian entanglement of formation. *Phys. Rev. A*, 69(5):052320, 2004.

[238] BP Lanyon, C Maier, M Holzäpfel, T Baumgratz, C Hempel, P Jurcevic, I Dhand, AS Buyskikh, AJ Daley, M Cramer, et al. Efficient tomography of a quantum many-body system. *arXiv:1612.08000*, 2016.

[239] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Neural-network quantum state tomography. *Nat. Phys.*, 14(5):447, 2018.

[240] Giacomo Torlai and Roger G Melko. Latent space purification via neural density operators. *arXiv preprint arXiv:1801.09684*, 2018.

[241] I. Bengtsson and K. Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge Univ Pr, 2006.

[242] See the appendix section 'Comparison with approximate state reconstruction methods'.

[243] Hilary A Carteret. Noiseless quantum circuits for the peres separability criterion. *Phys. Rev. Lett.*, 94(4):040502, 2005.

[244] Jianming Cai and Wei Song. Novel schemes for directly measuring entanglement of general states. *Phys. Rev. Lett.*, 101(19):190503, 2008.

[245] Karol Bartkiewicz, Paweł Horodecki, Karel Lemr, Adam Miranowicz, and Karol Życzkowski. Method for universal detection of two-photon polarization entanglement. *Phys. Rev. A*, 91:032315, Mar 2015.

[246] Juan Carrasquilla and Roger G Melko. Machine learning phases of matter. *Nat. Phys.*, 2017.

[247] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[248] Alexander Hentschel and Barry C Sanders. Machine learning for precise quantum measurement. *Phys. Rev. Lett.*, 104(6):063603, 2010.

[249] Leonardo Banchi, Nicola Pancotti, and Sougato Bose. Quantum gate learning in qubit networks: Toffoli gate without time-dependent control. *npj Quantum Inf.*, 2:16019, 2016.

[250] GA Viano. Solution of the hausdorff moment problem by the use of pollaczek polynomials. *J. Math. Anal. Appl.*, 156(2):410–427, 1991.

[251] Swapan Rana. Negative eigenvalues of partial transposition of arbitrary bipartite states. *Phys. Rev. A*, 87(5):054301, 2013.

[252] Artur K Ekert, Carolina Moura Alves, Daniel KL Oi, Michał Horodecki, Paweł Horodecki, and Leong Chuan Kwek. Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.*, 88(21):217901, 2002.

[253] Michael Reck, Anton Zeilinger, Herbert J Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73(1):58, 1994.

[254] Dmitry A Abanin and Eugene Demler. Measuring entanglement entropy of a generic many-body system with a quantum switch. *Phys. Rev. Lett.*, 109(2):020504, 2012.

[255] Cristiano De Nobili, Andrea Coser, and Erik Tonni. Entanglement entropy and negativity of disjoint intervals in cft: some numerical extrapolations. *J. Stat. Mech: Theory Exp.*, 2015(6):P06021, 2015.

[256] Nello Cristianini and John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.

[257] Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(8):832–844, 1998.

[258] Raúl Rojas. *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.

[259] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Netw.*, 61:85–117, 2015.

[260] Peter Barmettler, Matthias Punk, Vladimir Gritsev, Eugene Demler, and Ehud Altman. Relaxation of antiferromagnetic order in spin-1/2 chains following a quantum quench. *Phys. Rev. Lett.*, 102(13):130603, 2009.

[261] Arun Nanduri, Hyungwon Kim, and David A Huse. Entanglement spreading in a many-body localized system. *Phys. Rev. B*, 90(6):064201, 2014.

[262] Sandu Popescu, Anthony J Short, and Andreas Winter. Entanglement and the foundations of statistical mechanics. *Nat. Phys.*, 2(11):754–758, 2006.

[263] Alioscia Hamma, Siddhartha Santra, and Paolo Zanardi. Quantum entanglement in random physical states. *Phys. Rev. Lett.*, 109(4):040502, 2012.

[264] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemp. Phys.*, 56(2):172–185, 2015.

[265] Stefan Trotzky, Patrick Cheinet, Simon Fölling, M Feld, Ute Schnorrberger, Ana Maria Rey, Anatoli Polkovnikov, EA Demler, MD Lukin, and I Bloch. Time-resolved observation and control of superexchange interactions with ultracold atoms in optical lattices. *Science*, 319(5861):295–299, 2008.

[266] Tatsuma Nishioka, Shinsei Ryu, and Tadashi Takayanagi. Holographic entanglement entropy: an overview. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504008, 2009.

[267] Filippo Caruso, Alex W Chin, Animesh Datta, Susana F Huelga, and Martin B Plenio. Entanglement and entangling power of the dynamics in light-harvesting complexes. *Phys. Rev. A*, 81(6):062346, 2010.

