Quantum information theory and many-body physics

UNIVERSITEIT VAN AMSTERDAM

QuSoft
Research Center for Quantum Software

CWI

Quantum information theory and many-body physics

Promotiecommissie

Promotores:	prof. dr. M. Walter	Ruhr-Universität Bochum
	prof. dr. E.M. Opdam	Universiteit van Amsterdam
Overige leden:	prof. dr. J.-S. Caux	Universiteit van Amsterdam
	prof. dr. M. Cheng	Universiteit van Amsterdam
	prof. dr. M. Christandl	Københavns Universitet
	prof. dr. R. T. König	Technische Universität München
	dr. M. Ozols	Universiteit van Amsterdam
	prof. dr. C. J. M. Schoutens	Universiteit van Amsterdam
	prof. dr. E. Verlinde	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

List of publications

This dissertation is based on the following papers. The authors of these papers have equally contributed to the obtained results.

[WW21c] ***Bosonic entanglement renormalization circuits from wavelet theory***
Freek Witteveen, Michael Walter
SciPost Physics **10**, 6 (2021)

[WSSW22] ***Quantum circuit approximations and entanglement renormalization for the Dirac field in 1+ 1 dimensions***
Freek Witteveen, Volkher Scholz, Brian Swingle, Michael Walter
Communications in Mathematical Physics **389** 1 (2022)

[RWW20] ***A converse to Lieb-Robinson bounds in one dimension using index theory***
Daniel Ranard, Michael Walter, Freek Witteveen
Accepted for publication in Annales Henri Poincaré

[CPWW] ***Random tensor networks with nontrivial link spectra***
Newton Cheng, Geoff Penington, Michael Walter, Freek Witteveen
In preparation

The author has also co-authored to the following papers in the course of his PhD, which are not included in this dissertation.

[GWSG20] ***Signal processing techniques for efficient compilation of controlled rotations in trapped ions***
Koen Groenland, Freek Witteveen, Kareljan Schoutens, Rene Gerritsma
New Journal of Physics **22** 6 (2020)

[WW21a] ***Hypergraph min-cuts from quantum entropies***
Michael Walter, Freek Witteveen
Journal of Mathematical Physics **62** 9 (2021)

You fling the book on the floor, you would hurl it out of the window, even out of the closed window, through the slats of the Venetian blinds; let them shred its incongruous quires, let sentences, words, morphemes, phonemes gush forth, beyond recomposition into discourse; through the panes, and if they are of unbreakable glass so much the better, hurl the book and reduce it to photons, undulatory vibrations, polarized spectra; through the wall, let the book crumble into molecules and atoms passing between atom and atom of the reinforced concrete, breaking up into electrons, neutrons, neutrinos, elementary particles more and more minute; through the telephone wires, let it be reduced to electronic impulses, into flow of information, shaken by redundancies and noises, and let it be degraded into a swirling entropy.

— Italo Calvino, If on a Winter's Night a Traveler

Contents

III   Entanglement in random tensor networks

Interactions between quantum information and many-body physics

The advent of computers has revolutionized the field of physics, in the first place by enabling the numerical simulation of complex physical models that were out of reach of theoretical understanding. Besides this, the *theory* of computation as well as the closely related field of *information theory* have had a profound impact on our understanding of nature, especially with regards to statistical mechanics and thermodynamics [Jay57, Lan91]. Conversely it has been fruitful to use insights and principles from physics in understanding the theory of computation [MM11, MM09] for instance in the use of models from statistical mechanics to understand approximation and learning algorithms.

Our best theories of small-scale fundamental physics are *quantum theories*, as developed from the early 20th century onwards. We have an excellent understanding of how to formulate such theories and their equations of motion (the *Schrödinger equation*). Nevertheless, actual computation of the results of these equations of motion can be challenging, even more so than for classical physics. This is especially the case for *quantum many-body systems*, which are quantum systems consisting of a large number of local degrees of freedom or particles. For instance, one can think of the electrons in a solid metal which can 'move around' and interact with one another, of the structure of a large molecule, or of the quantum particles described by the Standard Model. In such systems the number of parameters needed to describe the system (the wave function) typically scales exponentially with the number of particles. This poses fundamental computational problems. While surmountable in some cases by using appropriate approximations, there are many systems of interest which remain extremely challenging to simulate. This has led to the proposal of a *quantum computer*, a computer consisting of quantum particles and which has as its logical operations local quantum operations. The initial idea was that such a quantum computer could potentially simulate quantum systems more efficiently [F$^+$82], while later it was realized that quantum computers may have unrelated applications and provide algorithmic speed-ups for various problems (the most famous of which may be Shor's algorithm for period finding and prime factorization, and Grover's algorithm for black-box search [NC02]). Similarly, the fact that nature appears to be of a quantum nature has motivated the development of a theory of *quantum information* [Wat18, Wil13] which has fundamental differences with the classical theory of information.

If quantum computers and quantum communication are to be realized at large scale, will this again revolutionize physics? While there is strong evidence that the development of quantum computers would tremendously improve our abilities to simulate quantum systems, there are still many interesting questions to answer in this regard: which quantum systems can be simulated accurately and efficiently, and what are good principles for designing quantum simulation algorithms? In some ways, the *theory of quantum information and quantum computation* has already had a profound impact on quantum many-body physics, as it has been very fruitful to study many-body physics through the lens of quantum information theory, providing both new techniques and new questions to ask.

In the introduction to this dissertation we identify three broad themes in the interaction between quantum information theory and many-body physics:

   (i) The structure of many-body ground states

  (ii) Understanding unitary quantum dynamics

 (iii) The relation between quantum information and quantum gravity.

Each of these comprises a large body of research; in this dissertation we have tried to add a (minor) new insight to each. The aim of this introduction is to provide in a pedagogical manner a context for the results in this dissertation and, highlighting the guiding questions we have tried to answer. As opposed to most of the remainder of the dissertation, we will be rather informal. Each of the three parts of this dissertation provides an introduction where we focus further on the relevant objects of study.

We assume that the reader is familiar with the basic theory of quantum computation and quantum information theory, for instance at the level of standard works such as [NC02, Wat18] or the lecture notes [DW19, Pre98] A textbook providing an overview of quantum information theory and many-body physics is [ZCZ$^+$19].

1.1   Many-body ground states

Many quantum mechanical phenomena are determined by their low-energy physics. In this section we will describe the problem of finding the lowest energy states of a quantum mechanical system and how tools from quantum computation and quantum information theory are relevant to this problem. Quantum mechanical systems are described by Hamiltonians, which describe the interactions between different components of the system. Formally speaking, if $\mathcal{H}$ is the Hilbert space of quantum states, a Hamiltonian is a self-adjoint operator H acting on $\mathcal{H}$. The spectrum of H is interpreted as the energy levels of the system. A minimal assumption is that the spectrum of H is bounded from below, and in this case the lowest eigenvalue is the ground state energy of the system. Models of physical systems are typically such that the Hilbert space is decomposed into local Hilbert spaces. For instance, one can have a lattice spin system, where we have sites in a lattice Γ, and a local Hilbert space $\mathcal{H}_x = \mathbb{C}^{d_x}$ at each lattice site $x \in \Gamma$. For convenience assuming a finite lattice, the full Hilbert space is then given

by the tensor product[1]

$$\mathcal{H} = \bigotimes_{x \in \Gamma} \mathcal{H}_x. \tag{1.1}$$

If the system consists of a large number of subsystems, we speak of a *many-body* system and we will call the number of sites in Γ the system size. The Hamiltonian will typically be k-local for some k, meaning that

$$H = \sum_{X \subset \Gamma} H_X \tag{1.2}$$

where the sum is over subsets $X \subseteq \Gamma$ of size at most k where $H_X = I_{X^c} \otimes h_X$ is a Hermitian operator which acts only on the sites in X. Moreover, physical systems often have *geometrically local* Hamiltonians, meaning that sites which are far apart do not interact, or interact only weakly. This can be formalized by demanding that the H_X are zero for sets X which have diameter larger than some prescribed value, or by demanding that the size of the local terms $\|H_X\|$ decays with the size of the diameter of X.

Let us give two paradigmatic sources of examples.

(i) In *quantum chemistry* the most important Hamiltonians are those describing electrons in a molecule. The electrons have 2-local interactions with the nuclei and with each other. These models have limited geometric locality. In this case, we have a fermionic system, with degrees of freedom labelled by the orbitals of the atoms in the molecule.

(ii) Lattice spin systems modelling a *condensed matter system*. In this case we may take a regular lattice, and the Hamiltonian has geometric locality. A basic example is the one-dimensional Ising model

$$H = -J \sum_{n \in \mathbb{Z}} X_n X_{n+1} + h Z_n \tag{1.3}$$

where X and Z are the Pauli matrices and h and J are real parameters.

In this introduction we will focus on lattice spin systems, rather than problems from quantum chemistry. Such systems are not just important for condensed matter physics and material science, but can also be used as a discretization for quantum field theories. For instance, QCD (which models quarks in the standard model) is often numerically studied by approximating it by a model of lattice fermions [Smi02]. We order the spectrum of the Hamiltonian H as $E_0 \leq E_1 \leq \dots$. Then a *ground state* $|\psi\rangle$ is an eigenvector with minimal eigenvalue $H|\psi\rangle = E_0|\psi\rangle$, where E_0 is the ground state energy. The ground state energy could be degenerate (but in the following we assume for convenience that it is unique). The difference $\Delta = E_1 - E_0$ is the *ground state energy gap*. We say that a family of Hamiltonians of increasing system sizes $|\Gamma|$ is *gapped* if Δ is lower bounded by a constant. From the perspective of statistical physics, the state $|\psi\rangle$ describes the system at zero temperature. For low temperature physics, it is therefore crucial to understand the ground state and the ground state energy gap. For example, the Ising model (which can be analytically solved) has two parameter regimes with

---

[1]Another possibility is that one has indistinguishable particles with bosonic or fermionic statistics, in which case the notion of locality is different (which we will ignore for the purpose of this introduction).

different behaviour and a *phase transition* in between. However, for other Hamiltonians finding ground states or understanding the phase diagram can be a difficult problem, analytically as well as numerically. The Hilbert space has exponential dimension in the system size, so naive linear algebra methods are typically limited to small system sizes. Many useful classical algorithms have been designed to nevertheless simulate many-body quantum systems, but numerical simulation of strongly interacting quantum systems remains a highly challenging problem. For instance, there are powerful methods based on quantum Markov Chain Monte Carlo sampling; however, in strongly interaction systems the so-called sign problem can make this method unreliable [TW05]. On the other hand, one could use quantum computers for the simulation of quantum ground state properties. This seems to be a promising direction, both for quantum chemistry and condensed matter or field theories [CRO+19, Pre18]. However, as we will see in a moment, finding ground states is not an easy problem, even for a quantum computer.

1.1.1   Quantum computation and ground states

What can we say in general about ground states of physical (condensed matter) systems, that is, of geometrically local Hamiltonians? It can be proven that this is a hard question in a precise sense. Recall that QMA is, informally speaking, the class of decision problems for which a quantum computer can efficiently verify positive instances, given access to a (quantum) witness. In other words, it is the quantum generalization of the class NP, and QMA-complete problems are widely believed to be hard to solve for quantum computers. A seminal result in quantum complexity theory is that for general local Hamiltonians it is QMA-hard to find its ground state energy [KSVV02, KKR06]. To be more precise, let H be a 2-local Hamiltonian on n sites, and assume that we know that E_0 is either smaller than 0, or $E_0 = \Omega(\frac{1}{\text{poly}(n)})$. Then it is QMA-hard to decide which of these two is the case. This remains true in various situations where one restricts the class of Hamiltonians, for instance to nearest-neighbour Hamiltonians on a two-dimensional lattice of qubits [OT05] as well as various other physically realistic models [SV09, CGW14]. This result strongly indicates limits to the application of quantum computers to simulate ground state physics. On the other hand, it does not exclude the possibility of efficient quantum algorithms for finding ground states for more restricted classes of physical systems. There are systems which can be solved exactly *analytically* (such as the one-dimensional Ising model in Eq. (1.3)), there are systems which can be solved numerically by a *classical* computer (such as one-dimensional systems with a constant gap, as we will see in Section 1.1.2) and, as we saw above, there are systems which are likely hard for a *quantum* computer. An important open question is whether there are classes of interesting and physically relevant quantum systems for which one can approximate ground state physics more efficiently with a quantum computer than with a classical computer. There are various 'generic' algorithms which find approximations to the ground state (all of which have limitations, in line with the hardness of the problem). Here we discuss two of the most well-known approaches.

(i) The most well-established algorithms are based on a version of *phase estimation*. Suppose that U is a unitary, and $|\psi\rangle$ is an eigenvector with eigenvalue $e^{i\phi}$. Phase estimation is a quantum algorithm which, given access to an oracle implementing the unitary U, and an initial state $|\chi\rangle$ which has nonzero overlap with $|\psi\rangle$, allows

you to estimate ϕ. One may now, for instance, implement (an approximation of) the unitary $U = e^{iH}$ in some way on a quantum computer, which will have an eigenvalue e^{iE_0}, and apply phase estimation to U to find E_0 [AL99]. While a powerful generic algorithm and running in polynomial time in the number of qubits n, it requires one to have a sufficiently good starting state to have a reasonable probability of success (for instance, it may be hard to find an initial state $|\chi\rangle$ which has $\Omega(\frac{1}{\text{poly}(n)})$ overlap with $|\psi\rangle$).

(ii) A second, more heuristic approach is the class of *variational quantum eigensolvers* (VQE). Here, one considers a class of ansatz states which can be efficiently prepared on a quantum computer, for instance a collection of short depth local circuits with some free continuous parameters in the circuit gates. One then uses a hybrid quantum-classical algorithm to optimize the parameters in this ansatz class: one uses the quantum computer to compute energy expectation values and gradients with respect to the parameters, and a classical algorithm (for instance gradient descent) to find a (local) minimum in the parameter space. This is a classical optimization problem, using a quantum computer only for a subroutine which evaluates energies (and energy gradients), which means that one needs a relatively small number of qubits and relatively short coherence times. This provides the main advantage of this approach: it may be relatively amenable to near-term noisy devices, which are not yet able to perform fully fault-tolerant quantum computation. The main downside is that it is hard to get (rigorous) guarantees for good approximations and we do not yet know whether using such methods will be accurate in practice. First of all, one has to choose an ansatz class which can indeed approximate the ground state to some accuracy, which is often not clear, and secondly, the optimization problem may be non-convex and the minimization algorithms may not converge. Finally, when using noisy devices, one needs to understand the dependence of the noise on the accuracy of the approximation. For a review and references, see [CAB$^+$21].

In conclusion, an important question with respect to the use of quantum computers for ground state finding is to determine which systems are amenable to certain types of simulations and how we can design useful algorithms. A useful lesson seems to be that in the design of quantum algorithms (and this is especially true for VQE) we should try to use physical insights and principles. See [BBMC20] for a good overview of the prospects for quantum computing in many-body physics.

1.1.2 Tensor networks

Quantum information theory has had a profound impact on *classical* simulation methods for ground states as well, especially in the form of tensor network methods. A typical property of gapped geometrically local Hamiltonians on a lattice is that they satisfy an *area law* for the entanglement entropy [ECP10]. Here by an area law we mean that if we have a domain X in the lattice (where $|X|$ is much smaller than the system size) then the entanglement entropy $H(X)_\psi$ scales approximately as $|\delta X|$, the size of the boundary of X. A closely related fact is that we expect exponentially decaying correlations between different sites. That is, if O_x and O_y are hermitian operators on sites x and y, then we

expect

$$|\text{tr}[O_x O_y \psi] - \text{tr}[O_x \psi]\,\text{tr}[O_y \psi]| = \mathcal{O}(e^{-\xi d(x,y)}) \tag{1.4}$$

where $d(x,y)$ is the distance between x and y and ξ is the correlation length. This has motivated the introduction of *tensor network states*. A good introduction can be found in [Orú14]. We will sketch the basic concepts. Suppose we have a quantum state $|\psi\rangle$ on a tensor product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$, where $\mathcal{H}_k = \mathbb{C}^{d_k}$, then we may expand $|\psi\rangle$ in a product basis

$$|\psi\rangle = \sum_{\{i_k\}} \psi_{i_1 i_2 \ldots i_n} |i_1\rangle \otimes |i_2\rangle \otimes \ldots \otimes |i_n\rangle.$$

The collection of numbers $\psi_{i_1 i_2 \ldots i_n}$ defines a tensor. Of course, in general, the size of this collection is exponential in n. Tensor networks provide a method to parametrize a relevant subset of tensors in an efficient manner, by 'breaking up ψ in smaller tensors'. Given two tensors A and B with coefficients $A_{i_1,i_2,\ldots,i_n}$ and $B_{j_1,j_2,\ldots,j_m}$ we may for instance contract A and B along the first indices (provided the corresponding dimensions are equal) to get a tensor with coefficients

$$\sum_{i_1} A_{i_1,i_2,\ldots,i_n} B_{i_1,j_2,\ldots,j_m}$$

Now, if we are given a collection of tensors, then we may contract indices along a graph which is defined by letting the tensors correspond to vertices with a number of dangling half-edges corresponding to the number of indices the tensor has; we then indicate which indices are contracted by connecting half-edges to form an edge in the graph. The resulting tensor has uncontracted indices on all unconnected half-edges. The dimensions along the contracted edges are called *bond dimensions*, while the dimensions along the uncontracted edges are the *physical dimensions*. See Fig. 1.1 for a simple illustration. This notation is consistent with the usual diagrammatic notation for quantum circuits (interpreting the gates as tensors). An equivalent perspective on this contraction scheme is that one places maximally entangled states along the edges of the graph and projects the states defined by the local tensors onto these maximally entangled states. For this reason, such tensor network states are also known as *projected maximally entangled pair states* (PEPS). The key idea behind the use of PEPS is that if you know that the state you are interested in has some restrictions on its entanglement structure, there may be a natural tensor network ansatz which approximates the state. An example of a two-dimensional PEPS tensor network is given in Fig. 1.2. In general, if we have a tensor network state $\rho = |\psi\rangle\langle\psi|$, and we let A denote a subset of the dangling half-edges, then it is easy to see that an upper bound to the rank of ρ_A is given by $D^{|\gamma_A|}$, where $|\gamma_A|$ is the number of bond edges we have to cut in order to disconnect A from its complement A^c in the tensor network graph, and where D is the bond dimension. It follows that the entanglement entropy $H(\rho)_A$ is upper bounded by $|\gamma_A|\log(D)$. From Fig. 1.2 it is clear that on a lattice this upper bound matches an area law.

Depending on the graph and the bond dimensions, a tensor network representation can be much more efficient than simply encoding the full state as a tensor. This can be exploited numerically by performing a variational optimization over this class of tensor network states to find approximations to the ground state.
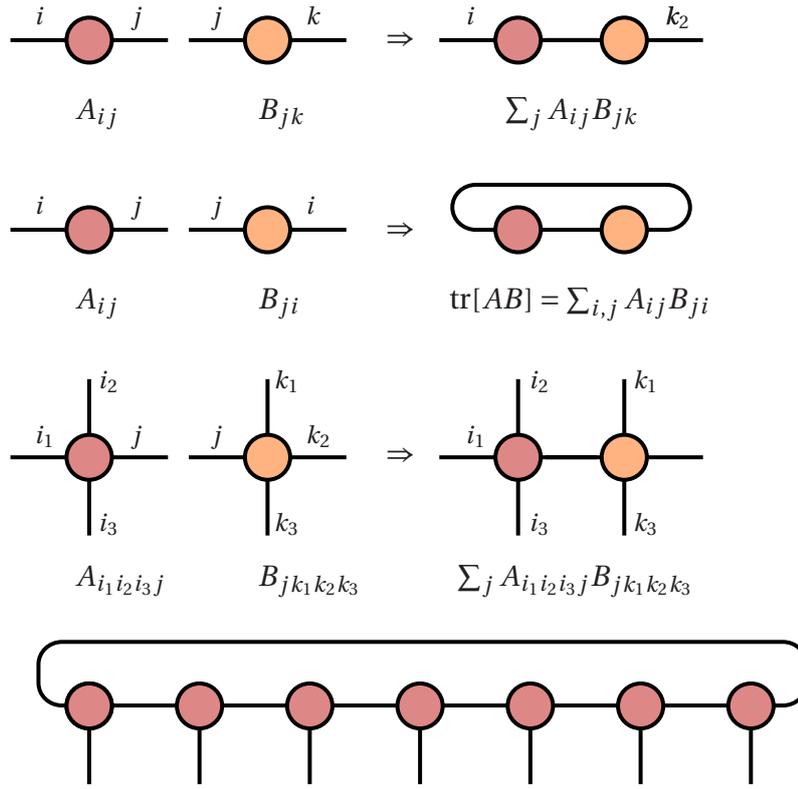
Figure 1.1: Examples of tensor network contractions, in particular corresponding to matrix multiplication, the trace of a product of matrices, and in the bottom row, a matrix product state.

To make this concrete, we will briefly discuss one of the most powerful examples, which is the class of *matrix product states* (MPS), which is the one-dimensional version of PEPS. In this case we consider a one-dimensional chain of N qudits. For convenience of notation we assume periodic boundary conditions. Fix a local physical dimension d and a bond dimension D. Then an MPS state is defined by N tensors of size $d \times D \times D$, which we write as $A^{(n)}_{ijk}$ for $n = 1, \dots, N$ (so $i = 0, \dots, d-1$ and $j, k = 0, \dots, D-1$). Denote by $A^{(n)}_i$ the $D \times D$ matrix with entries $\{A^{(n)}_{ijk}\}_{j,k}$. Then the associated MPS state is a state of N qudits, defined by the following product of matrices

$$|\psi\rangle = \sum_{\{i_k\}} \text{tr}[A^{(1)}_{i_1} A^{(2)}_{i_2} \dots A^{(N)}_{i_N}] |i_1 i_2 \dots i_N\rangle.$$

The associated tensor network graph is illustrated in Fig. 1.1. For MPS, if we let $A \subseteq [N]$ be an interval, then the entanglement entropy $H(A)_\rho$ for $\rho = |\psi\rangle\langle\psi|$ is upper bounded by $2\log(D)$ (since we only need to cut two edges to separate A from its complement) which again is consistent with an area law, since A has a constant size boundary, no matter the length of the interval! We would like to emphasize that the total number of parameters in the description of the MPS state $|\psi\rangle$ is NdD^2, so for fixed D this number of parameters is *linear* in the system size rather than exponential. If one increases the bond dimension D to a size exponential in N one can write any quantum state as an MPS (but one is usually interested in the regime where D is constant or polynomial in N). For one-dimensional spin systems, the space of ground states of gapped Hamiltonians
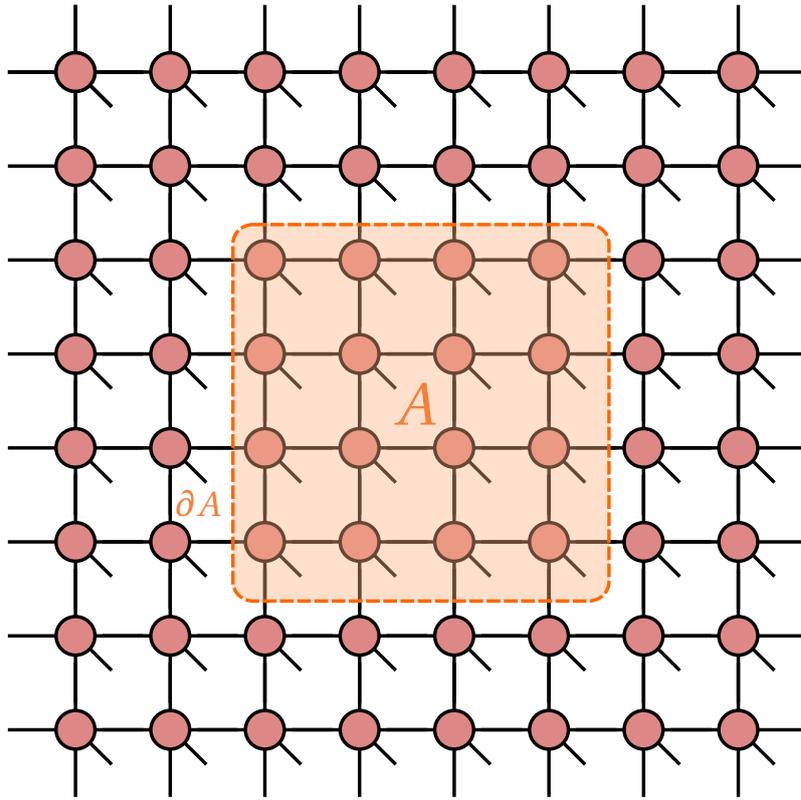
Figure 1.2: A PEPS tensor network on a two-dimensional lattice (with open boundary conditions). We see that if we take a subsystem A, we can separate A from A^c by cutting a number of bonds proportional to the boundary size $|\partial A|$, which matches area law behaviour of the entanglement entropy.

is well understood. If H is gapped, it is known that the ground state satisfies both an area law and exponential decay of correlations. One can approximate the ground state to precision ε by an MPS state with bond dimension $D = \mathcal{O}(\mathrm{poly}(N, \frac{1}{\varepsilon}))$ [Has07], see also [DB19] for approximations with constant bond dimension which are locally accurate. Moreover, such an approximation can be found in $\mathcal{O}(\mathrm{poly}(N, \frac{1}{\varepsilon}))$ by a classical computer [LVV15][2]. From this we conclude that, while in general the problem of finding ground states of local Hamiltonians is QMA-hard, when we restrict to one-dimensional gapped Hamiltonians the problem can be solved in polynomial time on a classical computer.

In higher spatial dimensions PEPS tensor networks (as in Fig. 1.2) can also be used to approximate ground states of Hamiltonians. Nevertheless, this poses various challenges. On the theoretical side, there are no rigorous approximation results available and there are various open questions with regards to which states can be accurately represented by PEPS [CGRPG19]. Applying numerical techniques is also much more challenging. To compare with MPS, if we would like to contract an MPS state and, for instance, compute the value of a particular coefficient of the resulting tensor, we simply have to

---

[2]Historically, the development of tensor networks started with the *density matrix renormalization group* (DMRG) algorithm [Whi92] which, in hindsight, was an algorithm for finding MPS approximations to ground states [Sch11]. While yielding excellent results in practice, there is no rigorous proof for the correctness of DMRG.
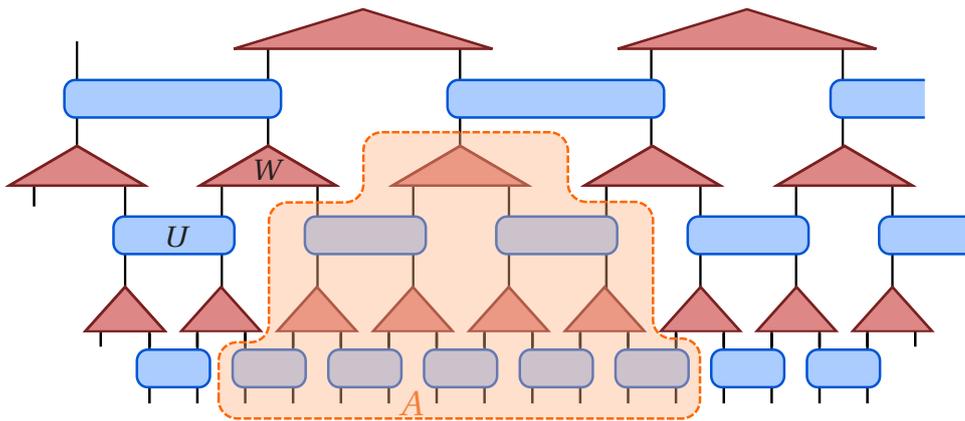
Figure 1.3: A MERA tensor network. The blue tensors are unitaries U, whereas the red tensors are isometries W. Each layer implements a real-space renormalization by a factor of two. If we want to cut the tensor network to separate a subsystem A we can do so by cutting a number of bonds logarithmic in $|A|$, which allows for a logarithmic correction to the one-dimensional area law. If the tensor network describes a state which is translation invariant and scale invariant, all the unitaries may be taken to be a fixed unitary, and all the isometries may be taken as a fixed isometry.

multiply matrices. For general PEPS the contraction of the tensor network state can be much more complicated (in general, computing the contracted PEPS tensor from a collection of N tensors is #P-hard [SWVC07, HHEG20]). Nevertheless, the use of PEPS tensor network states has become a crucial tool both for theoretical developments in condensed matter physics (for instance to obtain better understanding of topological order, see [CPGSV21] for an overview and references) and for numerical computations in strongly interacting two-dimensional systems.

1.1.3 Entanglement renormalization

When we discussed MPS we focussed on gapped systems. In condensed matter systems gapless systems typically occur at a phase transition (a critical point) between two different gapped phases. Such systems have polynomially decaying correlation functions, and in one spatial dimension they have a logarithmic correction to the area law for the entanglement entropy, that is, the entanglement entropy of an interval A scales with the logarithm of the size of the interval $H(A)_\rho \sim \log(|A|)$. One can still use MPS states with bond dimension increasing with the system size to accommodate this entropy scaling [VC06]. Another approach is to use that critical points typically satisfy a form of *scale invariance* and their continuum limit is described a conformal field theory (CFT) [FMS12]. Such systems are described by a fixed point of the renormalization group. The idea of the *multi-scale entanglement renormalization ansatz* (MERA) is to implement the renormalization group by a quantum circuit, which can also be seen as a tensor network [Vid08, Vid07]. This structure is illustrated in Fig. 1.3. We will introduce MERA in more detail in Chapter 2.

As we saw there is a well-developed theory for MPS states in one-dimensional systems. For MERA, this theory is much less developed and our understanding is largely based on numerical experiments. A second motivation to obtain deeper understanding

of MERA comes from quantum computing. As we saw in Section 1.1, future applications of quantum computers to simulation of ground state physics might benefit from carefully considering physical principles in designing algorithms. Real-space renormalization could be such a foundational principle, again especially for critical systems. As MERA is a tensor network with a unitary structure it is well suited to use in a quantum computer for VQE algorithms, an idea which has been advocated in [KS17a]. Thus, there is s strong twofold motivation for developing theory and design principles for MERA, leading to one of the guiding questions for Part I of this dissertation:

> Can we *prove* that MERA is an accurate ansatz class for critical quantum systems?

In this dissertation we work out an idea from [EW16, HSW$^+$18]: for free theories one can analytically construct MERA tensor networks by second quantization of discrete *wavelet transforms.* A wavelet transform is a localized analog of the Fourier transform, and has a similar dyadic recursive structure as MERA. We show that the connection between MERA and wavelets can be extended to a wide class of free bosonic systems.

Numerically, it can be seen that many properties of the limiting CFT of a critical lattice system are approximately encoded in a MERA tensor network for the lattice model. This leads to a second question we explore in Part I:

> What is the relation between entanglement renormalization and quantum field theory?

We show that there is a close relation to the quantum field theory limit of the lattice models we consider and the *continuous wavelet transform,* which is itself a continuous limit of the discrete wavelet transform. We use this correspondence to show that MERA can be used to approximate correlation functions of these field theories, providing a new approach to quantum simulation of quantum field theories. While this work is for a class of free theories (which can therefore be simulated efficiently in any case), we hope this provides a foundation for further work on understanding entanglement renormalization for other classes of critical quantum systems. In Part I we start with an introduction and review of entanglement renormalization in Chapter 2 and we review relevant aspects of wavelet theory in Chapter 3. Then, in Chapter 4 we come to our first main result: a construction of fermionic entanglement renormalization circuits and an analysis of their accuracy for a fermionic quantum field theory. Our second main result is in Chapter 5, where we work out an analogous construction for bosonic systems.

1.2   Unitary quantum dynamics

In Section 1.1 we discussed the problem of understanding the structure of ground states in many-body quantum physics, and using quantum and classical computers to simulate these states. We now move to the topic of understanding the *dynamics* of quantum systems. Given a Hamiltonian H on a Hilbert space $\mathcal{H}$, and an initial state $|\psi\rangle$, after time t the system will have evolved (assuming we have a closed system which is

completely described by H) to a state $|\psi(t)\rangle = \exp(-iHt)|\psi\rangle = U_t|\psi\rangle$ where U_t is the unitary $\exp(-iHt)$. Or, if we start with a density matrix ρ, the time evolved density matrix is given by $U_t\rho U_t^*$

More generally, if $H(t)$ is time dependent, the state evolves along the Schödinger equation

$$\partial_t|\psi(t)\rangle = -iH(t)|\psi(t)\rangle.$$

with initial condition $|\psi(0)\rangle = |\psi\rangle$. The solution to the Schödinger equation defines unitary operators U_t such that $|\psi(t)\rangle = U_t|\psi\rangle$ for any initial state $|\psi\rangle$. The above describes time evolution in the *Schrödinger picture*, where one keeps track of how the quantum state changes. Equivalently, one can keep the initial state fixed, and instead evolve observables. Recall that if O is some self-adjoint operator on $\mathcal{H}$ then the expected value of O in the state ρ is given by $\mathrm{tr}[\rho O]$. If we time evolve

$$\mathrm{tr}[\rho(t)O] = \mathrm{tr}[U_t\rho U_t^* O] = \mathrm{tr}[\rho U_t^* O U_t] = \mathrm{tr}[\rho O(t)]$$

where we defined the time evolved observable as $O(t) = U_t^* O U_t$. This is called the *Heisenberg picture* and for technical reasons this perspective can be more convenient for infinite dimensional systems (which we encounter in Part II).

1.2.1 Lieb-Robinson bounds

In Section 1.1 we saw that if we consider geometrically local Hamiltonians, the ground states have certain restrictions (such as an area law). What are the implications of the geometrically local nature of Hamiltonians to time evolution? Intuitively, if we have a system with only local interactions, information will travel through the system at most at some finite speed. This intuition is correct and is quantified by the so-called Lieb-Robinson bounds [LR72]. We study this question in the Heisenberg picture. Let H be a geometrically local Hamiltonian on a lattice Γ. Consider an operator O_X which has support on a finite subset of sites $X \subset \Gamma$ (that is, O_X is a tensor product of some operator on X and the identity operator on $\Gamma \setminus X$). Then, we claim that after time evolution along H for time t, $O_X(t)$ is approximately supported in a ball of some finite radius around X. Equivalently, $O_X(t)$ approximately commutes with all operators which are outside this ball. The precise formulation is given by the Lieb-Robinson bounds, which state that for any O_Y supported on sites $Y \subseteq \Gamma$ it holds that

$$\|[O_X(t), O_Y]\| \le Ce^{-\alpha(d(X,Y)-vt)}\|O_X\|\|O_Y\| \tag{1.5}$$

where α, v and C are constants (which depend on Γ and H and where C may depend on the size of the sets X and Y) and $d(X,Y)$ is the distance between the sets X and Y. From Eq. (1.5) we see that when Y has distance from X greater than vt, the commutator decreases exponentially, so $O_X(t)$ is approximately supported on a ball of radius of the order vt. The constant v is called the Lieb-Robinson velocity and bounds the information propagation speed.

It may be instructive to compare these results with the situation for a quantum field theory. There, one assumes a different form of locality which implies a strict lightcone of influence. That is, v is analogous to the speed of light c, but in a quantum field theory $[O_X(t), O_Y] = 0$ if the distance between X and Y is greater than ct. We

have presented a simple version of the Lieb-Robinson bounds, but there are many variations, for instance for interactions which are not strictly local but have (exponential or polynomial) decay with the interaction distance, see [NSY19] for a general discussion of Lieb-Robinson bounds. We provide further detail in Section 9.1. The Lieb-Robinson bounds show that in quantum lattice systems, while there is no fundamental speed limit, an approximate speed limit nevertheless arises (as it should, as one expects that the continuum limit of the spin system can be approximated by a quantum field theory).

The Lieb-Robinson bounds are a fundamental result, and have found various applications, amongst which a rigorous proof for the decay of correlations as in Eq. (1.4) for gapped systems [HK06].

1.2.2   Quantum simulation of dynamics

The possibility of simulating unitary quantum dynamics given a description of the Hamiltonian is potentially one of the most important applications of a quantum computer. It was in fact one of the original motivations for Feynman to propose the concept of quantum computation [F$^+$82]. It is crucial for a good understanding of quantum many-body systems, and moreover the problem is BQP-complete, and therefore likely to be hard in general on a classical computer. The difference between classical and quantum computers is therefore more clearly visible for this problem than for problems involving ground states: while we do not know of any subexponential classical algorithms to simulate quantum dynamics, a quantum computer is (almost by definition) well-suited to perform such simulations. In our description of using phase estimation for ground state finding we needed to implement time evolutions along some Hamiltonian as well, and therefore these algorithms also benefit from accurate and fast quantum algorithms for Hamiltonian evolution.

There are two main approaches to simulation of evolution along a local Hamiltonian on a quantum computer.

(i) The first method is known as *Trotterization* and was first suggested as a quantum algorithm in [Llo96]. The basic idea is that if we are given a local Hamiltonian H as in Eq. (1.2) where the local terms $\|H_X\|$ have bounded norm. Suppose that we can write $H = H_1 + H_2$ where H_1 and H_2 consist of local terms which do not overlap. For example, for the Ising Hamiltonian in Eq. (1.3) we could take H_1 to be the sum over all terms with n even and H_2 the sum over all terms with n odd. By the Baker-Campbell-Hausdorff we have the Suzuki-Trotter expansion given by $e^{iHt} = (e^{iH_1 t/N} e^{iH_2 t/N})^N + \mathcal{O}(\frac{t^2}{N})$. If we take $N = \mathcal{O}(\frac{t^2}{\varepsilon})$ we thus obtain an $\mathcal{O}(\varepsilon)$-approximation to the time evolution. Since H_1 and H_2 are sums of disjoint local terms we can write $e^{iH_1 t/n}$ and $e^{iH_1 t/n}$ as a layer of quantum gates. This yields a quantum circuit of depth $N = \mathcal{O}(\frac{t^2}{\varepsilon})$ which approximates the unitary time evolution. There are many refinements and generalizations of this procedure, in particular for Hamiltonians which need not be geometrically local, and approaches involving higher order commutators in the Baker-Campbell-Hausdorff formula leading to better scaling (for instance, a (near) linear scaling in t is desirable). See [CST$^+$21] for state of the art results on Trotterization.

(ii) A second approach uses Lieb-Robinson bounds to approximate dynamics for a geometrically local Hamiltonian [HHKL18, TGS$^+$19]. In this approach one uses that

if A, B and C are lattice subsystems with $A \cup B \cup C = \Gamma$, and we write H_Y, for the restrictions of the Hamiltonian to a subset $Y \subseteq \Gamma$ (that is, it consists of all terms which have support inside Y), then we can approximate $e^{-iHt} \approx e^{-iH_{A \cup B} t} e^{iH_B t} e^{-iH_{B \cup C} t}$ if we choose the regions such that A and C are separated by a sufficiently large distance. The approximation error is then bounded by the Lieb-Robinson bounds. This decomposition can be used in an iterative fashion to decompose e^{-iHt} into a quantum circuit.

There are also various other approaches, for instance methods based on qubitization [LC19] and quantum signal processing [LC17].

1.2.3 Quantum cellular automata

A useful model to study locality in quantum dynamics is that of a *quantum cellular automaton* (QCA). A QCA of radius R is a unitary which maps any local operator O_X supported on a set X to an operator supported on sites within radius R of X. Straightforward examples are local quantum circuits and translations. In one spatial dimension, QCAs are completely classified [GNVW12]. Using an index which measures an information flow it was shown that any such QCA can be written as a composition of a local circuit and a shift operator. Here the index can be thought of as an obstruction to writing the QCA as a circuit. As we saw that physical dynamics typically only preserve locality *approximately*, it is interesting to generalize the notion of a QCA to a version with approximate locality. In Part II we introduce this generalization as *approximately locality preserving unitaries* (ALPUs). As for QCAs, there is no continuous time evolution, but rather a single time step. The analog of quantum circuits is formed by time evolutions along time-dependent Hamiltonians. This leads to the following natural question:

> Given a unitary satisfying Lieb-Robinson bounds (i.e. an ALPU), can it be generated by some time-dependent local Hamiltonian? If not, what are the obstructions?

We show that in one spatial dimension, the index theory of [GNVW12] can be extended to ALPUs, leading to a similar classification, with Hamiltonian evolutions being the analog of quantum circuits. In particular, this gives a criterion for when an ALPU can be written as a Hamiltonian evolution.

As we saw, Hamiltonian evolutions can be approximated to arbitrary accuracy by quantum circuits. This raises a second fundamental question:

> Can any ALPU be approximated to arbitrary accuracy by a QCA?

Since the ALPU consists of a single time step and there is no Hamiltonian, Trotterization is not applicable. The Lieb-Robinson bounds based methods for simulation are already closer in spirit, but also in that case one fundamentally uses the Hamiltonian. However, we are nevertheless able to affirmatively answer this question for one-dimensional systems.

In Chapter 6 we introduce the theory of QCAs. In Chapter 7 we review elements of the theory of operator algebras and a result on perturbations of algebras which is

the key technical ingredient for our results. Next, in Chapter 8 we review the index theory of [GNVW12] and give a new expression for this index in terms of the mutual information. Chapter 9 contains our main results: the approximation of ALPUs by QCAs and an extension of the index theory to ALPUs.

1.3   Quantum information and quantum gravity

A third domain in which there is a lively dialogue between quantum information theory and many-body physics is in the study of quantum gravity. Quantum gravity poses one of the deepest open problems in theoretical physics. Of the fundamental physical forces known to us it is the only one which has not been incorporated in the Standard Model of particle physics. In almost all practical situations gravity is a very weak force and can be treated semi-classically. Nevertheless, understanding quantum gravity is required to understand some of the most fascinating phenomena in the universe, such as the Big Bang and black holes, both of which involve situations where gravity is a strong force compared to the length scales at which quantum effects become relevant, and a fully quantum theory of gravity is necessary. Naive approaches to formulating a theory of quantum gravity immediately run into problems: quantizations of general relativity are in general not renormalizable, and hence do not provide a well-defined quantum theory. The fact that gravity is such a weak force also implies that it is extremely challenging to study quantum gravity empirically, as one would need access to energy scales which are far out of reach for current experiments. One particular physical phenomenon where quantum gravity plays a crucial rol is in black holes. While classically, any observer passing the black hole horizon will never be able to return, it has been shown that black holes are radiating objects quantumly, emitting *Hawking radiation* [Haw75]. However, it appears that this radiation is completely thermal, containing no information about the mass that collapsed to form the black hole. This poses a puzzle: the black hole can in fact completely evaporate, but it seems there is *information loss* which is incompatible with global unitary evolution of the universe.