[268] Viktor Eisler and Zoltán Zimborás. Entanglement negativity in the harmonic chain out of equilibrium. *New J. Phys.*, 16(12):123020, 2014.

[269] Xueda Wen, Po-Yao Chang, and Shinsei Ryu. Entanglement negativity after a local quantum quench in conformal field theories. *Phys. Rev. B*, 92(7):075109, 2015.

[270] Nicholas E Sherman, Trithep Devakul, Matthew B Hastings, and Rajiv RP Singh. Nonzero-temperature entanglement negativity of quantum spin models: Area law, linked cluster expansions, and sudden death. *Physical Review E*, 93(2):022128, 2016.

[271] Abolfazl Bayat. Scaling of tripartite entanglement at impurity quantum phase transitions. *Phys. Rev. Lett.*, 118(3):036102, 2017.

[272] Cornelius Lanczos. *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*. United States Governm. Press Office Los Angeles, CA, 1950.

[273] Gene H Golub and Gérard Meurant. Matrices, moments and quadrature. *Pitman Research Notes in Mathematics Series*, pages 105–105, 1994.

[274] Shashanka Ubaru, Jie Chen, and Yousef Saad. Fast estimation of tr(f(a)) via stochastic lanczos quadrature. *SIAM Journal on Matrix Analysis and Applications*, 38(4):1075–1099, 2017.

[275] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77(8):1413, 1996.

[276] Insu Han, Dmitry Malioutov, Haim Avron, and Jinwoo Shin. Approximating the spectral sums of large-scale matrices using chebyshev approximations. *arXiv:1606.00942*, 2016.

[277] Lin Lin, Yousef Saad, and Chao Yang. Approximating spectral densities of large matrices. *SIAM review*, 58(1):34–65, 2016.

[278] Michael F Hutchinson. A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines. *Commun. Stat. Simul. Comput.*, 19(2):433–450, 1990.

[279] M Bellalij, L Reichel, Giuseppe Rodriguez, and H Sadok. Bounding matrix functionals via partial global block lanczos decomposition. *Applied Numerical Mathematics*, 94:127–139, 2015.

[280] Jacob Biamonte and Ville Bergholm. Tensor networks in a nutshell. *arXiv preprint arXiv:1708.00006*, 2017.

[281] Edo Liberty, Franco Woolfe, Per-Gunnar Martinsson, Vladimir Rokhlin, and Mark Tygert. Randomized algorithms for the low-rank approximation of matrices. *Proceedings of the National Academy of Sciences*, 104(51):20167–20172, 2007.

[282] Franco Woolfe, Edo Liberty, Vladimir Rokhlin, and Mark Tygert. A fast randomized algorithm for the approximation of matrices. *Applied and Computational Harmonic Analysis*, 25(3):335–366, 2008.

[283] Per-Gunnar Martinsson, Vladimir Rokhlin, and Mark Tygert. A randomized algorithm for the decomposition of matrices. *Applied and Computational Harmonic Analysis*, 30(1):47–68, 2011.

[284] Matthew B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

[285] Peter Pippan, Steven R White, and Hans Gerd Evertz. Efficient matrix-product state method for periodic boundary conditions. *Phys. Rev. B*, 81(8):081103, 2010.

[286] Paola Ruggiero, Vincenzo Alba, and Pasquale Calabrese. Entanglement negativity in random spin chains. *Phys. Rev. B*, 94(3):035152, 2016.

[287] Glen Bigan Mbeng, Vincenzo Alba, and Pasquale Calabrese. Negativity spectrum in 1d gapped phases of matter. *Journal of Physics A: Mathematical and Theoretical*, 50(19):194001, 2017.

[288] Udaysinh T Bhosale, Steven Tomsovic, and Arul Lakshminarayan. Entanglement between two subsystems, the wigner semicircle and extreme-value statistics. *Phys. Rev. A*, 85(6):062331, 2012.

[289] Yasuhiro Sekino and Leonard Susskind. Fast scramblers. *Journal of High Energy Physics*, 2008(10):065, 2008.

[290] Andrea Coser, Erik Tonni, and Pasquale Calabrese. Entanglement negativity after a global quantum quench. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(12):P12017, 2014.

[291] Haim Avron and Sivan Toledo. Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. *Journal of the ACM (JACM)*, 58(2):8, 2011.

[292] Moritz August, Mari Carmen Bañuls, and Thomas Huckle. On the approximation of functionals of very large hermitian matrices represented as matrix product operators. *Electronic Transactions on Numerical Analysis*, 46:215–232, 2017.