Below we will discuss one specific approach to quantum gravity, which is strongly motivated by understanding black hole physics.

1.3.1   Entropies in holography

An important observation from the study of black holes and Hawking radiation is that black holes appear to have an entropy which scales with the surface area of the black hole (as can be deduced from the temperature of the Hawking radiation) and has later been confirmed by string theory calculations [Haw76, Bek20, SV96]. This is rather remarkable: naively one expects that the entropy of matter scales with volume rather than area. This has led to the proposal that quantum gravity could be of a *holographic* nature: if we consider some volume of four-dimensional space-time, the full information content of this region can be described by information on the three-dimensional boundary, or in other words, the four space-time dimensions we observe are really an emergent phenomenon of a theory which is fundamentally three-dimensional. This idea has found considerable theoretical support from the discovery of the AdS/CFT correspondence [Mal99], in which it can be shown that various string theories on a $d+1$-dimensional Anti-de Sitter (bulk) space are equivalent to a nongravitational conformal
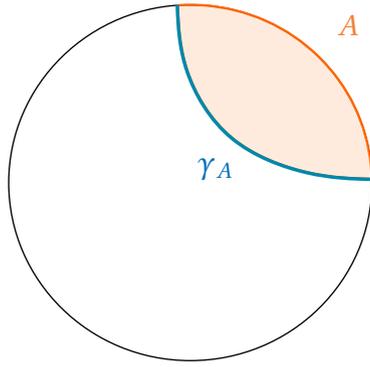
(a) Illustration of the RT-formula Eq. (1.6) for a 2-dimensional spatial bulk slice. Note that one should think of this disk as a hyperbolic disk, so γ_A is a geodesic. The orange shaded region is the entanglement wedge for the boundary region A.
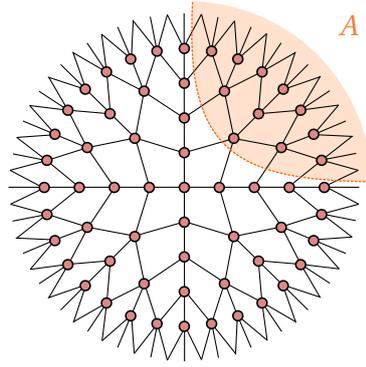
(b) A tensor network toy model for a holographic state. The physical Hilbert space is parametrized by the dangling half-edges at the 'boundary', and the internal nodes can be thought of as the 'bulk'. The graph structure can be taken to be a discretization of a hyperbolic space, and minimal cuts in the graph correspond to minimal surfaces.

field theory (CFT) on a d-dimensional (boundary) space. Here the AdS space is a type of space which has negative curvature (as opposed to the observed universe) and has the property that light rays reach its conformal boundary in finite time. A useful way of thinking about this property is that AdS is gravity 'confined in a box' to make the analysis easier. Intriguingly, in the AdS/CFT correspondence, weak and strong interaction are reversed: a strongly coupled boundary CFT corresponds to bulk gravity with weak (semi-classical) gravity. This means that AdS/CFT is not only an example of a well-defined theory of quantum gravity, but also a powerful tool in the study of certain strongly coupled quantum field theories. See [AE15] for an introduction and applications of holography. Quantum information theory has taken a centre stage in understanding how the bulk space-time arises from the boundary theory. One crucial insight which ties quantum information theory to holographic quantum gravity is the *Ryu-Takayanagi (RT) formula* [RT06a, RT06b]. This formula states that if one takes a state ρ which is dual to a stationary bulk space-time, and we consider a subregion A of a spatial slice of the boundary theory (which is the boundary of a bulk spatial slice), then the entanglement entropy $H(A)_\rho$ can be computed (on suitable regularization) as

$$H(A)_\rho = \min \frac{\text{Area}(\gamma_A)}{4G_N} + ... \tag{1.6}$$

where we consider the minimization over the areas of surfaces γ_A in a spatial slice of the bulk space-time which separate A from its complement, as in Fig. 1.4a. In this formula G_N is the Newton constant and the corrections to the area term are quantum corrections which are small for large G_N. A further question is what 'information' about the bulk can be reconstructed from ρ_A. Since the bulk and boundary theories are equivalent, one can reconstruct the bulk state from knowing the full boundary state. If we restrict to the reduced state ρ_A it turns out we can 'reconstruct' the part of space which is bounded by the minimal surface in Eq. (1.6). This region is called the *entanglement wedge*, and the corresponding reconstruction procedure is called *entanglement wedge*

reconstruction. See for instance [CMMN19] for a quantum information perspective on entanglement wedge reconstruction. These considerations make clear that quantum information in the boundary theory plays a crucial role in the spatial structure of the bulk gravity and motivates the slogan 'space-time emerges from the entanglement structure of the boundary theory'. In the above discussion of the RT formula, we restricted to the most straightforward situation where we have a stationary state; generalizations for general time-dependent states and states with entropic matter in the bulk have been proposed as well [HRT07, FLM13, EW15]. These generalizations have played an important role in recent advances in understanding the black hole information paradox.

1.3.2   Holographic toy models and random tensor networks

Holographic theories of quantum gravity may be relatively well-controlled, they are still complicated strongly-coupled quantum field theories. For this reason, it has been very helpful in our understanding of quantum information principles in holography to study simple toy models which mimic holographic theories in important aspects. One example are relatively simple models of gravity (such as JT gravity, and its approximation by the SYK model as a boundary theory) in which more explicit computations are possible. Other prominent examples of toy models are constructed as tensor networks. Recall that for a tensor network, if we look at a subsystem A, any edge cut γ_A in the network separating A from its complement gives an upper bound on the entropy of $\log(D)|\gamma_A|$, where D is the bond dimension. Thus, we get an upper bound given by the *minimal cut* in the tensor network. If this upper bound is (approximately) saturated, one could think of this as a version of the RT-formula, where one has dangling edges (on which the state lives) as the 'boundary' and internal nodes of the network as the 'bulk'. This analogy is illustrated in Fig. 1.4b. There are various methods to construct such tensor networks. One of the first developments in this regard was the observation that the structure of MERA bears similarity to the holographic principle [Swi12a]: it models a CFT, and the circuit extends into an additional 'scale dimension'. Moreover, the tensor network is reminiscent of a hyperbolic (AdS) space, and the cut shown in Fig. 1.3 is similar to an RT surface. Another example is provided by the HaPPY tensor network [PYHP15] which can be constructed on an arbitrary graph with vertices with fixed degree, and which uses perfect tensors (tensors which are isometries under any bipartition of the half-edges) to ensure that for relevant cuts the minimal cut upper bound is saturated. This can also be used as a toy model for bulk reconstruction: one can construct tensor networks which also have degrees of freedom in the bulk. This perspective has been helpful in thing about bulk reconstruction in terms of an *error correcting code*: a code space of bulk degrees of freedom is encoded in the boundary Hilbert space, and restricting to a boundary subregion corresponds to an erasure error on the complement [Har17]. The error correction perspective has been useful for understanding the emergence of locality in the bulk [ADH15, Har18].

Finally, a very useful and versatile toy model has been the class of *random tensor networks* [HNQ$^+$16], which will be the topic of Part III. This is simply a PEPS tensor network where the tensors are taken to be uniformly random tensors. It can be shown that, with high probability, these models approximately satisfy the RT formula and have the desired error correction (bulk reconstruction) properties. It has been argued that there is a strong correspondence between *fixed-area states* (which do not take into

account that there are fluctuations in the size of the minimal surface and have a flat entanglement spectrum) in holography and random tensor network states [DHM19]. One may wonder:

> Can random tensor networks be used as a toy model for holographic quantum gravity beyond fixed area states?

We study a model where the maximally entangled states in the PEPS construction are replaced by arbitrary entangled states. This reproduces non-flat entanglement spectra. Moreover, this is particularly well suited to study interesting recent developments in the connection between one-shot quantum information theory and holography [AP20, AP22] which allows for a more general understanding of the RT formula and bulk reconstruction. To test the usefulness of this model, we study the following fundamental holographic question:

> What happens at a phase transition between two different RT surfaces?

That is, what happens to the RT formula and its generalization when there are two competing minimal surfaces of approximately the same minimal size. This is an important question: such a phase transition is the basis for holographic computations of the Page curve. Also, it is close related to a phase transition between two disjoint regions, which are either entangled or not (their entanglement wedge is connected or disconnected). We study this question for random tensor network states with two competing minimal cuts, and we show that there are corrections to the RT formula and the entanglement spectrum, confirming and mirroring computations for holographic systems [PSSY19, MWW20, AP20], using methods from random matrix theory and one-shot quantum information theory.

In Chapter 10 we introduce aspects of quantum information theory in holographic quantum gravity, elaborating in Chapter 11 on the use of the replica trick. Then, in Chapter 12 we introduce the random tensor network model formally and discuss the replica trick for random tensor networks. Chapter 13 and Chapter 14 contain our main results of Part III: proofs of convergence of the entanglement spectra of random tensor networks with nontrivial link states in two different regimes.

Notation and conventions

We now introduce some general notations and conventions we will use throughout the dissertation. Given a Hilbert space $\mathcal{H}$, we write $\langle \cdot, \cdot \rangle$ for the inner product and $\|\cdot\|$ for the norm of vectors. We denote by $B(\mathcal{H})$ the space of bounded operators on $\mathcal{H}$ and the operator norm of an operator A by $\|A\|$. We denote Hermitian adjoints by A^*, and we write $A \leq A'$ if the difference $A' - A$ is positive semidefinite. We denote identity operators by $I_{\mathcal{H}}$ and leave out the subscript if the Hilbert space is clear from the context. For a linear operator A we write $\|A\|_p = \mathrm{tr}[(A^*A)^{\frac{p}{2}}]^{\frac{1}{p}}$ for the Schatten p-norm (mostly the cases $p = 1, 2$ will be relevant). For the finite dimensional Hilbert space $\mathbb{C}^n$, we

use bra-ket notation and write $|0\rangle, \ldots, |n-1\rangle$ for the standard basis. If $|\phi\rangle$ is a vector in bra-ket notation, we use the convention that $\phi = |\phi\rangle\langle\phi|$ is the associated projector (quantum state if $|\phi\rangle$ is normalized). For $n \in \mathbb{N}$, we write $[n] := \{1, \ldots, n\}$ and for a set X we write $|X|$ for the number of elements of the set. Apart from these generalities, we will introduce notation and recall definitions throughout the main body.

Part I

Entanglement renormalization and wavelets

Introduction to entanglement renormalization

One of the most promising applications of quantum computers is the simulation of strongly correlated quantum systems. A problem of special interest is finding the *ground state* of some given Hamiltonian. As we saw in Section 1.1 this is generally a difficult problem, but there is good hope that ground states for physically relevant quantum systems can be found efficiently by a quantum computer. In this light, a closely related (weaker) question is whether the ground state of a Hamiltonian of interest can be prepared or approximated by a quantum circuit of low depth. From the perspective of classical computation, a similar question is whether this ground state can be approximated by a tensor network state in an ansatz class with a relatively small number of parameters.

In this part of the dissertation, which is based on [WW21c, WSSW22], we will address this question for a very specific class of one-dimensional quantum systems (free fermions and free bosons) and provide a theoretical analysis of one approach to such quantum circuits (and associated tensor networks) which is known as *entanglement renormalization* using wavelets. In the current chapter we provide a brief introduction to entanglement renormalization and give an informal overview of our results. In Chapter 3 we review basic wavelet theory and provide proofs of various useful results. We first discuss *fermionic* entanglement renormalization in Chapter 4. Then, in Chapter 5 we discuss the *bosonic* version.

2.1 Entanglement renormalization

An important task in the study of quantum many-body systems is finding useful parameterizations of physically relevant quantum states. In Section 1.1.2 we encountered one particular approach, which is by using tensor network states, which are defined by contractions of local tensors according to a network or graph structure. This gives a natural way to prescribe the entanglement structure of the state, while retaining the ability to describe interesting states such as low energy states of local Hamiltonians. Tensor networks are particularly useful to implement real-space renormalization inspired numerical methods for strongly interacting quantum many-body systems. In one spatial dimension, prominent examples are the density matrix renormalization group [Whi92], with the associated tensor network class of matrix product states (MPS) we defined
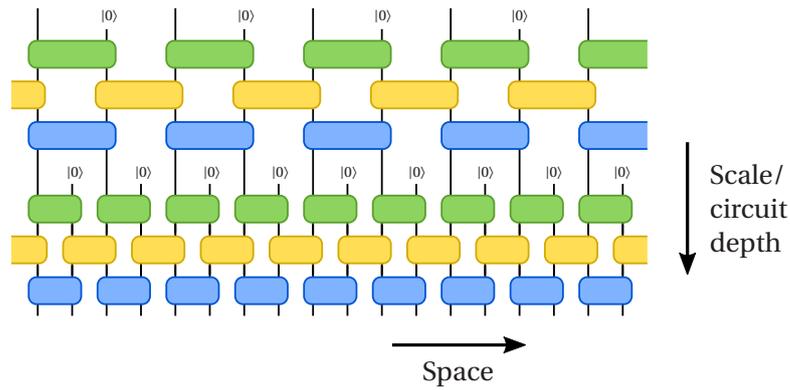
Figure 2.1: The structure of an entanglement renormalization circuit. Each layer is a constant depth quantum circuit that is supposed to implement a real-space renormalization. Every layer takes as input the output of the previous layer and a product state, resulting in an entangled quantum state at the bottom. Layers further up in the figure correspond to structure at larger scales.

in Section 1.1.2 and *entanglement renormalization* [Vid07], with the corresponding multiscale entanglement renormalization ansatz (MERA) states [Vid07, Vid08]. The idea of entanglement renormalization is to perform real-space renormalization (so the renormalization is spatially local) in a unitary fashion. Intuitively, it can be described as follows. We start with the ground state $|\psi\rangle$ of a Hamiltonian H on a one-dimensional lattice $\mathbb{Z}$, which is a potentially highly entangled state. We apply a local quantum quantum circuit to $|\psi\rangle$ which is such that it maps $|\psi\rangle \mapsto |0_{\text{odd}}\rangle \otimes |\phi_{\text{even}}\rangle$. That is, $|\psi\rangle$ is mapped to a state which is a product state (e.g. the $|0\rangle$ state) on the odd sublattice, and to a state $|\phi\rangle$ on the even sublattice which is still entangled, and can be thought of as the ground state of a new (renormalized) Hamiltonian. If the model was at a critical point, this could be the same Hamiltonian. We can interpret this process as 'disentangling the short range entanglement' and $|\phi\rangle$ contains the remaining long-range entanglement. We may then apply the same procedure to $|\phi\rangle$ and iterate this process, each time reducing the nontrivial degrees of freedom by a factor of two. If the model is critical, we may take each layer to be identical. Importantly, this process can be applied in two different directions. Either one can see it (as described above) as taking the state $|\psi\rangle$ and disentangling it layer by layer. Alternatively, we can read the circuit in the other direction, and use it to prepare the state $|\psi\rangle$ starting from a product state on each layer and some appropriate state at the 'top' layer. Typically, given a (critical) Hamiltonian one can not prepare the ground state exactly in this way, but only approximately, where the approximation becomes more accurate with increasing circuit depth.

Importantly, this process can be applied in two different directions. Either one can see it (as describe above) as taking the state $|\psi\rangle$ and disentangling it layer by layer. Alternatively, we can of course read the circuit in the other direction, and use it to prepare the state $|\psi\rangle$ starting from a product state. Typically, given a (critical) Hamiltonian one can not prepare the ground state exactly in this way, but only approximately, where the approximation becomes more accurate with increasing circuit depth. We have not yet specified what the state at the 'top' layer of the circuit should be. However, one can show that for a sufficiently large number of layers (logarithmic in the system size), the choice of state on the top level does not influence the resulting MERA state by much [KK17].
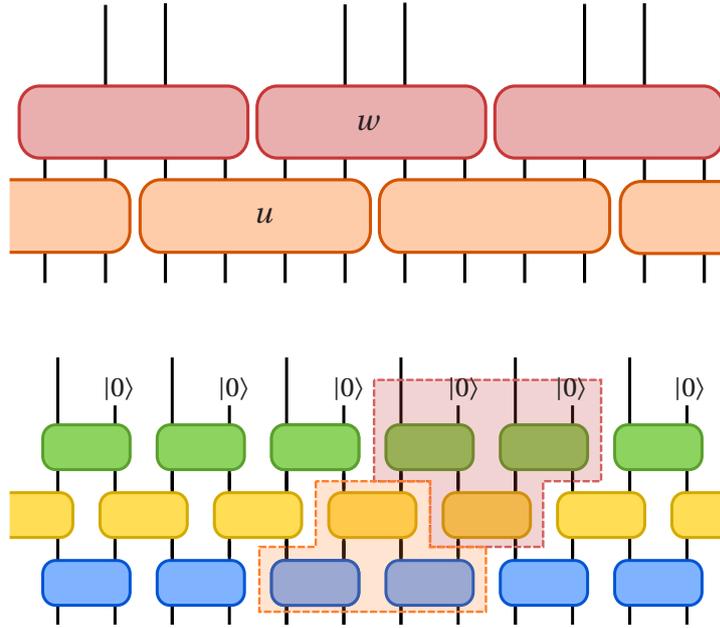
Figure 2.2: (a) A single layer in the MERA ansatz. Here u is a unitary, and w an isometry. If the local degrees of freedom are qubits, this would be a MERA with bond dimension $\chi = 4$. (b) Every entanglement renormalization circuit can be written as a MERA (the converse need not be true) by grouping the unitaries in the circuit appropriately as shown here, see also [KK17].

In general MERA tensor networks, the renormalization layer is often described as the application of first a *unitary* to a number of neighbouring sites (which can be thought of as disentangling the sort range entanglement of these sites) and then applying (the adjoint of) an *isometry* to the output of neighbouring unitaries (which can be thought of as projecting onto the renormalized state). This structure is shown in Fig. 1.3. In other words, a MERA tensor network state prepares an (approximate) ground state through a series of layers, each of which consists of isometries followed by local unitary transformations. The dimension of the input Hilbert space for the isometry is called the *bond dimension*. For a scale-invariant theory, each of the layers can be taken identical. The MERA described here is a so-called binary MERA, other structures (for instance ternary MERA [EV09a]) are also possible.

The way in which MERA is typically used for computational problems is in variational algorithms. That is, one takes the MERA states as an ansatz class and then minimizes the energy $\langle \psi_{\text{MERA}} | H | \psi_{\text{MERA}} \rangle$ over this class. A crucial aspect of MERA which allows one to perform this optimization relatively efficiently is that the unitary structure of the network gives rise to a causal cone, so one can perform the optimization with respect to local terms in the Hamiltonian on a local patch of the tensor network. See for instance [EV09a, HVDH21] for gradient descent based numerical methods to perform this optimization.

Unfortunately our analytic understanding of entanglement renormalization is still limited (as compared to for instance MPS). Ideally, one would like to know conditions under which the ground state of a Hamiltonian can be approximated as a MERA state, and how the approximation accuracy scales with the bond dimension.

2.1.1 Entanglement renormalization and quantum computing

In our constructions we actually obtain low-depth local quantum circuits implementing the entanglement renormalization layer, as shown in Fig. 2.1, and for this reason we will often use the term *entanglement renormalization circuits* rather than MERA throughout this part of the dissertation. This is a relevant distinction, as this class of states can be prepared efficiently on a quantum computer, which makes them a promising ansatz class for variational optimization on a quantum computer. This latter perspective was introduced in [KS17a], where the corresponding class was called *DMERA*. In Fig. 2.2 it is illustrated how an entanglement renormalization circuit is a special case of MERA, by blocking together gates in the circuit to form the unitaries and isometries in the usual binary MERA. One also sees that increasing circuit depth corresponds to exponentially increasing bond dimension. The main bottleneck in classical algorithms for MERA is the contraction of the tensor network state to compute for instance expectation values of observables (which is necessary for all variational algorithms). These contractions scale polynomially in the bond dimension χ, for instance in practice as $\mathcal{O}(\chi^9)$ for the binary MERA [EV09a]. While polynomial in χ, this quickly becomes computationally expensive. If we consider an entanglement renormalization circuit, χ is exponential in the circuit depth, so the contraction cost using known classical contraction algorithms increases exponentially with increasing circuit depth. On the other hand, the ansatz is already given as a quantum circuit, so the computation of expectation values of observables is efficient on a quantum computer: the number of layers required is typically logarithmic in the system size, and the circuit depth of a single layer typically scales polylogarithmically in the desired error. This suggests that one may use a quantum computer to perform the tensor network contractions, and a classical computer to perform the optimization over the parameters in the circuit [KS17a], in other words, to use entanglement renormalization circuits for a VQE as discussed in Section 1.1. Note that entanglement renormalization circuits are a subclass of general MERA, and potentially MERA could be more expressive with the same number of parameters. However, from numerical evidence [HGPC21] it appears that the accuracy of entanglement renormalization circuits (DMERA) still scales favorably with the number of variational parameters compared to regular MERA.

A final appealing property of entanglement renormalization circuits is that they are robust to small errors (a phenomenon we will also encounter in our results). Naively, applying multiple short-depth circuits would lead to an accumulation of errors. However, the scale-invariant nature of MERA ensures that errors in 'deep' layers will not contribute too much. This property makes them interesting candidates for noisy intermediate-scale quantum (NISQ) devices [KS17a, BCSF21, Pre18] which are not capable of fault-tolerant quantum computation.

2.1.2 Entanglement renormalization and quantum field theory

Entanglement renormalization is especially interesting and useful for critical systems (i.e. systems at a phase transition, where the energy gap from the ground state to the first excited state vanishes). In physical systems, at a critical point, it typically happens that the system becomes *scale invariant*. Intuitively, this means that the system is invariant under 'zooming in and out' when renormalizing (that is, the state is a fixed point of the renormalization group). This means that the behaviour of the
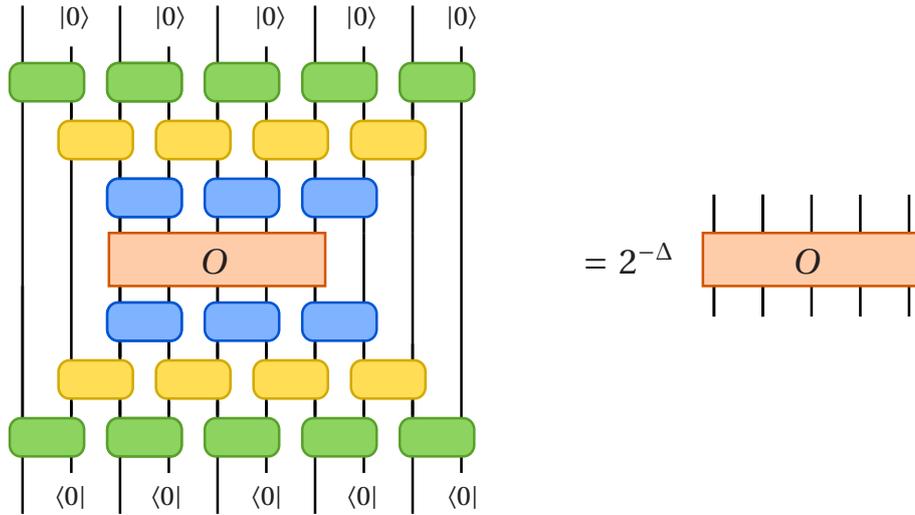
Figure 2.3: Under the MERA superoperator a discretization O of a field operator with scaling dimension Δ is an eigenvector with eigenvalue approximately equal to $2^{-\Delta}$.

system at this point is described by a scale-invariant quantum field theory, a *conformal field theory* (CFT) [FMS12]. Correspondingly, one can take identical layers in a MERA description of the state. Interestingly, one can reconstruct various properties of the CFT that describes the state from the MERA superoperator. The MERA superoperator is the operation of applying a single layer of entanglement renormalization. From the field theory perspective this superoperator should correspond to 'zooming out' by a factor of two. In a CFT, the fields are conveniently organized by their behaviour under rescaling transformations, determined by the *scaling dimensions*. If ϕ_Δ is a field with with scaling dimension Δ, one expects that under rescaling by a factor 2 the field transforms as $\phi(x) \mapsto 2^{-\Delta}\phi(\frac{x}{2})$. It turns out that if one computes at eigenvalues of the MERA superoperator one approximately recovers some of the scaling of the theory, as illustrated in Fig. 2.3. It is also possible to recover OPE coefficients [EV13]. This phenomenon is intuitively plausible, but a thorough theoretical understanding of which conformal data are captured, and how this depends on the bond dimension, is still lacking. We will see that in the wavelet construction of entanglement renormalization for free theories there is a natural connection to the corresponding quantum field theories and one can exactly reproduce various scaling dimensions (with a clear interpretation).

To provide a broader perspective, we mention that several approaches have been proposed to extend the notion of quantum circuits and more generally of tensor networks to quantum field theories. Roughly speaking there are two distinct routes: one is to define a variational class of continuum states, whereas the other is to consider a restricted set of observables and try to approximate correlation functions of these observables. An example of the former is cMERA [HOVV13], which defines a class of states that arise from a real-space renormalization procedure. In this case the 'quantum circuit' that performs the entanglement renormalization is also continuous. Another example is cMPS [VC10], which can be interpreted as a path integral [BHJ$^+$12]. Both cMERA and cMPS have been successfully demonstrated numerically for free theories, and these classes of states have also been used as a basis for perturbation theory [CMMN19] and variational algorithms [HCO$^+$10] for 1+1 dimensional quantum field theories. In

particular, cMPS can be used as a variational ansatz to study interacting field theories at very high precision, see for instance [VC10, GRV17]. Yet rigorous proofs have largely been elusive. We will follow the second route, by considering correlation functions of smeared operators. These operators are discretized at an appropriate scale and an ordinary (entanglement renormalization) quantum circuit is used to prepare a state with which to compute their correlation functions.

A final motivation to investigate tensor networks for conformal field theories is provided by the wish to study holography and the AdS/CFT correspondence, which we encountered in Section 1.3. It has been remarked that entanglement renormalization has a structure reminiscent of this duality [Swi12b], as the circuit reorganizes a critical one-dimensional system to a two-dimensional structure that is a discretization of AdS space, although the precise connection to holographic theories is still being developed [BCC+15, MV18a]. Any MERA tensor network can be extended to a unitary quantum circuit by extending the isometries to unitaries with an auxiliary input, so that the MERA is recovered by applying the circuit to an appropriate product state. Such extensions are not unique. In contrast, the wavelet based constructions naturally yield unitary quantum circuits that reorganizes the degrees of freedom of the Dirac theory in one higher dimension, by position and scale, cleanly separating positive and negative energy modes of the Dirac fermion. Thus it can be seen as a circuit realization of a holographic mapping for an actual conformal field theory, complementing tensor network toy models of holographic mappings as proposed in [PYHP15, YHQ16, HNQ+16, NW16]. Of course, the relation to holography should be seen here as merely an analogy, as actual holographic quantum field theories with an AdS dual gravity theory are strongly interacting theories rather than free field theories.

2.1.3   Entanglement renormalization and wavelets

Unfortunately our analytic understanding of entanglement renormalization is still limited (as compared to for instance MPS). Ideally, one would like to know conditions under which the ground state of a Hamiltonian can be approximated as a MERA state, and how the approximation accuracy scales with the bond dimension. One direction in which progress to such analytic understanding has been made, and as will be described in detail in this part of the dissertation, is in connection to wavelets.

Wavelet transforms decompose a signal as a linear combination of localized wave packets or 'wavelets' at different scales (as compared to the Fourier transform, which uses plane waves). We will review the wavelet transform in some detail in Chapter 3. For now we give a brief informal description so we can explain the relation to entanglement renormalization.

In each step of the *discrete wavelet transforms* a signal f is decomposed into a high-frequency component f_w (the 'details' of the signal) and a low-frequency component f_s (the 'large scale structure' of the signal) by applying a high-frequency (wavelet) filter g_w and low-frequency (scaling) filter g_s and subsampling the result on the even sublattice. The wavelet transform then proceeds iteratively on the low-frequency component of the signal. The original signal can be reconstructed using the filters g_w and g_s as well. In fact, if one iterates this reconstruction procedure, one can construct a limiting continuous function ψ from the filters. This leads to the *continuous wavelet transform*: it turns out that shifted and scaled versions of ψ, given by $\psi_{j,k}(x) = 2^{\frac{j}{2}} \psi(2^j x - k)$ form
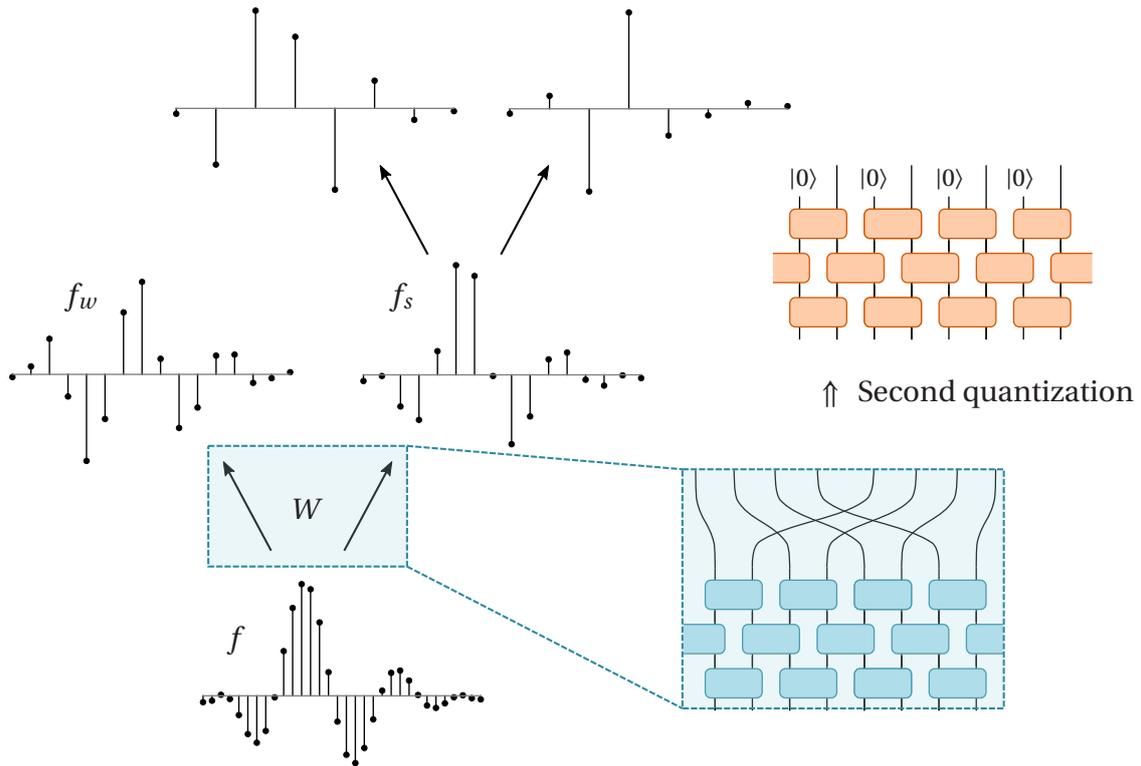
Figure 2.4: An illustration of the (discrete) wavelet decomposition map W. It decomposes a signal f into a scaling (low-frequency) signal f_s and a wavelet (high-frequency) signal f_w. This can be applied iteratively to the scaling signal; each layer of W 'zooms out' by a factor of 2. The transform W can be written as a classical circuit which can be second quantized to obtain an entanglement renormalization circuit.

an orthonormal basis for the space of square-integrable functions. The fact that these functions are compactly supported, yet allow for a Fourier-type function decomposition has led to the nomenclature *wavelet function* for ψ.

The structure and philosophy of the discrete wavelet transform is very similar to real-space renormalization, and its original development was partially motivated by applications in real-space renormalization [Bat99]. In particular, the iterative nature of the wavelet transform bears a close similarity to entanglement renormalization: in this case each layer of the circuit disentangles the 'high frequency' (short scale) entanglement, and then proceeds iteratively with the remaining 'low frequency' (long scale) degrees of freedom, which are now organized on a coarse-grained lattice. Moreover, to deepen the analogy, the discrete wavelet transform can be written as a 'circuit', which is a 'classical' (or single-particle) circuit in the sense that it acts on direct sums of degrees of freedom rather than a tensor product. In [EW16, HSW⁺18] it was shown that this more than just an analogy: second quantization of an appropriate pair of discrete wavelet transforms gives rise to fermionic entanglement renormalization circuits for a critical lattice fermion. In other words, one considers the second quantization of a *single-particle circuit*, which is given by a wavelet transform. Through a Jordan-Wigner transformation this is equivalent to a quantum circuit on a chain of qubits approximating the ground state of the critical Ising model. These ideas are illustrated in Fig. 2.4. In fact, the wavelet-based entanglement renormalization circuits from [EW16] have

recently been implemented on an ion trap quantum computer in [SJ21].

2.2 Summary of contributions

In this part of the dissertation we will outline how second quantization of wavelet transforms gives rise to entanglement renormalization for one-dimensional free fermions and bosons. Here, a *free* Hamiltonian is a Hamiltonian which is quadratic in the (fermionic or bosonic) creation and annihilation operators. In practice, ground states for one-dimensional systems can be simulated very accurately using MPS methods (such as DMRG) even at critical points (which are numerically more challenging than gapped phases). Also, the Hamiltonians of the free systems we study in this part of the dissertation are actually analytically solvable. Nevertheless, our results provide a potentially useful perspective on the simulation of many-body quantum systems: as discussed above using entanglement renormalization circuits as a variational ansatz can potentially speed up computations for (critical) lattice systems in one dimension for non-free systems as well. Secondly, entanglement renormalization can also be applied in higher spatial dimensions [EV09b], in which case classical tensor network methods quickly become computationally expensive. Understanding in detail the solution for free systems is a crucial first step to obtain better understanding of these applications.

2.2.1 Informal statement of results

Fermionic entanglement renormalization

In Chapter 4 we investigate entanglement renormalization for a free fermionic system. From previous work [EW16, HSW⁺18] it was known that from second quantization of an appropriate 'Hilbert pair' of discrete wavelet transforms one obtains fermionic entanglement renormalization circuits for a critical lattice fermion.

In Chapter 4, based on [WSSW22], we extend the results of [EW16, HSW⁺18] by showing that the continuum limit of the wavelet transform, and the associated wavelet and scaling functions have a natural interpretation for the continuum limit of the fermion, a Dirac quantum field theory [FMS12].

To be precise, we consider the free massless Dirac fermion in 1+1 dimensions, with action

$$S(\Psi) = \frac{1}{2} \int \Psi^\dagger \gamma^0 \gamma^\mu \partial_\mu \Psi \, \mathrm{d}x \mathrm{d}t$$

for a two-component complex fermionic field Ψ on the line (or on a circle). We take the algebraic approach to quantum field theories, which allows for a rigorous description of (free) quantum field theories [Haa12]. To have bounded operators we 'smear' the fields. For a smearing function f we define the bounded operator $\Psi(f)$ which should be thought of (informally) as

$$\Psi(f) = \int f(x) \Psi(x) \mathrm{d}x.$$

From a physical perspective the smearing function is justified by the fact that one can only probe the system at some finite scale.

We will now describe a procedure which approximates correlation functions of smeared operators. Informally, the procedure is that we first discretize the operators at some scale (i.e., we impose a UV cut-off), and then, in order to obtain the free fermion ground state, we need to 'fill the Dirac sea' up to the relevant scale. So, the circuit, starting from the Fock vacuum, has to fill all the negative energy modes over the range of scales that are relevant for the inserted operators, directly analogous to a real-space renormalization procedure.

We know the negative energy states explicitly in Fourier space, but the non-trivial problem is that we want to construct a *local* circuit, while the Fourier basis for the negative energy solutions is very non-local. In order to obtain a circuit that is compatible with scale invariance and translation invariance, but is still local, we thus use a wavelet basis for the space of negative energy solutions. It is not possible to construct a basis that is both completely local and consists of exactly negative energy solutions, but it turns out it is approximately possible by using a pair of wavelets that approximately satisfy a certain phase relation, leading to a so-called *Hilbert pair* of wavelets. This construction takes as input two integer parameters K and L, such that the support of the wavelet is of size $2(K + L)$, and there is an approximation parameter ε which measures how accurately the phase relation is satisfied. The wavelet functions give rise to a 'classical' circuit, which implements the decomposition of a function in the wavelet basis at different scales. This circuit should be thought of as a circuit on the single-particle level, and the fermionic quantum circuit is obtained as its second quantization.

Now let $\{O_i\}$, $i = 1, \ldots, n$ be a set of smeared operators that are either linear in the fields or normal-ordered quadratic operators, and which are compactly supported. We would like to compute correlation functions

$$G(\{O_i\}) = \langle O_1 \cdots O_n \rangle. \tag{2.1}$$

The procedure sketched above discretizes the operators O_i and constructs a quantum circuit that computes an approximation $G_{\mathscr{L},\varepsilon}^{\mathrm{MERA}}(\{O_i\})$ of the correlation function, where $\mathscr{L}$ is the number of layers of the circuit, and ε is an error parameter.

The following is a simplified version of our main result. A precise formulation is given by Theorem 4.7, where we also specify precisely which operators we consider and give explicit bounds for the approximation error.

We assume that we are given a family of wavelet filters with uniformly bounded scaling functions, of support N and approximating the Hilbert pair relation to accuracy ε. The constructed circuits have depth $D = \lceil \frac{N}{2} \rceil$ for a single circuit layer, and the bond dimension of the corresponding MERA tensor network is given by $\chi = 2^D$. For simplicity we consider a two-point function (see Theorem 4.7 for the error scaling with the number of operator insertions).

Theorem (Informal). *Let O_1, O_2 be Dirac field creation or annihilation operators or normal-ordered quadratic operators, smeared by a differentiable function with compact support. Then the approximation error for a MERA state with $\mathscr{L}$ layers is bounded by*

$$|\langle O_1 O_2 \rangle - \langle O_1 O_2 \rangle_{\mathrm{MERA}}| = \mathscr{O}(N^2 2^{-\frac{\mathscr{L}}{3}}) + \mathscr{O}(\varepsilon \log \tfrac{N}{\varepsilon}).$$

The constants in the $\mathscr{O}$-notation depend the support and smoothness of the O_i.

Our main theorem provides a justification for the numerical success of MERA for quantum field theories by providing rigorous bounds on the approximation of correlation functions. A Dirac fermion can be decomposed into two Majorana fermions. Our

construction is compatible with this decomposition, so we also obtain quantum circuits for Majorana fermions. It remains an open problem to prove analytic bounds on the decay of ε with N. For the construction of [Sel02] where $N = 2(K + L)$, numerically the parameter ε is seen to decrease exponentially with $\min\{K, L\}$.

As mentioned above, the wavelet construction gives rise to an entanglement renormalization circuit rather than an (unstructured) MERA tensor network in a canonical way. We also show that the scaling dimensions of the fermionic fields (and of a number of descendents) are captured exactly by the MERA superoperator. Our work therefore elucidates the relation between entanglement renormalization and quantum field theory, and provides a new perspective on quantum simulation of quantum field theories.

Bosonic entanglement renormalization

In [EW16] it was suggested that wavelet constructions could also be relevant for free bosonic systems, and in Chapter 5, based on [WW21c], we confirm this suggestion. In order to formulate what this entails, we will work with bosonic quantum circuits. Just as for free fermionic systems, we will look at second quantizations of single-particle circuits, or Gaussian circuits. This is also known as linear optics circuits, and means that each local operation is implemented by time evolution along a quadratic Hamiltonian. This is subclass of all bosonic quantum circuits which can be efficiently simulated (upon adding a single non-Gaussian bosonic quantum gate to the set of allowed operations, however, bosonic quantum circuits are able of universal quantum computation [KLM01]). In contrast to more usual notions of quantum circuits and tensor networks, the Hilbert spaces are infinite dimensional. In particular, the usual definition of a tensor network with a finite bond dimension has no immediate analogue. However, finite-depth quantum circuits such as entanglement renormalization circuits of the form of Fig. 2.1 are still meaningful even in this infinite-dimensional bosonic setup. The notion of Gaussian bosonic entanglement renormalization has been introduced and studied in [EV10a], in which an extensive explanation of the formalism can be found.

In Chapter 5 we show that one can indeed construct a Gaussian bosonic entanglement renormalization scheme for bosonic quadratic one-dimensional Hamiltonians, using the second quantization of *biorthogonal wavelet filters*, a generalization of the usual orthogonal wavelet filters. Interestingly, in the bosonic case our results are not restricted to the scale-invariant case, but can be used to construct entanglement renormalization circuits for arbitrary translation invariant quadratic bosonic Hamiltonians. Given such a Hamiltonian, we explain how a corresponding (approximate) entanglement renormalization circuit can be found by solving a filter design problem. We also give a general heuristic for constructing such filters in Section 3.4, similar to the construction of the Daubechies wavelets. This is in contrast to the fermionic case, where the only known constructions are for massless (critical) fermions. Finally, as in the fermionic case, the continuum limit of the discrete system is directly related to the continuous scaling and wavelet functions, associated to the continuous limit of the discrete wavelet transform. Moreover, for the free massless boson our construction reproduces various scaling dimensions exactly. If the system is not scale-invariant, we explain how one can still define versions of the wavelet and scaling functions which are not scale-invariant.