[293] Y-Y Shi, L-M Duan, and Guifre Vidal. Classical simulation of quantum many-body systems with a tree tensor network. *Phys. Rev. A*, 74(2):022320, 2006.

[294] Valentin Murg, Frank Verstraete, Örs Legeza, and Rein-hard M Noack. Simulating strongly correlated quantum systems with tree tensor networks. *Phys. Rev. B*, 82(20):205105, 2010.

[295] Guifre Vidal. Entanglement renormalization. *Phys. Rev. Lett.*, 99(22):220405, 2007.

[296] Guifré Vidal. Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.*, 101(11):110501, 2008.

[297] Igor L Markov and Yaoyun Shi. Simulating quantum computation by contracting tensor networks. *SIAM Journal on Computing*, 38(3):963–981, 2008.

[298] Edwin Pednault, John A Gunnels, Giacomo Nannicini, Lior Horesh, Thomas Magerlein, Edgar Solomonik, and Robert Wisnieff. Breaking the 49-qubit barrier in the simulation of quantum circuits. *arXiv preprint arXiv:1710.05867*, 2017.

[299] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, and Hartmut Neven. Simulation of low-depth quantum circuits as complex undirected graphical models. *arXiv preprint arXiv:1712.05384*, 2017.

[300] Jianxin Chen, Fang Zhang, Mingcheng Chen, Cupjin Huang, Michael Newman, and Yaoyun Shi. Classical simulation of intermediate-size quantum circuits. *arXiv preprint arXiv:1805.01450*, 2018.

[301] Igor L Markov, Aneeqa Fatima, Sergei V Isakov, and Sergio Boixo. Quantum supremacy is both closer and farther than it appears. *arXiv preprint arXiv:1807.10749*, 2018.

[302] *Guide to NumPy*. CreateSpace Independent Publishing Platform, USA, 2nd edition, 2015.

[303] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. [Online; accessed 17 August 2018].

[304] Siu Kwan Lam, Antoine Pitrou, and Stanley Seibert. Numba: A llvm-based python jit compiler. In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, page 7. ACM, 2015.

[305] Vicente Hernandez, Jose E Roman, and Vicente Vidal. Slepc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Transactions on Mathematical Software (TOMS)*, 31(3):351–362, 2005.

[306] Lisandro D Dalcin, Rodrigo R Paz, Pablo A Kler, and Alejandro Cosimo. Parallel distributed computing using python. *Advances in Water Resources*, 34(9):1124–1139, 2011.

[307] Lisandro Dalcín, Rodrigo Paz, and Mario Storti. Mpi for python. *Journal of Parallel and Distributed Computing*, 65(9):1108–1115, 2005.

[308] Daniel GA Smith and Johnnie Gray. opt_einsum - a python package for optimizing contraction order for einsum-like expressions, june 2018.

[309] Eugene F Dumitrescu, Allison L Fisher, Timothy D Goodrich, Travis S Humble, Blair D Sullivan, and Andrew L Wright. Benchmarking treewidth as a practical component of tensor-network–based quantum simulation. *arXiv preprint arXiv:1807.04599*, 2018.

[310] Vedika Khemani, Frank Pollmann, and SL Sondhi. Obtaining highly excited eigenstates of many-body localized hamiltonians by the density matrix renormalization group approach. *Physical review letters*, 116(24):247204, 2016.

[311] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS-W*, 2017.

[312] Andreas Sorge. pyfssa 0.7.6, December 2015.

[313] Jean-Paul Blaizot and Georges Ripka. *Quantum theory of finite systems*, volume 3. MIT press Cambridge, 1986.

[314] M Veldhorst, HGJ Eenink, CH Yang, and AS Dzurak. Silicon cmos architecture for a spin-based quantum computer. *arXiv preprint arXiv:1609.09700*, 2016.

[315] Lawrence R Mead and Nikos Papanicolaou. Maximum entropy in the problem of moments. *J. Math. Phys.*, 25(8):2404–2417, 1984.

[316] Viktor Eisler and Zoltán Zimborás. On the partial transpose of fermionic gaussian states. *New J. Phys.*, 17(5):053048, 2015.

[317] Guillaume Aubrun. Partial transposition of random states and non-centered semicircular distributions. *Random Matrices: Theory Appl.*, 1(02):1250001, 2012.

[318] Motohisa Fukuda and Piotr Śniady. Partial transpose of random quantum states: Exact formulas and meanders. *J. Math. Phys.*, 54(4):042202, 2013.

[319] James Bergstra, Daniel Yamins, and David Cox. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International Conference on Machine Learning*, pages 115–123, 2013.

[320] François Chollet et al. Keras. `https://github.com/fchollet/keras`, 2015.