We prove anr approximation theorem, formally stated as Theorem 5.3, for the correlation functions of the MERA state, given biorthogonal wavelet filters with certain

properties, similar to the fermionic results of [HSW$^+$18]. In this case we focus on the discrete setting. In Theorem 5.3, we prove a general result which applies to an arbitrary quadratic Hamiltonian. In the particular case of the harmonic chain with Hamiltonian

$$H = \frac{1}{2}\Big(\sum_{n\in\mathbb{Z}} p_n^2 + m^2 q_n^2 + \frac{1}{4}(q_n - q_{n+1})^2\Big),$$

specializes as follows. We measure the error in the two-point functions $\langle p_i p_j \rangle$ and $\langle q_i q_j \rangle$ (or covariance matrix). We assume that we are given an appropriate wavelet pair (which we are not guaranteed to exist in general, but for which we provide a numerical construction) which satisfies a certain relation up to error ε, with support N and which is stable in the sense that the wavelet decomposition map and scaling functions are uniformly bounded. Then we have the following:

Theorem (Informal). *For the harmonic chain with mass m, the approximation error using the MERA state resulting from $\mathscr{L}$ layers of entanglement renormalization is bounded by*

$$|\langle p_i p_j \rangle_{\text{exact}} - \langle p_i p_j \rangle_{\text{MERA}}| = \big(\mathscr{O}(N^{\frac{3}{2}} 2^{-\frac{\mathscr{L}}{2}}) + \mathscr{O}(\varepsilon \log \tfrac{N}{\varepsilon})\big)\sqrt{m^2 + 1},$$

$$|\langle q_i q_j \rangle_{\text{exact}} - \langle q_i q_j \rangle_{\text{MERA}}| = \big(\mathscr{O}(N^{\frac{3}{2}} 2^{-\frac{\mathscr{L}}{2}}) + \mathscr{O}(\varepsilon \log \tfrac{N}{\varepsilon})\big)\frac{1}{m},$$

the latter assuming $m > 0$. In the massless case, the latter bound is replaced by

$$|\langle q_i q_j \rangle_{\text{exact}} - \langle q_i q_j \rangle_{\text{MERA}}| = \big(O(N^{\frac{3}{2}} 2^{-\frac{\mathscr{L}}{2}}) + \mathscr{O}(\varepsilon \log \tfrac{N}{\varepsilon})\big)\sqrt{|i - j|}.$$

In the massless case, there is an IR divergence and $\langle q_i q_j \rangle$ is only defined up to a constant, so we define $\langle q_i q_j \rangle$ by subtracting the divergence; see Eq. (5.33) in Section 5.3 for details.

The intuition behind the proof of both the fermionic and the bosonic result is that the contribution of the $\mathscr{L}$-th layer to the correlation function is bounded by $\mathscr{O}(2^{-\frac{\mathscr{L}}{2}})$, so we need $\mathscr{O}(\log \frac{1}{\delta})$ layers to get within error δ (even with perfect filters). In principle, each layer contributes an additive factor of $\mathscr{O}(\varepsilon)$, due to the error in the filter relation, so naively this would yield an error bound of size $\mathscr{O}(\mathscr{L}\varepsilon)$, which does not scale in the desired way for large $\mathscr{L}$. However, deeper layers do not contribute much to the correlation function, so we effectively have only $\log(\frac{1}{\varepsilon})$ layers where the inaccuracy of the phase relation is relevant, leading to an error of size $\mathscr{O}(\varepsilon \log(\frac{1}{\varepsilon}))$. If we denote by N the circuit depth of a single layer, then we find numerically that ε can be made exponentially small as a function of N, whereas the other wavelet-dependent parameters we have suppressed above only grow polynomially. Hence, the total required depth of a single layer of entanglement renormalization for a desired error is polylogarithmic in ε^{-1}. This shows that our entanglement renormalization circuits prepare the ground state very efficiently: a circuit of overall depth $\mathscr{O}(\text{polylog}(\frac{1}{\delta}))$ achieves an accuracy δ on the correlation functions. Overall, we believe that our work, in [WSSW22, WW21c], together with [EW16, HSW$^+$18], essentially completes our conceptual understanding of Gaussian entanglement renormalization for free one-dimensional theories as the second quantization of wavelet decompositions.

2.2.2 Prior work

The only rigorous results and constructions that are known for entanglement renormalization and MERA in one-dimensional critical systems rely on wavelet theory. Our

work adds extensively to this line of research by extending the work of [EW16, HSW$^+$18] to the continuum and to bosonic systems. The idea to use wavelet theory for renormalization is very intuitive and in fact dates from the early phases of wavelet theory (see for instance [Bat99]). In Refs. [Qi13, Lee17] Haar wavelets were used as a fermionic holographic mapping, and in Ref. [SB16] Daubechies wavelets were used as a bosonic holographic mapping. The connection to quantum circuits and entanglement renormalization was made in Refs. [EW16, EW18]. For us Ref. [HSW$^+$18] is especially relevant. In this work a systematic method to construct circuits for lattice fermions was described based on the discrete wavelet transform of a pair of wavelets with special properties, a so-called approximate Hilbert pair. It is a well-known that there is both a *discrete* and a *continuous* perspective on wavelet transforms, but from previous work it was unclear whether and how continuous wavelet theory relates to the quantum field theory limit of the lattice systems. We show that the extension of the discrete to the continuous wavelet transform precisely relates the entanglement renormalization circuits in a natural way to the Dirac fermion. This offers a new perspective on the relation between MERA and its continuum limit. For instance, it becomes very clear why the entanglement renormalization superoperator captures some scaling dimensions of the theory exactly in this construction.

For the bosonic case a natural idea is to simulate bosonic quantum field theories on a bosonic quantum computer [MPSW15], and wavelets are a very efficient choice of basis to discretize a quantum field theory for this purpose [BRSS15]. We explain that for any free 1+1-dimensional bosonic field theory, one can use suitably chosen biorthogonal wavelets to discretize the theory and use the corresponding wavelet decomposition to prepare its (approximate) ground state using the bosonic Gaussian entanglement renormalization circuit. The idea to use wavelets to discretize a field theory is been suggested before, see for instance [BP13, BRSS15, SMMT20] for some recent discussions of discretizing bosonic field theories using wavelets. Our approach however fundamentally differs from these works in that we use *biorthogonal* wavelets (as is natural in the bosonic setting), which moreover are specifically designed to target the Hamiltonian of the field theory (rather than using off-the-shelf wavelets such as the Daubechies wavelets).

Wavelet theory

Our entanglement circuits will be obtained by second quantization of a wavelet transformation. In Chapter 4 and Chapter 5 we will explain fermionic and bosonic second quantization and how it leads to entanglement renormalization. In the current chapter we review the basic theory of wavelets. We will use *orthogonal* wavelets for fermionic circuits and *biorthogonal* wavelets for bosonic circuits. In Section 3.1 we explain the definition of an orthogonal wavelet basis, and how a choice of wavelet basis stratifies a function space into different scales. The difference with Fourier analysis is that the basis functions are not plain waves but 'wavelet' functions, which are spatially localized wave packets. We also explain how different scales in this decomposition are related through a filtering procedure. Then, in Section 3.2 we generalize this to biorthogonal wavelet bases. In Section 3.3 we discuss how to construct (biorthogonal) wavelet filters which satisfy certain prescribed phase relations, which is crucial to the entanglement renormalization construction. In particular we define and construct *(approximate) Hilbert pairs* of wavelet filters. Next, we explain how (biorthogonal) wavelet transforms are implemented by a classical linear circuit (which can be interpreted as a single-particle quantum circuit) in Section 3.5. Finally, in Section 3.6 we will prove various technical results on approximations using wavelets which are the core ingredient for our entanglement renormalization approximation theorems. For a more extensive introduction to wavelets we refer the reader to, e.g., Chapter 7 in [Mal08]. Most of the material in this section is a review of wavelet theory. However, the general construction of circuits for biorthogonal wavelet filters, the construction of biorthogonal wavelet filters in Section 3.3 and the results of Section 3.6.1 are our own contribution.

Notation

We now introduce some notation and conventions with respect to function spaces and Fourier analysis that we will use throughout Part I. We define the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ as the interval $[0, 1]$ with endpoints identified. We write $L^2(\mathbb{R})$, $L^2(\mathbb{S}^1)$, etc. for Hilbert spaces of square-integrable functions with respect to the Lebesgue measure that assigns unit measure to unit intervals, and we denote by $\ell^2(\mathbb{Z})$ the Hilbert space of square-integrable sequences. The Fourier transform of a function $\phi \in L^2(\mathbb{R})$ is denoted by $\hat{\phi} \in L^2(\mathbb{R})$ and is given by $\hat{\phi}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-ix\omega}dx$ if ϕ is absolutely integrable. Similarly, the Fourier transform of a function $\phi \in L^2(\mathbb{S}^1)$ is denoted by $\hat{\phi} \in \ell^2(\mathbb{Z})$ and can be computed as $\hat{\phi}(n) = \int_0^1 f(x)e^{-ix2\pi n}dx$. Lastly, we define the Fourier transform of a sequence $f \in$

$\ell^2(\mathbb{Z})$ to be the 2π-periodic function $\hat{f} \in L^2(\mathbb{R}/2\pi\mathbb{Z})$ given by $\hat{f}(\theta) = \sum_{n \in \mathbb{Z}} f[n]e^{-i\theta n}$.[1]
For $\mathcal{H} = L^2(\mathbb{R})$, $L^2(\mathbb{S}^1)$, or $\ell^2(\mathbb{Z})$, and $\hat{\lambda}$ a bounded function on the Fourier domain,
i.e. an element of $L_\infty(\mathbb{R})$, $L_\infty(\mathbb{Z})$ or $L_\infty(\mathbb{S}^1)$ respectively, we will denote by $m(\hat{\lambda}) \in B(\mathcal{H})$
the *Fourier multiplier* with symbol $\hat{\lambda}$, defined by multiplication with $\hat{\lambda}$ in the Fourier
domain (equivalently, convolution with the inverse Fourier transform λ in the original
domain). In order for $m(\hat{\lambda})$ to be a bounded operator we need that $\hat{\lambda}$ is a bounded
function On $\ell^2(\mathbb{Z})$, we define the *downsampling* operator $\downarrow$ by $(\downarrow f)[n] = f[2n]$; its
adjoint is the *upsampling* operator $\uparrow$ given by $(\uparrow f)[2n] = f[n]$ and $(\uparrow f)[2n+1] = 0$
for $f \in \ell^2(\mathbb{Z})$. We will also use the Sobolev spaces $H^K(\mathbb{R})$ and $H^K(\mathbb{S}^1)$, which consist of
functions that have a square-integrable weak K-th derivative, denoted $f^{(K)}$. All p-norms
for $p \neq 2$ will be denoted by $\|f\|_p$, while we write $\|f\| = \|f\|_2 = \sqrt{\langle f, f \rangle}$. We write $\mathbf{1}$ for
the constant function equal to one, and $\mathbf{1}_X$ for the indicator function of a set X.

3.1   Wavelet bases

An (orthogonal) wavelet basis is an orthonormal basis for $L^2(\mathbb{R})$ consisting of scaled
and translated versions of a single localized function $\psi \in L^2(\mathbb{R})$, called the (orthogonal)
wavelet function. In this case we define

$$W_j = \{\psi_{j,k} : k \in \mathbb{Z}\}, \quad \text{where} \quad \psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k),$$

so $L^2(\mathbb{R}) = \bigoplus_j W_j$. We can therefore interpret of W_j as the *space of functions at scale j*,
also called the detail space at scale j, where large j corresponds to fine scales and
small j to coarse scales.

   In signal processing, wavelet bases are often constructed from an auxiliary func-
tion $\phi \in L^2(\mathbb{R})$, known as the *scaling function*. We let

$$V_j = \text{span}\{\phi_{j,k} : k \in \mathbb{Z}\}, \quad \text{where} \quad \phi_{j,k}(x) = 2^{\frac{j}{2}}\phi(2^j x - k)$$

and we demand that the V_j form a complete filtration of $L^2(\mathbb{R})$, i.e.,

$$\{0\} \subseteq \ldots \subseteq V_j \subseteq V_{j+1} \subseteq \ldots \subseteq L^2(\mathbb{R}), \quad \overline{\bigcup_j V_j} = L^2(\mathbb{R}).$$

One can show that in this case there exists an associated wavelet function ψ which is
such that the wavelets at scale j span exactly the orthogonal complement of V_j in V_{j+1}:

$$V_{j+1} = W_j \oplus V_j \tag{3.1}$$

for all $j \in \mathbb{Z}$. A sequence of subspaces $\{V_j\}_{j \in \mathbb{Z}}$ as above is said to form a *multiresolution
analysis*, since Eq. (3.1) allows to recursively decompose a signal in some V_j scale by
scale. It follows that V_j is the space spanned by all $\psi_{l,k}$ for $l < j$, i.e.

$$V_j = \bigoplus_{l<j} W_l$$

---

[1] Note the choice of factors of 2π in these conventions. These have been chosen so as to be compatible
with the integer shift structure of the wavelet basis.

(a)

(b)
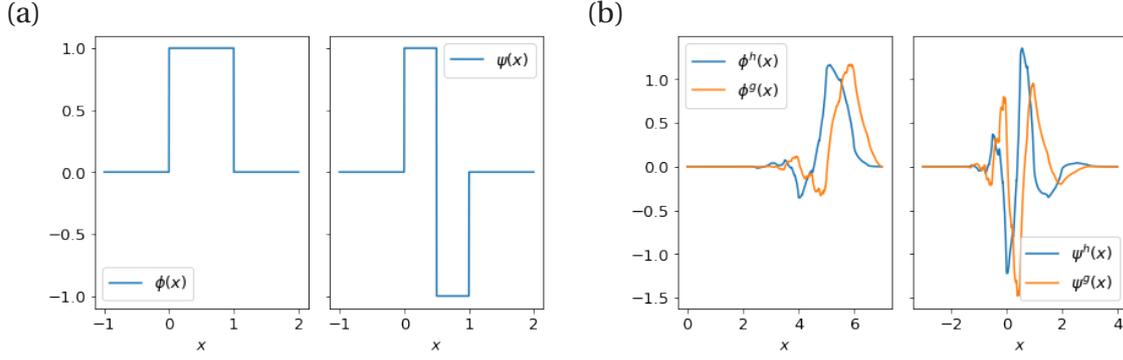
Figure 3.1: (a) Scaling and wavelet function for the Haar wavelet ($\phi = \mathbf{1}_{[0,1)}$, $\psi = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}$). (b) Scaling and wavelet functions for the approximate Hilbert pair with parameters $K = L = 2$ due to Selesnick (ϕ^h, ψ^h in black; ϕ^g, ψ^g in gray). See Section 3.4 and Table 3.1 for further detail.

and the $\phi_{j,k}$ are an orthonormal basis for V_j. Note that for any fixed j

$$L^2(\mathbb{R}) = V_j \oplus \bigoplus_{l \geq j} W_l.$$

Thus we can decompose any function $f \in L^2(\mathbb{R})$ as

$$f = \sum_{k \in \mathbb{Z}} s_k \phi_{j,k} + \sum_{l=j}^{\infty} \sum_{k \in \mathbb{Z}} w_{l,k} \psi_{l,k}$$

where

$$s_k = \langle \phi_{j,k}, f \rangle \qquad w_{l,k} = \langle \psi_{l,k}, f \rangle.$$

In conclusion, we can interpret V_j as the space of *functions up to (but excluding) details at scale j*. Intuitively, the scaling function can be thought of as a 'bump function' and the wavelet function as a 'wave packet'. This intuition, and the orthogonality between scaling and wavelet function are well-illustrated by the *Haar wavelet* (see Fig. 3.1, (a)). We will use pairs of wavelets that are tailored to target the vacuum of the Dirac theory (see Section 3.3 below).

Wavelet bases as above can be obtained by deriving them from filters. A real-valued sequence $g_s \in \ell^2(\mathbb{Z})$ is called a *scaling filter* (or *low-pass filter*) if its Fourier transform satisfies, for all $\theta \in \mathbb{R}/2\pi\mathbb{Z}$,

$$|\hat{g}_s(\theta)|^2 + |\hat{g}_s(\theta + \pi)|^2 = 2 \quad \text{and} \quad \hat{g}_s(0) = \sqrt{2}. \tag{3.2}$$

Under mild technical conditions on g_s (see, e.g., [Mal08, Thm 7.2]), which we always assume to be satisfied, we can define scaling and wavelet functions $\phi, \psi \in L^2(\mathbb{R})$ by

$$\psi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} g_w[n] \phi(2x - n),$$

$$\phi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} g_s[n] \phi(2x - n)$$

and these functions satisfy the conditions above and hence define a wavelet basis (and associated multiresolution analysis). The sequence $g_w \in \ell^2(\mathbb{Z})$ is known as the *wavelet filter* (or *high-pass filter*) and it can be computed via

$$\hat{g}_w(\theta) = e^{-i\theta}\overline{\hat{g}_s(\theta + \pi)}, \quad \text{i.e.} \quad g_w[n] = (-1)^{1-n}\bar{g}_s[1-n]. \tag{3.3}$$

Thus, the expansion coefficients of the wavelet and scaling function at scale $j = 0$ in terms of scaling functions at scale $j = 1$ are precisely given by the wavelet and scaling filters, respectively (cf. Eq. (3.1)). This generalizes immediately to arbitrary scales: For all $j, k \in \mathbb{Z}$,

$$\psi_{j,k} = \sum_{n \in \mathbb{Z}} g_w[n]\phi_{j+1,2k+n}, \tag{3.4}$$

$$\phi_{j,k} = \sum_{n \in \mathbb{Z}} g_s[n]\phi_{j+1,2k+n}. \tag{3.5}$$

In Fourier space, these relations read

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}_w\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right), \tag{3.6}$$

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}_s\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \tag{3.7}$$

for all $\omega \in \mathbb{R}$. The Fourier transform of the scaling function can be expressed as an infinite product of evaluations of the scaling filter:

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} \frac{1}{\sqrt{2}}\hat{g}_s(2^{-k}\omega) \tag{3.8}$$

In particular, it is bounded by one, i.e., $\|\hat{\phi}\|_{\infty} = 1$. It is also useful to note that the wavelet function averages to zero, i.e., $\int_{-\infty}^{\infty} \psi(x)\,dx = 0$.

Throughout this article, we will always work with filters of *finite length* (the length of a sequence $f \in \ell^2(\mathbb{Z})$ is defined as the minimal number N such that f is supported on N consecutive sites). Specifically, we will assume that the support of the scaling filter is $\{0, \ldots, N-1\}$. In the signal processing literature, such filters are called finite impulse response (FIR) filters with N taps. It is clear from Eq. (3.3) that in this case the wavelet filter is supported in $\{2 - N, \ldots, 1\}$, hence has finite length N as well. If the filters have finite length then the wavelet and scaling functions are compactly supported on intervals of width N [Mal08, Prop 7.2].

3.1.1 Wavelet decomposition and reconstruction

Suppose that we would like to express a given function $f \in L^2(\mathbb{R})$ in a wavelet basis. As a first step, we replace f by $P_j f \in V_j$, where $P_j \colon L^2(\mathbb{R}) \to V_j$ denotes the orthogonal projection onto the space of functions below scale j. This is corresponds to removing high frequency components (in signal processing) or to a UV cut-off (in physics). We explain in Lemma 3.3 below how to bound the error $\|f - P_j f\|$ in terms of a Sobolev norm. To express $P_j f$ in terms of the orthonormal basis $\{\phi_{j,k}\}_{k \in \mathbb{Z}}$ of V_j, define the partial isometries

$$\alpha_j \colon L^2(\mathbb{R}) \to \ell^2(\mathbb{Z}), \qquad (\alpha_j f)[k] = \langle \phi_{j,k}, f \rangle, \tag{3.9}$$

where we note that $P_j = \alpha_j^* \alpha_j$. We show below that, if f is sufficiently smooth, the scaling coefficients $\alpha_j f$ can be well-approximated by sampling f on a uniform grid with spacing 2^{-j} (Lemma 3.4).

Next, we iteratively obtain the wavelet coefficients of $P_j f$ at all scales $n < j$. For this purpose, let

$$\beta_j \colon L^2(\mathbb{R}) \to \ell^2(\mathbb{Z}), \quad (\beta_j f)[k] = \langle \psi_{j,k}, f \rangle,$$

and define the unitary operator

$$W \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2, \qquad W f = (\downarrow m(\overline{\hat{g}_w}) f) \oplus (\downarrow m(\overline{\hat{g}_s}) f), \qquad (3.10)$$

where the *downsampling* operator is given by $(\downarrow f)[n] = f[2n]$. Then, Eqs. (3.4) and (3.5) imply that

$$W \alpha_j f = \beta_{j-1} f \oplus \alpha_{j-1} f$$

for all $f \in L^2(\mathbb{R})$ and $j \in \mathbb{Z}$. That is, applying W to the scaling coefficients at some scale j yields in the first component the wavelet coefficients and in the second component the scaling coefficients at one scale coarser. Note that, due to the scale invariance of the wavelet basis, the operator W does *not* depend explicitly on j. We can iterate this procedure to obtain a map

$$W^{(\mathscr{L})} \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}+1}, \qquad W^{(\mathscr{L})} = (I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}-1}} \oplus W) \cdots (I_{\ell^2(\mathbb{Z})} \oplus W) W, \quad (3.11)$$

which decomposes through successive filtering the scaling coefficients at scale $j + 1$ into the wavelet coefficients at scales j to $j - \mathscr{L} + 1$ and the scaling coefficients at scale $j - \mathscr{L} + 1$. That is:

$$W^{(\mathscr{L})} \alpha_j = \beta_{j-1} \oplus W^{(\mathscr{L}-1)} \alpha_{j-1} = \cdots = \beta_{j-1} \oplus \beta_{j-2} \oplus \ldots \oplus \beta_{j-\mathscr{L}} \oplus \alpha_{j-\mathscr{L}},$$

or

$$W^{(\mathscr{L})} \alpha_j f = \sum_{l=0}^{\mathscr{L}-1} \beta_{j-l} f \otimes |l\rangle + \alpha_{j-\mathscr{L}} f \otimes |\mathscr{L}\rangle$$

for all $f \in L^2(\mathbb{R})$. The unitaries W and $W^{(\mathscr{L})}$ are known as ($\mathscr{L}$ layers of) the (orthogonal) *discrete wavelet transform*. $W^{(\mathscr{L})}$ can be readily implemented by a scale-invariant linear circuit consisting of convolutions and downsampling circuit elements (see Section 4.2.4 and Fig. 3.2 for a visualization). The wavelet decomposition map W is an orthogonal map, i.e. $W^{-1} = W^\mathsf{T} = W^*$ (note that the filters are real valued). The inverse can be thought of as a *reconstruction* of the signal from the scaling and wavelet components and is given by

$$W^\mathsf{T} \colon \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \to \ell^2(\mathbb{Z}), \qquad W^\mathsf{T}(f_w \oplus f_s) = (m(\hat{g}_w) \uparrow f_w) + (m(\hat{g}_s) \uparrow f_s), \qquad (3.12)$$

where the *upsampling* operator is given by $(\uparrow f)[2n] = f[n]$ and $(\uparrow f)[2n+1] = 0$.
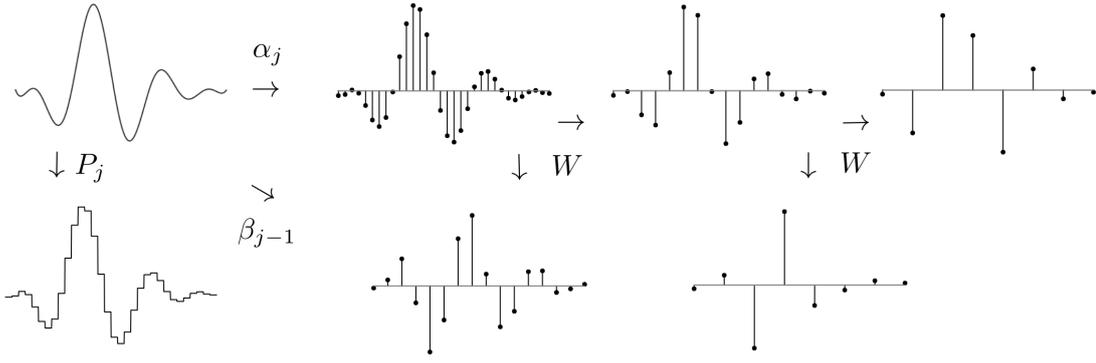
Figure 3.2: Illustration of the various maps defined in Section 3.1.1 in the case of the Haar wavelet.

3.1.2 Periodic wavelets

Given a wavelet ψ on $\mathbb{R}$ with scaling function ϕ and filters g_s and g_w, one can construct a corresponding family of *periodic* wavelet and scaling functions on the circle $\mathbb{S}^1$. Following [Mal08, Section 7.5], we define for $j \geq 0$ and $k = 1,\dots,2^j$ the functions

$$\psi_{j,k}^{\mathrm{per}}(x) = \sum_{m \in \mathbb{Z}} \psi_{j,k}(x + m),$$

$$\phi_{j,k}^{\mathrm{per}}(x) = \sum_{m \in \mathbb{Z}} \phi_{j,k}(x + m)$$

in $L^2(\mathbb{S}^1)$, which again form orthogonal sets of functions. If we set

$$V_j^{\mathrm{per}} = \mathrm{span}\{\phi_{j,k}^{\mathrm{per}} : k = 1,\dots,2^j\}, \qquad W_j^{\mathrm{per}} = \mathrm{span}\{\psi_{j,k}^{\mathrm{per}} : k = 1,\dots,2^j\}$$

then we have

$$\mathbb{C}\mathbf{1} = V_0 \subseteq V_1 \subseteq \cdots \subseteq L^2(\mathbb{S}^1), \quad \overline{\bigcup_{j \geq 0} V_j} = L^2(\mathbb{S}^1), \quad \text{and} \quad V_{j+1} = W_j \oplus V_j.$$

The space V_0 is one-dimensional and consists of the constant functions, i.e., $\phi_{0,1}^{\mathrm{per}}(x) = 1$ for $x \in \mathbb{S}^1$). Thus, the wavelet functions $\{\psi_{j,k}^{\mathrm{per}}\}_{j \geq 0, k=1,\dots,2^j}$ together with $\phi_{0,1}^{\mathrm{per}} = \mathbf{1}$ form an orthonormal basis of $L^2(\mathbb{S}^1)$. Similarly to before, we denote by $\alpha_j^{\mathrm{per}}, \beta_j^{\mathrm{per}} : L^2(\mathbb{S}^1) \to \mathbb{C}^{2^j}$ denote the partial isometries that send a function to its expansion coefficients with respect to the periodized scaling and wavelet basis functions (for fixed j), and we denote by $P_j^{\mathrm{per}} = (\alpha_j^{\mathrm{per}})^* \alpha_j^{\mathrm{per}}$ to be the orthogonal projection onto $V_j \subseteq L^2(\mathbb{S}^1)$.

Since the radius of the circle sets a coarsest length scale, the corresponding filters are now scale-dependent and given by

$$g_{s,j}^{\mathrm{per}}[n] = \sum_{m \in \mathbb{Z}} g_s[n + 2^j m]$$

$$g_{w,j}^{\mathrm{per}}[n] = \sum_{m \in \mathbb{Z}} g_w[n + 2^j m]$$

for $j \geq 0$ and $n = 1,\dots,2^j$. As before, they give rise to unitary maps

$$W^{(\mathscr{L}),\mathrm{per}} : \mathbb{C}^{2^\mathscr{L}} \to \bigoplus_{j=0}^{\mathscr{L}-1} \mathbb{C}^{2^j} \oplus \mathbb{C}, \quad W^{(\mathscr{L}),\mathrm{per}} \alpha_\mathscr{L}^{\mathrm{per}} = \beta_{\mathscr{L}-1}^{\mathrm{per}} \oplus \dots \oplus \beta_0^{\mathrm{per}} \oplus \alpha_0^{\mathrm{per}} \qquad (3.13)$$

that expand a signal at a certain scale into (all) its wavelet coefficients and the remaining scaling coefficient (which is the average of f).

We note that $g_{s,j}^{\text{per}} = g_s$ and $g_{w,j}^{\text{per}} = g_w$ for sufficiently large j (namely when 2^j is at least as large as the cardinality of the filters' supports). This is intuitive since at sufficiently fine scales the periodicity of the circle is no longer visible.

3.2 Biorthogonal wavelets

In Section 3.1 the wavelet functions ψ were such that $\psi_{j,k}$ formed an *orthogonal* basis. The wavelet decomposition operator W, which consisted of filtering and upsampling a signal, was an orthogonal map and was such that it could be inverted by upsampling and filtering by the same filter. These two properties are generalized by *biorthogonal wavelets*. In this case we consider a pair of real-valued sequences $g_s, h_s \in \ell^2(\mathbb{Z})$, as the scaling (or low-pass) filters. We demand that these filters satisfy the *perfect reconstruction condition* on their Fourier transforms, generalizing Eq. (3.2)

$$\hat{g}_s(\theta)\overline{\hat{h}_s(\theta)} + \hat{g}_s(\theta + \pi)\overline{\hat{h}_s(\theta + \pi)} = 2, \tag{3.14}$$

and we define the corresponding wavelet (or high-pass) filters by

$$\hat{g}_w(\theta) = e^{-i\theta}\overline{\hat{h}_s(\theta + \pi)} \quad \text{and} \quad \hat{h}_w(\theta) = e^{-i\theta}\overline{\hat{g}_s(\theta + \pi)}. \tag{3.15}$$

As in the case of the orthogonal discrete wavelet transform these filters can be used to separate a signal $f \in \ell^2(\mathbb{Z})$ into a low-frequency and a high frequency component, and conversely to reconstruct the original signal from these components. For this, we let

$$W_g \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2, \qquad W_g f = (\downarrow m(\overline{\hat{g}_w})f) \oplus (\downarrow m(\overline{\hat{g}_s})f)$$

as in Eq. (3.10) and we similarly define W_h using the filters h_s and h_w in place of g_s and g_w, respectively. Iterating these maps we also obtain $W_g^{(\mathscr{L})}$ and $W_h^{(\mathscr{L})}$. It follows from Eq. (3.14) that f can be reconstructed from its decomposition $W_g f$ by applying the transposed operation $W_h^\top$, so $W_g^{-1} = W_h^\top$. The roles of g and h can be exchanged in this procedure.

3.2.1 Biorthogonal scaling and wavelet functions

Given biorthogonal wavelet filters g, h, the associated scaling functions ϕ^g, ϕ^h are defined in Fourier space for $a = g, h$ by

$$\hat{\phi}^a(\omega) = \prod_{n=1}^{\infty} \frac{\hat{a}_s(2^{-n}\omega)}{\sqrt{2}} \tag{3.16}$$

and the associated *wavelet functions* ψ^g, ψ^h by

$$\hat{\psi}^a(\omega) = \frac{1}{\sqrt{2}}\hat{a}_w\left(\frac{\omega}{2}\right)\hat{\phi}^a\left(\frac{\omega}{2}\right) \tag{3.17}$$

generalizing Eq. (3.6) and Eq. (3.7). Both have compact support if the filters are finite; an example is shown in Fig. 3.5. Again, we can define rescaled and shifted versions

$$\psi^a_{j,k}(x) = 2^{-\frac{j}{2}} \psi^a(2^{-j}x - k),$$

$$\phi^a_{j,k}(x) = 2^{-\frac{j}{2}} \phi^a(2^{-j}x - k).$$

It then follows that the sets $\{\psi^g_{j,k}\}_{j,k\in\mathbb{Z}}$ and $\{\psi^h_{j,k}\}_{j,k\in\mathbb{Z}}$ form a dual basis pair for $L^2(\mathbb{R})$, in the sense that

$$\langle \psi^g_{j,k}, \psi^h_{j',k'} \rangle = \delta_{j,j'} \delta_{k,k'}.$$

Moreover,

$$\langle \phi^g_{j,k}, \phi^h_{j,k'} \rangle = \delta_{k,k'}.$$

If the filters are finite and the scaling functions are square-integrable functions (which is closely related to the discrete wavelet decomposition being sufficiently stable) the sets $\{\psi^g_{j,k}\}_{j,k\in\mathbb{Z}}$ and $\{\psi^h_{j,k}\}_{j,k\in\mathbb{Z}}$ form a Riesz basis of $L^2(\mathbb{R})$ [CDF92]. This means that we can write any function $f \in L^2(\mathbb{R})$ as

$$f = \sum_{j,k} \langle \psi^g_{j,k}, f \rangle \psi^h_{j,k} = \sum_{j,k} \langle \psi^h_{j,k}, f \rangle \psi^g_{j,k}.$$

By construction of the scaling and wavelet functions, these are such that if

$$f = \sum_n s[n] \phi^g_{0,n}$$

then we can rewrite

$$f = \sum_{l=0}^{\mathscr{L}-1} \sum_n w[j,k] \psi^g_{j,k} + \sum_k \tilde{s}[k] \phi^g_{\mathscr{L},k} \tag{3.18}$$

where we find the coefficients $w[j,k]$ and $\tilde{s}[k]$ precisely by applying the discrete wavelet transformation W_h to the signal s.

3.3 Wavelet pairs with phase relations

In order to obtain entanglement renormalization circuits from wavelet transforms we will need wavelet filters which satisfy certain relations. For fermionic entanglement renormalization, we need to find a pair of orthogonal wavelets with wavelet filters g_w and h_w which are such that

$$\hat{h}_w(\theta) \approx -i \operatorname{sgn}(\theta) e^{i\frac{\theta}{2}} \hat{g}_w(\theta) \tag{3.19}$$

for $\theta \in (-\pi, \pi)$ as we will explain in Chapter 4. For definiteness we choose the convention that $\operatorname{sgn}(0) = 1$, but note that here this choice is inconsequential, as $\hat{h}_w(0) = \hat{g}_w(0) = 0$. Such a pair of wavelets is called an approximate Hilbert pair.

For biorthogonal wavelets, given a 2π-periodic function $E(\theta)$ we would like to find a pair filters g_w and h_w which are such that they form a biorthogonal pair of wavelet filters and

$$\hat{g}_w(\theta) \approx E(\theta) \hat{h}_w(\theta). \tag{3.20}$$

The function $E(\theta)$ will have the interpretation of the dispersion relation of a bosonic Hamiltonian. We will start by discussing in detail the notion of an approximate Hilbert pair. In Section 3.4 we provide an explicit construction method to find filters which satisfy Eq. (3.19) and Eq. (3.20).

3.3.1 Approximate Hilbert pairs

The construction of fermionic entanglement renormalization circuits is based on approximating the *Hilbert transform* using wavelets, which is closely related to the phase relation Eq. (3.19). The Hilbert transform is the unitary operator on $L^2(\mathbb{R})$ defined by

$$\widehat{\mathcal{H}f}(\omega) = -i\,\mathrm{sgn}(\omega)\hat{f}(\omega).$$

We will need a pair of wavelet and scaling filters g_w, g_s and h_w, h_s (note that this refers two *orthogonal* wavelets, and not to a biorthogonal pair of wavelets) such that the associated wavelet functions ψ^g and ψ^h satisfy

$$\psi^h = \mathcal{H}\psi^g. \tag{3.21}$$

Such a pair of wavelets is called a *Hilbert pair,* and this condition is equivalent to the filter condition in Eq. (3.19). We can also formulate this as a condition on the scaling filters, which leads to the following two equivalent conditions to generate a Hilbert pair [Sel01]

$$\hat{h}_s(\theta) = \mu_s(\theta)\hat{g}_s(\theta), \tag{3.22}$$
$$\hat{h}_w(\theta) = \mu_w(\theta)\hat{g}_w(\theta),$$

where μ_s and μ_w are periodic functions in $L^\infty(\mathbb{R}/2\pi\mathbb{Z})$ defined by

$$\mu_s(\theta) = e^{-i\frac{\theta}{2}}, \tag{3.23}$$
$$\mu_w(\theta) = -i\,\mathrm{sgn}(\theta)e^{i\frac{\theta}{2}}$$

for $|\theta| < \pi$. Note that there is a discontinuity at $\theta = \pm\pi$ (but it suffices to define μ_s and μ_w almost everywhere). In this situation, Eqs. (3.6) and (3.7) implies that the scaling functions ϕ_g and ϕ_h will be related by

$$\hat{\phi}^h(\omega) = \lambda_s(\omega)\hat{\phi}^g(\omega), \tag{3.24}$$

where $\lambda_s \in L^\infty(\mathbb{R})$ is defined by

$$\lambda_s(\omega) = -i\,\mathrm{sgn}(\omega)\mu_w(-\omega) \tag{3.25}$$

and we have extended μ_w to a periodic function in $L^\infty(\mathbb{R})$. We refer to [Sel01, Sel02] for further detail. Since the Hilbert transform does not preserve compact support, we can not hope for exact Hilbert pair wavelets using compactly supported wavelets. However, an approximate version can be realized. The following definition describes the notion of approximation that is appropriate in our context.

Definition 3.1. An ε-*approximate Hilbert pair* consists of a pair of wavelet and scaling filters, g_w, g_s, h_w, h_s, with corresponding wavelet functions ψ^g, ψ^h and scaling functions ϕ^g, ϕ^h, such that

$$\|\hat{h}_s - \mu_s\hat{g}_s\|_\infty \leq \varepsilon. \tag{3.26}$$

That is, the error in the phase relation (3.22) for the scaling filters is bounded by ε. This condition can be checked numerically. In Section 3.6.1 we will show that this condition translates into a good approximation of the Hilbert pair relation in Eq. (3.21). If we periodize an (approximate) Hilbert pair as described in Section 3.1.2, we get periodic wavelets that are (approximately) related by the Hilbert transform on the circle.

We can also formalize Eq. (3.20). Let $\varepsilon > 0$ and $E\colon (-\pi, \pi) \to \mathbb{C}$.

Definition 3.2. An ε-*approximate E-dispersion pair* consists of a biorthogonal wavelet and scaling filters, g_w, g_s, h_w, h_s, with corresponding wavelet functions ψ^g, ψ^h and scaling functions ϕ^g, ϕ^h, such that

$$\|\hat{h}_s - E\hat{g}_s\|_\infty \leq \varepsilon. \tag{3.27}$$

That is, the error in the dispersion relation (3.20) for the wavelet filters is bounded by ε.

3.4   Construction of filters

We will explain how to construct filter pairs that yield a good approximation of a given phase relation as in Section 3.3. We start with a construction due to Selesnick [Sel02, Sel01], who introduced the notion of an approximate Hilbert pair. His construction depends on two parameters, K and L, where K is the number of vanishing moments of the wavelets (relevant for the approximation power of the wavelet decomposition and for the smoothness of the wavelets) and where L is essentially the number of terms in a Taylor expansion of the relation in Eq. (3.22) at $\theta = 0$. By construction, the filters are real and have finite length $N = 2(K + L)$, so the wavelet and scaling functions are compactly supported on intervals of width N. In this construction, we let $d[n]$ be a so-called maximally flat all-pass filter with delay $\frac{1}{4}$ of degree L, given by

$$d[n] = (-1)^n \binom{n}{L} \prod_{k=1}^{n-1} \frac{\frac{1}{2} - L + k}{\frac{3}{2} + k}$$

for $n = 1, \ldots, L$. This is such that $e^{-iL\theta}\hat{d}(-\theta)/\hat{d}(\theta) \approx e^{-i\frac{\theta}{2}} = \mu_s(\theta)$ for $\theta \in (-\pi, \pi)$ where the approximation is accurate around $\theta = 0$. The idea is now that in order to get $\hat{h}_s(\theta) = \mu_s(\theta)\hat{g}_s(\theta)$, we impose the relation

$$\hat{d}(\theta)\hat{h}_s(\theta) = e^{-iL\theta}\hat{d}(-\theta)\hat{g}_s(\theta)$$

This is ensured if we make the following ansatz for the filters

$$\hat{h}_s(\theta) = e^{-iL\theta}\hat{d}(-\theta)(1 + e^{i\theta})^K \hat{f}(\theta)$$
$$\hat{g}_s(\theta) = \hat{d}(\theta)(1 + e^{i\theta})^K \hat{f}(\theta)$$

where $\hat{f}$ is the Fourier transform of a finite sequence which needs to be determined. The parameter K determines the number of vanishing moments of the biorthogonal wavelets, just as in the Daubechies wavelet construction. We need to solve for f such that the resulting filters are indeed wavelet filters. The condition in Eq. (3.2) translates to

$$\hat{s}(\theta)\hat{f}(\theta)\hat{f}(-\theta) + \hat{s}(\theta + \pi)\hat{f}(\theta + \pi)\hat{f}(\pi - \theta) = 2$$
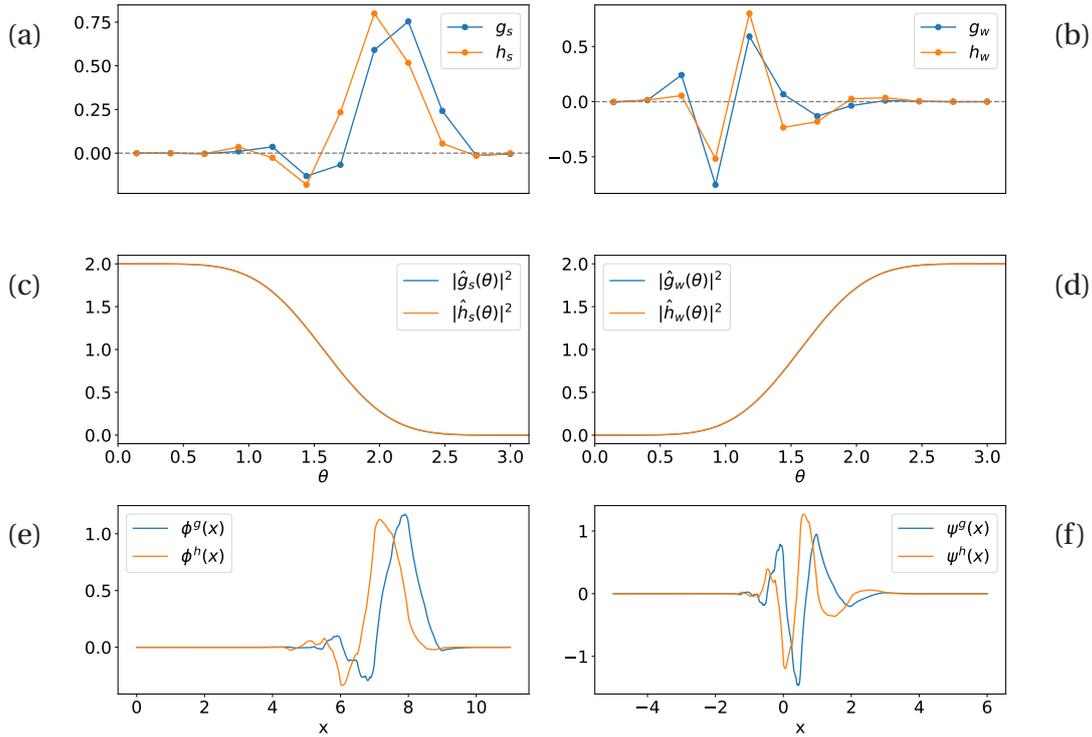
Figure 3.3: The results of using $K = 2$, $L = 4$ in the Selesnick construction: (a) scaling filters g_s and h_s, (b) wavelet filters g_w and h_w, (c) absolute value squared of the Fourier transforms of the scaling filters $|g_s(\theta)|^2$ and $|h_s(\theta)|^2$, (d) absolute value squared of the Fourier transforms of the wavelet filters $|g_w(\theta)|^2$ and $|h_w(\theta)|^2$, (e) scaling functions ϕ^g and ϕ^h, and (f) wavelet functions ψ^g and ψ^h.

where $s(\theta) = \hat{d}(\theta)\hat{d}(-\theta)(2\cos(\frac{\theta}{2}))^{2K}$. We then try to solve this by letting $\hat{r}(\theta) = \hat{f}(\theta)\hat{f}(-\theta)$ and take r to be a solution to

$$\hat{s}(\theta)\hat{r}(\theta) + \hat{s}(\theta + \pi)\hat{r}(\theta + \pi) = 2$$

which is equivalent to the linear system

$$\sum_l s[2n - l]r[l] = \delta_0[n].$$

Finally, if possible, we perform a spectral factorization $\hat{r}(\theta) = \hat{f}(\theta)\hat{f}(-\theta)$. A necessary and sufficient condition for this is that $\hat{r}(\theta) \geq 0$ for all θ. Unfortunately, there is no guarantee for this (although there is upon a small modification of the procedure [ACGR20]).

Numerically, one can see that Selesnick's construction described above produces ε-approximate Hilbert pairs where ε decays exponentially with $\min\{K, L\}$ [HSW+18]. In Section 3.6 we will encounter various other relevant parameters for error bounds which depend on the wavelet filters. It can be seen numerically that these parameters remain bounded or increase only slowly, as shown in Table 3.1. For a more extensive treatment of the theory of approximate Hilbert pairs, see [SBK05, YO05, CU09, CU10] and the more recent work [ACGR20]. See Fig. 3.3 for an illustration of Selesnick's wavelets with parameters $K = L = 2$ and Fig. 3.4 for a numerical evaluation of the approximation errors.
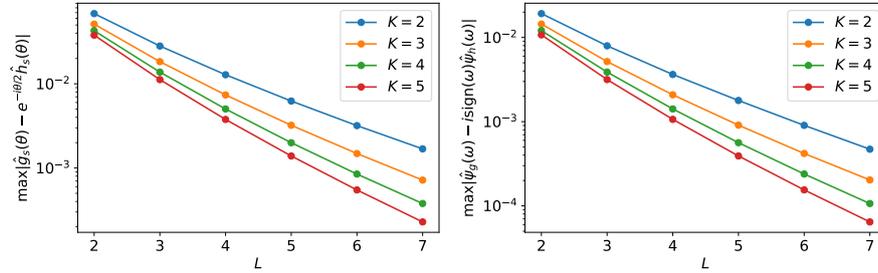
Figure 3.4: Approximation errors for $\varepsilon = \max_\theta |g_w(\theta) - |\sin(\frac{\theta}{2})| h_w(\theta)|$ and $\max_\theta |\psi^g(\theta) - \frac{|\theta|}{4}\psi^h(\theta)|$ for different values of K and L for approximate Hilbert pairs arising from the Selesnick construction. For fixed K the error appears to decrease exponentially in L.

K	L	N	ε	C_{UV}	C_{IR}	C_χ	C'_χ	C_ϕ
1	1	4	0.264099	0.619741	2.542073	1.166423	1.142220	1.254999
2	2	8	0.068221	0.622182	1.217454	1.155488	0.295133	2.296890
3	3	12	0.018338	0.624782	1.190944	1.154757	0.079283	2.116091
4	4	16	0.005020	0.626782	1.150151	1.154705	0.021691	1.251461
5	5	20	0.001389	0.628374	1.130260	1.154701	0.005999	2.120782
6	6	24	0.000387	0.629686	1.120354	1.154701	0.001671	2.106891
7	7	28	0.000108	0.630795	1.114293	1.154701	0.000468	1.234832
8	8	32	0.000030	0.631752	1.108135	1.154701	0.000132	2.434899
9	9	36	0.000009	0.632674	1.106718	1.154701	0.000037	1.923738
10	10	40	0.000003	0.638023	1.440101	1.154701	0.000011	5.752427

Table 3.1: Numerical values of various constants for Selesnick's approximate Hilbert pairs with parameters $K = L$. It appears that ε decays exponentially with increasing $K = L$, while the other constants from Lemma 3.3, Lemma 3.4, Lemma 3.5 and Lemma 3.8 are well-behaved.

A straightforward generalization of this procedure allows us to construct an ε-approximate E-dispersion pair of biorthogonal filters as in Definition 3.2, that is, such that Eq. (3.20) holds, for a broad class of functions $E(\theta)$. Let us assume that $E(\theta) = E(-\theta)$, and $E(\theta)$ is real-valued. We would like to construct a *biorthogonal filter pair* (g, h) such that

$$\hat{g}_w(\theta) \approx E(\theta) h_w(\theta) \tag{3.28}$$

or equivalently

$$\hat{h}_s(\theta) \approx E(\theta + \pi)\hat{g}_s(\theta) \tag{3.29}$$

To modify the procedure above, we need an arbitrary rational approximation

$$E(\theta + \pi) \approx \frac{\hat{a}(\theta)}{\hat{b}(\theta)},$$

where a and b are real finite symmetric sequences on $[-L, L]$. The approximation only has to be accurate around $\theta = 0$. We then make the following ansatz for the Fourier transform of the scaling filters

$$\begin{aligned} \hat{g}_s(\theta) &= \hat{b}(k)(1 + e^{i\theta})^K \hat{f}(\theta), \\ \hat{h}_s(\theta) &= \hat{a}(k)(1 + e^{i\theta})^K \hat{f}(\theta) \end{aligned} \tag{3.30}$$

where again $\hat{f}(\theta)$ is the Fourier transform of a real finite sequence $f[n]$ that still needs to be determined and K is the number of vanishing moments. By construction, $\hat{g}_s(\theta)$ and $\hat{h}_s(\theta)$ are small near $\theta = \pi$, and Eq. (3.29) is satisfied. In order for Eq. (3.30) to generate biorthogonal wavelet filters, they need to satisfy the condition in Eq. (3.14) which now translates to

$$\hat{s}(\theta)\hat{f}(\theta)\hat{f}(-\theta) + \hat{s}(\theta + \pi)\hat{f}(\theta + \pi)\hat{f}(\pi - \theta) = 2$$

where $\hat{s}(\theta) = \hat{a}(\theta)\hat{b}(\theta)(2\cos(\frac{\theta}{2}))^{2K}$. One can try to solve this in exactly the same way as before by introducing $\hat{r}(\theta) = \hat{f}(\theta)\hat{f}(-\theta)$, solving

$$\sum_l s[2n - l]r[l] = \delta_0[n]$$

and, if possible, performing a spectral factorization to obtain f. Unfortunately, we do not know of a condition on the product $\hat{a}\hat{b}$ that guarantees the existence of a solution. If a solution exists, the resulting filters (g, h) will have support of size $2N$ where $N = K + 2L$.

One particularly useful phase relation is when $E(\theta) = |\sin(\frac{\theta}{2})|$ (which will correspond to a harmonic massless bosonic chain). In this case, an appropriate choice for a and b is given by

$$\begin{aligned} \hat{a}(\theta) &= \frac{1}{2}(e^{iL\theta}\hat{d}(\theta)^2 + e^{-iL\theta}\hat{d}(-\theta)^2), \\ \hat{b}(\theta) &= \hat{d}(\theta)\hat{d}(-\theta). \end{aligned} \tag{3.31}$$

where $d[n]$ is as in the Selesnick construction. Indeed, we find that

$$\frac{\hat{a}(\theta)}{\hat{b}(\theta)} = \frac{1}{2}\left(e^{iL\theta}\frac{\hat{d}(\theta)}{\hat{d}(-\theta)} + e^{-iL\theta}\frac{\hat{d}(-\theta)}{\hat{d}(\theta)}\right) \approx \frac{1}{2}(e^{i\frac{\theta}{2}} + e^{-i\frac{\theta}{2}}) = \cos(\frac{\theta}{2})$$

K	L	N	B	ε	κ
1	1	6	5.508287	0.319609	1.505577
2	2	12	1.587766	0.060920	1.459265
3	3	18	1.321713	0.016613	1.490601
4	4	24	1.282064	0.004152	1.501892
5	5	30	1.211855	0.001087	1.513364
6	6	36	1.090349	0.000282	1.522019
7	7	42	1.853832	0.000063	1.596999

Table 3.2: Numerical values of various constants for the construction of biorthogonal wavelet filters in Eq. (3.31) with parameters $K = L$. Here $\varepsilon = \max_\theta |g_w(\theta) - |\sin(\frac{\theta}{2})|h_w(\theta)|$. The parameters $B = \max\{\|\phi^g\|_\infty, \|\phi^h\|_\infty\}$ and $\kappa = \max\{\|W_g\|, \|W_h\|\}$ are relevant for Theorem 5.3.

for $\theta \in (-\pi, \pi)$. In Fig. 3.6 we show the goodness of the approximation in Eqs. (3.28) and (3.29) as a function of K and L. The resulting filters and wavelets for $K = 2, L = 4$ are shown in Fig. 3.5. We remark that the construction in Eq. (3.31) is not necessarily optimal. From numerical evidence in Fig. 3.6 it again appears that the accuracy of the approximation improves exponentially with increasing support. An interesting open problem is to rigorously prove the existence of approximate solutions to Eq. (3.28) with (exponentially) improving approximation accuracy as the filter size increases.

3.5 Construction of linear circuits from filters

The construction of entanglement renormalization circuits will be by 'second quantizing' a single-particle circuit for the wavelet transforms, as will be explained in Chapter 4 and Chapter 5. We now explain how to explicitly construct a single-particle circuit for a (biorthogonal) wavelet transform. We emphasize that the type of circuit we construct is not a quantum circuit but will have the interpretation of being a circuit in the single-particle picture. By a single-particle nearest neighbour circuit of depth N we mean a composition $A_N \circ \ldots \circ A_1$ for a sequence of maps $A_1, \ldots A_N$ on $\ell^2(\mathbb{Z})$ which are such that

$$A_i = \bigoplus_{n=\text{even}} (a_i)_{n,n+1} = \begin{pmatrix} \ddots & & & \\ & (a_i)_{n,n+1} & 0 & \\ & 0 & (a_i)_{n+2,n+3} & \\ & & & \ddots \end{pmatrix}$$

for i even, and similarly a sum over odd terms if i is odd. Here each $(a_i)_{n,n+1}$ is a two by two matrix, acting on sites n and $n+1$. In other words, each layer of the circuit A_i is a block diagonal map, and the circuit acts as a linear map with a *direct sum* structure rather than a *tensor product* structure.

Motivated by the fermionic setting it has been extensively discussed in [EW18] how to construct *unitary* local circuits from *orthogonal* wavelet filters (so each A_i is unitary). The construction for biorthogonal wavelet filters is very similar and the symmetric case has already been discussed in [EW18], but for completeness we provide it here. Given a pair of biorthogonal filters (g_s, g_w) and (h_s, h_w) (in particular we could have $g = h$

Figure 3.5: The results of using $K = 2$, $L = 4$ in the construction of Eq. (3.31): (a) scaling filters g_s and h_s, (b) wavelet filters g_w and h_w, (c) absolute value squared of the Fourier transforms of the scaling filters $|g_s(\theta)|^2$ and $|h_s(\theta)|^2$, (d) absolute value squared of the Fourier transforms of the wavelet filters $|g_w(\theta)|^2$ and $|h_w(\theta)|^2$, (e) scaling functions ϕ^g and ϕ^h, and (f) wavelet functions ψ^g and ψ^h.
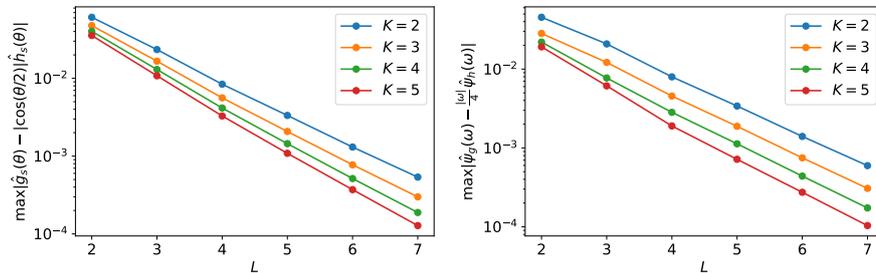
Figure 3.6: Approximation errors for $\varepsilon = \max_\theta |g_w(\theta) - |\sin(\frac{\theta}{2})| h_w(\theta)|$ and $\max_\theta |\psi^g(\theta) - \frac{|\theta|}{4} \psi^h(\theta)|$ for different values of K and L for a filter pair constructed using Eq. (3.31). For fixed K the error appears to decrease exponentially in L.

in which case we have an orthogonal wavelet filter) of support $2N$ we will construct a binary circuit of depth N that implements the wavelet decomposition map. We will assume that g_s and h_s are supported on $[-N+1, N]$, which we can always achieve by a shift. We let $W_g : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{C}^2)$ denote a single layer of a discrete wavelet transform, defined as in Eq. (3.10). By putting the scaling and wavelet outputs on the even and odd sublattice, respectively, we obtain a unitary

$$W' : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}), \quad W'_g := \iota W_g,$$

where

$$\iota : \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \to \ell^2(\mathbb{Z})$$
$$\iota(f_w \oplus f_s)[2n] = f_w[n],$$
$$\iota(f_w \oplus f_s)[2n+1] = f_s[n]$$

for all $f_w, f_s \in \ell^2(\mathbb{Z})$.

We would like to construct a circuit $A = A_N \circ \ldots \circ A_1$ which implements the wavelet reconstruction map in the sense that $A = (W'_g)^\mathsf{T}$ and $(A^\mathsf{T})^{-1} = (W'_h)^\mathsf{T}$, as in Fig. 5.1. By shift invariance this is equivalent to

$$
\begin{aligned}
A\delta_1 &= g_s \\
A\delta_2 &= g_w \\
(A^\mathsf{T})^{-1}\delta_1 &= h_s \\
(A^\mathsf{T})^{-1}\delta_2 &= h_w
\end{aligned}
\tag{3.32}
$$

as illustrated for g in Fig. 3.7. So, we need construct the matrices a_i given the filters g and h. We will need the perfect reconstruction condition Eq. (3.14) which becomes

$$\sum_l g_s[2n+l]h_s[l] = \delta_0[n]$$

upon applying the inverse Fourier transform. In particular, the vectors

$$(g_s[-N+1], g_s[-N+2])^\mathsf{T} \text{ and } (h_s[N-1], h_s[N])^\mathsf{T}$$

are orthogonal, and so are

$$(g_s[N-1], g_s[N])^\mathsf{T} \text{ and } (h_s[-N+1], h_s[-N+2])^\mathsf{T}.$$

Furthermore we will use that the wavelet filters are derived from the scaling filters as

$$
\begin{aligned}
g_w[n] &= (-1)^{(1-n)} h_s[1-n] \\
h_w[n] &= (-1)^{(1-n)} g_s[1-n]
\end{aligned}
\tag{3.33}
$$

which follows directly from Eq. (3.15). First suppose that $N = 1$. In that case we let

$$a_1 = \begin{pmatrix} g_s[0] & g_w[0] \\ g_s[1] & g_w[1] \end{pmatrix}.$$

Using the perfect reconstruction condition we may now check that

$$(a_1^\mathsf{T})^{-1} = \begin{pmatrix} h_s[0] & h_w[0] \\ h_s[1] & h_w[1] \end{pmatrix}$$
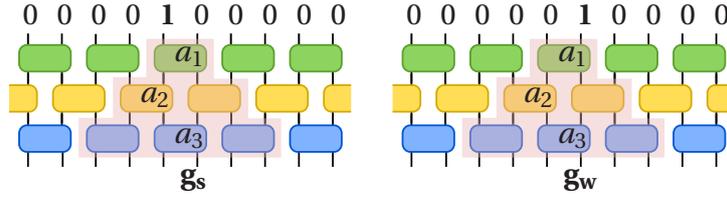
Figure 3.7: Illustration of Eq. (3.32), which gives the equations the a_i have to satisfy in order for the circuit to implement W_g.

so this satisfies Eq. (3.32). For $N > 1$ we will construct the A_i recursively. Let

$$g_N = \begin{pmatrix} g_s[N-1] & g_s[-N+2] \\ g_s[N] & g_s[-N+1] \end{pmatrix}$$

$$a_N = \frac{1}{\sqrt{\det(g_N)}} g_N$$

then it is clear that A_N^{-1} maps g_s to a sequence $g_s^{(N-1)}$ on $[-N+2, N-1]$ and A_N^{T} maps h_s to a sequence $h_s^{(N-1)}$ on $[-N+2, N-1]$ using the orthogonality properties derived from the perfect filter condition. In the non-generic degenerate case that $\det(g_N) = 0$, the size of the support can only be decreased by 1 and an additional layer is needed. Moreover, since A_N is invariant under shifts of 2, it is easy to see that $g_s^{(N-1)}$ and $h_s^{(N-1)}$ still satisfy the perfect reconstruction property. Finally, if we let α denote the map defined by $\alpha x[n] = (-1)^{(1-n)} x[1-n]$, then in order to see that A_N^{-1} maps g_w to the wavelet filter $g_w^{(N-1)}$ defined by $\alpha h_s^{(N-1)}$, it suffices to check that $A_N^{-1} \alpha = \alpha A_N^{\mathsf{T}}$ or equivalently $A_N^{-1} = \alpha A_N^{\mathsf{T}} \alpha$. This follows from the inversion formula for two by two matrices with determinant 1, i.e.,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Now we can recursively apply the same procedure to the filters $(g^{(N-1)}, h^{(N-1)})$ to construct $A_{N-1}, \ldots, A_1$. We have now seen that we can construct a nearest-neighbor single-particle circuit of depth N from a biorthogonal filter pair.

Conversely, given a nearest-neighbor single-particle circuit of depth N, which can be written as $A = A_N \circ \ldots \circ A_1$ where each A_i is invertible, we may *define* filters g and h by Eq. (3.32). We can then check that these filters form perfect reconstruction filters, in the sense that $W_h^{\mathsf{T}} = W_g^{-1}$. If we assume $\det(a_i) = 1$ for $i = 1, \ldots, N$, the wavelet and scaling filters are related as in Eq. (3.33).

3.6 Wavelet approximations

An important question in wavelet theory is how accurately a function f is approximated if it is truncated to a finite number of scales. In this section we will state three results that give quantitative bounds assuming that the wavelets are compactly supported and bounded. The results are standard, but we provide proofs in Section 3.6.2, in particular to keep careful track of the constants in the bounds. Using an argument from Fourier analysis in Lemma 3.10 we show in Lemma 3.3 an approximation result for a 'UV cut-off' for a sufficiently smooth f, where we discard all detail at fine scales, or

alternatively in Lemma 3.4, if we sample f on a grid. Next we show in Lemma 3.5 that for compactly supported functions we can also discard large scale wavelet components up to a small error, which should be thought of as an 'IR cut-off'. In this section we restrict to orthogonal wavelets. Under additional assumptions similar results are true for biorthogonal wavelets. In Section 3.6.1 we explain how the error in the filter relation for approximate Hilbert pair wavelets impacts other approximation properties.

We know that the shifted and rescaled wavelet functions form an orthonormal basis for $L^2(\mathbb{R})$, so we can write

$$f = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} (\beta_j f)[k] \psi_{j,k}.$$

We will be interested in approximating the function f over a finite number of scales, that is, we will try to approximate f by

$$f \approx \sum_{j=j_0}^{j_1} \sum_{k \in \mathbb{Z}} (\beta_j f)[k] \psi_{j,k}.$$

Recall that $j \gg 0$ corresponds to small scale structures in f, and $j \ll 0$ to large scale structure. In this section we will give quantitative bounds which show that one can approximate f to good precision using a finite number of scales if f is sufficiently smooth and compactly supported. To be precise, in Lemma 3.3 we show that large j can be discarded with small error if f is sufficiently smooth (so this bounds the error incurred by leaving out detail), and in Lemma 3.5 we show that we can truncate in the other direction if f is compactly supported (bounding the error of leaving out large scale structure).

We start with the first result (cf. Theorem 7.6 in [SN96]), corresponding to a UV cut-off. Recall that the Sobolev spaces $H^K(\mathbb{R})$ and $H^K(\mathbb{S}^1)$ consist of functions f with square-integrable weak K-th derivative, denoted $f^{(K)}$.

Lemma 3.3 (UV cut-off). *Assume that the Fourier transform of the scaling filter $\hat{g}_s(\theta)$ has a zero of order K at $\theta = \pi$. Then there exists a constant C_{UV} such that for every $f \in H^K(\mathbb{R})$ and $j \in \mathbb{Z}$, we have that*

$$\|P_j f - f\| \le 2^{-Kj} C_{\mathrm{UV}} \|f^{(K)}\|.$$

Similarly, for every $f \in H^1(\mathbb{S}^1)$ and $j \ge 0$, we have that

$$\|P_j^{\mathrm{per}} f - f\| \le 2^{-Kj} C_{\mathrm{UV}} \|f^{(K)}\|.$$

If the scaling filter is supported in $\{0, \dots, N-1\}$, we may take $K = 1$ and $C_{\mathrm{UV}} \le 2N^2$.

Under mild technical conditions the 'UV cut-off' from Lemma 3.3 can be well-approximated by sampling the function on a dyadic grid, as shown in the following lemma (see for instance Theorem 7.22 in [Mal08]).

Lemma 3.4 (Sampling error). *There exists a constant C_ϕ such that the following holds: For every $f \in H^1(\mathbb{R})$ and f_j the sequence defined by $(f_j)_k := 2^{-j/2} f(2^{-j} k)$ for $k \in \mathbb{Z}$ (we identify f with its unique representative as a continuous function), we have*

$$\|\alpha_j f - f_j\| \le 2^{-j} C_\phi \|f'\|.$$

Likewise, for every $f \in H^1(\mathbb{S}^1)$ and $f_j \in \mathbb{C}^{2^j}$ the vector with components

$$(f_j)_k := 2^{-j/2} f(2^{-j} k),$$

we have

$$\|\alpha_j^{\mathrm{per}} f - f_j\| \le 2^{-j} C_\phi \|f'\|.$$

If the scaling filter is supported in $\{0, \dots, N-1\}$, then these estimates hold with $C_\phi \le 2N^2$.

The final lemma of this section (adapted from [Woj97]) bounds the error incurred by leaving out coarse scale components from compactly supported functions, corresponding to an IR cut-off.

Lemma 3.5 (IR cut-off). *Assume that the scaling function ϕ satisfies*

$$\sqrt{\sum_{k \in \mathbb{Z}} |\phi(y-k)|^2} \le C_{\mathrm{IR}}$$

for all $y \in \mathbb{R}$. Then for every $f \in L^2(\mathbb{R})$ with compact support,

$$\|P_j f\| \le 2^{j/2} \sqrt{D(f)} C_{\mathrm{IR}} \|f\|$$

where $D(f)$ is the size of the smallest interval containing the support of f. If ϕ is bounded and supported in an interval of width N, we have $C_{\mathrm{IR}} \le \sqrt{N} \|\phi\|_\infty$.

3.6.1 Approximate Hilbert pair wavelets

The following lemma is an improved version of [HSW$^+$18, (A7)]. It controls the error incurred by using approximate instead of exact Hilbert pairs both on the line and on the circle.

Lemma 3.6. *Consider an ε-approximate Hilbert pair. Let $W_g^{(\mathscr{L})}$ and $W_h^{(\mathscr{L})}$ denote the corresponding wavelet transforms for $\mathscr{L}$ layers, defined as in Eqs. (3.10) and (3.11) using the filters g and h, respectively. Then,*

$$\|P_w\big(W_g^{(\mathscr{L})} - W_h^{(\mathscr{L})} m(\mu_w)\big)\| \le \varepsilon \mathscr{L}, \tag{3.34}$$

$$\|P_s\big(W_g^{(\mathscr{L})} - W_h^{(\mathscr{L})} m(\mu_w)\big)f\| \le \varepsilon \mathscr{L} \|f\| + 2\|P_s W_g^{(\mathscr{L})} f\| \qquad (\forall f \in \ell^2(\mathbb{Z})), \tag{3.35}$$

where $P_w = I_{\ell^2(\mathbb{Z})} \otimes \sum_{k=0}^{\mathscr{L}-1} |k\rangle \langle k|$ denotes the projection onto the wavelet coefficients and $P_s = I - P_w$ the projection onto the scaling coefficients.

Proof. As in Eq. (3.10), denote by $W_g, W_h \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$ the unitaries corresponding to a single layer of the wavelet transform:

$$W_g = (\downarrow m(\overline{\hat{g}_w})) \oplus (\downarrow m(\overline{\hat{g}_s})) \quad \text{and} \quad W_h = (\downarrow m(\overline{\hat{h}_w})) \oplus (\downarrow m(\overline{\hat{h}_s}))$$

One may easily verify the relation

$$m(\mu_w) \downarrow m(\mu_s) = \downarrow m(\mu_w).$$

This allows us to rewrite

$$
\begin{aligned}
W_h m(\mu_w) &= (\downarrow m(\overline{\hat{h}_w})) \oplus (\downarrow m(\overline{\hat{h}_s})) m(\mu_w) \\
&= (\downarrow m(\mu_w \overline{\hat{h}_w})) \oplus (\downarrow m(\mu_w) m(\overline{\hat{h}_s})) \\
&= (\downarrow m(\mu_w \overline{\hat{h}_w})) \oplus (m(\mu_w) \downarrow m(\mu_s \overline{\hat{h}_s})) \\
&= (I_{\ell^2(\mathbb{Z})} \oplus m(\mu_w)) \tilde{W}_h,
\end{aligned}
\tag{3.36}
$$

where we introduced

$$
\tilde{W}_h := (\downarrow m(\mu_w \overline{\hat{h}_w})) \oplus (\downarrow m(\mu_s \overline{\hat{h}_s})).
$$

Now consider $\mathscr{L}$ layers of the transform. For $l = 1, \dots, \mathscr{L}$, define $W_g^l := I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{l-1}} \oplus W_g$ and similarly W_h^l and $\tilde{W}_h^l$, so that $W_g^{(\mathscr{L})} = W_g^{\mathscr{L}} \cdots W_g^1$ etc. By using Eq. (3.36), we find that

$$
W_h^{(\mathscr{L})} m(\mu_w) = W_h^{\mathscr{L}} \cdots W_h^1 m(\mu_w) = \left(I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}}} \oplus m(\mu_w) \right) \tilde{W}_h^{\mathscr{L}} \cdots \tilde{W}_h^1
$$

Our assumption (3.26) on the scaling filter error in an approximate Hilbert pair implies that, for all l,

$$
\| W_g^l - \tilde{W}_h^l \| = \| W_g - \tilde{W}_h \| \le \max\{ \| \overline{\hat{g}_s} - \mu_s \overline{\hat{h}_s} \|_\infty, \| \overline{\hat{g}_w} - \mu_w \overline{\hat{h}_w} \|_\infty \} \le \varepsilon.
\tag{3.37}
$$

Next we write a telescoping sum

$$
\begin{aligned}
W_g^{(\mathscr{L})} - W_h^{(\mathscr{L})} m(\mu_w) &= W_g^{\mathscr{L}} \cdots W_g^1 - \left(I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}}} \oplus m(\mu_w) \right) \tilde{W}_h^{\mathscr{L}} \cdots \tilde{W}_h^1 \\
&= \left(I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}}} \oplus m(\mu_w) \right) \left(W_g^{\mathscr{L}} \cdots W_g^1 - \tilde{W}_h^{\mathscr{L}} \cdots \tilde{W}_h^1 \right) \\
&\quad + \left(0_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}}} \oplus (I - m(\mu_w)) \right) W_g^{\mathscr{L}} \cdots W_g^1 \\
&= \left(I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}}} \oplus m(\mu_w) \right) \sum_{l=1}^{\mathscr{L}} W_g^{\mathscr{L}} \cdots W_g^{l+1} \left(W_g^l - \tilde{W}_h^l \right) \tilde{W}_h^{l-1} \cdots \tilde{W}_h^1 \\
&\quad + \left(0 \oplus (I - m(\mu_w)) \right) W_g^{\mathscr{L}} \cdots W_g^1
\end{aligned}
$$

Using Eq. (3.37) and the fact that $\| W_g^l \| = \| \tilde{W}_h^l \| = 1$ for all l, we can therefore bound

$$
\| P_w \big(W_g^{(\mathscr{L})} - W_h^{(\mathscr{L})} m(\mu_w) \big) \| \le \sum_{l=1}^{\mathscr{L}} \| W_g^l - \tilde{W}_h^l \| \le \varepsilon \mathscr{L}
$$

and, since furthermore $\| m(\mu_w) \| = 1$,

$$
\begin{aligned}
\| P_s \big(W_g^{(\mathscr{L})} - W_h^{(\mathscr{L})} m(\mu_w) \big) f \| &\le \sum_{l=1}^{\mathscr{L}} \| W_g^l - \tilde{W}_h^l \| \| f \| + 2 \| P_s W_g^{\mathscr{L}} \cdots W_g^1 f \| \\
&\le \varepsilon \mathscr{L} \| f \| + 2 \| P_s W_g^{(\mathscr{L})} f \|.
\end{aligned}
$$

Thus we have established the desired bounds.   ∎

A completely similar argument establishes a version for the periodized wavelets:

Lemma 3.7. *Consider an ε-approximate Hilbert pair. Let $W_g^{(\mathcal{L}),\mathrm{per}}$ and $W_h^{(\mathcal{L}),\mathrm{per}}$ denote the periodized wavelet transforms for $\mathcal{L}$ layers, defined as in Eq. (3.13) using the periodizations of the filters g and h, respectively. Then,*

$$\|P_w^{\mathrm{per}}\big(W_g^{(\mathcal{L}),\mathrm{per}} - W_h^{(\mathcal{L})}m(\mu_{w,\mathcal{L}}^{\mathrm{per}})\big)\| \le \varepsilon\mathcal{L}, \tag{3.38}$$

and for all $f \in \mathbb{C}^{2^{\mathcal{L}}}$

$$\|P_s^{\mathrm{per}}\big(W_g^{(\mathcal{L}),\mathrm{per}} - W_h^{(\mathcal{L})}m(\mu_{w,\mathcal{L}}^{\mathrm{per}})\big)f\| \le \varepsilon\mathcal{L}\|f\| + 2\|P_s^{\mathrm{per}}W_g^{(\mathcal{L}),\mathrm{per}}f\|, \tag{3.39}$$

where $\mu_{w,\mathcal{L}}^{\mathrm{per}}[n] := \mu(2^{-\mathcal{L}+1}\pi n)$ and where P_w^{per} denotes the projection onto the $2^{\mathcal{L}} - 1$ many wavelet coefficients and $P_s^{\mathrm{per}} = I - P_w^{\mathrm{per}}$ the projection onto the scaling coefficient.

Next, we will show that expanding a function f in the scaling basis for an approximate Hilbert pair results in approximately the same coefficients as if one were to expand the function in the scaling basis for an exact Hilbert pair (cf. Eq. (3.24)).

Lemma 3.8. *Consider an ε-approximate Hilbert pair. Then there exists a constant $C_\chi > 0$, depending only on the scaling filters, such that the following holds: For every $f \in H^1(\mathbb{R})$,*

$$\|\alpha_j^h f - \alpha_j^g m(\lambda_{s,j})^* f\| \le 2^{-j}C_\chi\|f'\|.$$

where $\lambda_{s,j}(\omega) := \lambda_s(2^{-j}\omega)$. Similarly, for $f \in H^1(\mathbb{S}^1)$ we have that

$$\|\alpha_j^{h,\mathrm{per}} f - \alpha_j^{g,\mathrm{per}} m(\lambda_{s,j}^{\mathrm{per}})^* f\| \le 2^{-j}C_\chi\|f'\|.$$

where $\lambda_{s,j}^{\mathrm{per}}[n] := \lambda_{s,j}(2\pi n)$. If the scaling filters are supported in $\{0,\dots,N-1\}$ then these bounds hold with $C_\chi \le 2N^2$.

Proof. By Eqs. (3.2) and (3.23), $\hat{h}_s - \mu_s\hat{g}_s$ vanishes at $\theta = 0$, so there exists a constant $C > 0$ such that

$$\frac{1}{\sqrt{2}}|\hat{h}_s(\theta) - \mu_s(\theta)\hat{g}_s(\theta)| \le C|\theta| \tag{3.40}$$

for all $\theta \in [-\pi,\pi]$. As a consequence, we can derive the following bound on the Fourier transform of $\chi := \phi^h - m(\lambda_s)\phi^g$: For all $\omega \in [-\pi,\pi]$,

$$|\hat{\chi}(\omega)| = |\hat{\phi}^h(\omega) - \lambda_s(\omega)\hat{\phi}^g(\omega)| = |\prod_{k=1}^{\infty}\frac{1}{\sqrt{2}}\hat{h}_s(2^{-k}\omega) - \prod_{k=1}^{\infty}\frac{1}{\sqrt{2}}\mu_s(2^{-k}\omega)\hat{g}_s(2^{-k}\omega)|$$
$$\le \sum_{k=1}^{\infty}\frac{1}{\sqrt{2}}|\hat{h}_s(2^{-k}\omega) - \mu_s(2^{-k}\omega)\hat{g}_s(2^{-k}\omega)| \le \sum_{k=1}^{\infty}C|2^{-k}\omega| \le C|\omega| \tag{3.41}$$

using a telescoping series and the fact that $\|\hat{h}_s\|_\infty = \|\mu_s\hat{g}_s\|_\infty = \sqrt{2}$. Moreover, $\|\hat{\chi}\|_\infty \le 2$. Thus, Lemma 3.10 shows that, for all $f \in H^1(\mathbb{R})$,

$$\|\alpha_j^h f - \alpha_j^g m(\lambda_{s,j})^* f\|^2 = \sum_{k\in\mathbb{Z}}|\langle\phi_{j,k}^h, f\rangle - \langle\phi_{j,k}^g, m(\lambda_{s,j})^* f\rangle|^2$$
$$= \sum_{k\in\mathbb{Z}}|\langle\chi_{j,k}, f\rangle|^2$$
$$\le 2^{-2j}C_\chi^2\|f'\|^2,$$

where $C_\chi^2 = C^2 + 4/3$. The case when $f \in H^1(\mathbb{S}^1)$ works analogously.

Finally, assume that the scaling filters are supported in $\{0, \dots, N-1\}$. In this case, we know from Lemma 3.9 that, for all $\theta \in [-\pi, \pi]$,

$$\frac{1}{\sqrt{2}}|\hat{g}_s(\theta) - 1| \le \frac{N^2}{2}|\theta| \quad \text{and} \quad \frac{1}{\sqrt{2}}|\hat{h}_s(\theta) - 1| \le \frac{N^2}{2}|\theta|$$

and hence

$$\frac{1}{\sqrt{2}}|\hat{h}_s(\theta) - \mu_s(\theta)\hat{g}_s(\theta)| \le \frac{1}{\sqrt{2}}|\hat{h}_s(\theta) - 1| + \frac{1}{\sqrt{2}}|\hat{g}_s(\theta) - 1| + \frac{1}{\sqrt{2}}|1 - \mu_s(-\theta)|$$

$$\le \left(N^2 + \frac{1}{2}\right)|\theta|$$

for $\theta \in [-\pi, \pi]$. Thus Eqs. (3.40) and (3.41) hold with $C = N^2 + 1/2$, hence we have $C_\chi \le 2N^2$. ∎

The bounds in Lemma 3.8 hold for any pair of wavelets, not only for approximate Hilbert pairs. For the latter, not only is the constant C small in practice, but one can also use the relation between the filters Eq. (3.26) and a slightly adapted version of Lemma 3.10 to show that in fact Lemma 3.8 holds with

$$C_\chi' = 3(C + \varepsilon).$$

For the Selesnick approximate Hilbert pairs this leads to significantly smaller constants, see Table 3.1, but since this does not substantially impact our the scaling of our final bounds on correlation functions we do not pursue this direction further.

3.6.2 Proofs of wavelet lemmas

In this section we will prove some technical lemmas involving wavelets, amongst which Lemma 3.3, Lemma 3.4, Lemma 3.5. We first state a simple Lipschitz bound for the Fourier transforms of wavelet and scaling filters.

Lemma 3.9. *Let g_s be scaling filter supported in $\{0, \dots, N-1\}$. Then the corresponding wavelet filter g_w, defined in Eq. (3.3), is supported in $\{2-N, \dots, 1\}$ and we have that*

$$|\hat{g}_s(\theta) - \sqrt{2}| \le \frac{N^2}{\sqrt{2}}|\theta|,$$

$$|\hat{g}_w(\theta)| \le \frac{N(N+1)}{\sqrt{2}}|\theta|.$$

for all $\theta \in [-\pi, \pi]$.

Proof. By Eq. (3.2), $\|\hat{g}_s\|_\infty = \sqrt{2}$ and $\hat{g}_s(0) = \sqrt{2}$. Hence,

$$\|\hat{g}_s'\|_\infty \le \left(\sum_{n=0}^{N-1} n\right)\|g_s\|_\infty = \frac{N(N-1)}{2}\|g_s\|_\infty \le \frac{N(N-1)}{2}\|\hat{g}_s\|_\infty \le \frac{N^2}{\sqrt{2}},$$

where we used that $\|f\|_\infty \le \frac{1}{2\pi}\|\hat{f}\|_1 \le \|\hat{f}\|_\infty$ for any trigonometric polynomial. Therefore,

$$|\hat{g}_s(\theta) - \sqrt{2}| \le |\hat{g}_s(\theta) - \hat{g}_s(0)| \le \|\hat{g}_s'\|_\infty|\theta| \le \frac{N^2}{\sqrt{2}}|\theta|.$$

Now consider the corresponding wavelet filter g_w which by Eqs. (3.2) and (3.3) satisfies $\|\hat{g}_w\|_\infty = \sqrt{2}$ and $\hat{g}_w(0) = 0$ and is supported in $\{2 - N, \ldots, 1\}$. Then, similarly as above,

$$\|\hat{g}'_w\|_\infty \le \left(\sum_{n=2-N}^{1} |n| \right) \|g_w\|_\infty \le \frac{N(N+1)}{2} \|\hat{g}_w\|_\infty \le \frac{N(N+1)}{\sqrt{2}},$$

so we obtain

$$|\hat{g}_w(\theta)| = |\hat{g}_w(\theta) - \hat{g}_w(0)| \le \|\hat{g}'_w\|_\infty |\theta| \le \frac{N(N+1)}{\sqrt{2}} |\theta|.$$

∎

In practice, the bounds in Lemma 3.9 can be pessimistic. In principle, as the number of vanishing moments of the wavelets increases, one expects better dependence of the bounds on the size of the support. However, we are not aware of better bounds than those in Lemma 3.9 for approximate Hilbert pair wavelets.

We now proceed to prove the lemmas in Section 3.6. Our main tool is the following technical lemma.

Lemma 3.10. *Let $\chi \in H^{-K}(\mathbb{R})$ such that $\hat{\chi} \in L^\infty(\mathbb{R})$ and there exists a constant $C > 0$ such that $|\hat{\chi}(\omega)| \le C|\omega|^K$ for all $|\omega| \le \pi$. Define $C_\chi := (C^2 + \|\hat{\chi}\|_\infty^2 / 3)^{1/2}$. Then, for all $f \in H^K(\mathbb{R})$ and $j \in \mathbb{Z}$ we have that*

$$\sum_{k \in \mathbb{Z}} |\langle \chi_{j,k}, f \rangle|^2 \le 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2,$$

where $\chi_{j,k}(x) := 2^{\frac{j}{2}} \chi(2^j x - k)$. Similarly, for all $f \in H^K(\mathbb{S}^1)$ and $j \ge 0$ we have that

$$\sum_{k=1}^{2^j} |\langle \chi_{j,k}^{\mathrm{per}}, f \rangle|^2 \le 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2,$$

where $\chi_{j,k}^{\mathrm{per}}(x) = \sum_{m \in \mathbb{Z}} \chi_{j,k}(x + m)$.

Proof. For $f \in H^K(\mathbb{R})$, we start with

$$\sum_{k \in \mathbb{Z}} |\langle \chi_{j,k}, f \rangle|^2 = \sum_{k \in \mathbb{Z}} \left| \frac{1}{2\pi} \langle \widehat{\chi_{j,k}}, \hat{f} \rangle \right|^2$$

$$= \sum_{k \in \mathbb{Z}} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} 2^{-j/2} e^{i\omega 2^{-j} k} \overline{\hat{\chi}(2^{-j}\omega)} \hat{f}(\omega) d\omega \right|^2$$

$$= \sum_{k \in \mathbb{Z}} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} 2^{j/2} \overline{\hat{\chi}(\omega)} \hat{f}(2^j \omega) e^{i\omega k} d\omega \right|^2. \tag{3.42}$$

We can interpret this as the squared norm of the Fourier coefficients of the 2π-periodic function defined by

$$F(\theta) := \sum_{m \in \mathbb{Z}} 2^{j/2} \overline{\hat{\chi}(\theta + 2\pi m)} \hat{f}(2^j(\theta + 2\pi m)),$$

provided the latter is square integrable. To see this and obtain a quantitative upper bound, we note that, for every $\theta \in [-\pi, \pi]$,

$$
\begin{aligned}
|F(\theta)|^2 &\leq 2^j \sum_{m \in \mathbb{Z}} \left| \frac{\hat{\chi}(\theta + 2\pi m)}{(\theta + 2\pi m)^K} \right|^2 \sum_{m \in \mathbb{Z}} \left| (\theta + 2\pi m)^K \hat{f}(2^j(\theta + 2\pi m)) \right|^2 \\
&= 2^{-(2K-1)j} \sum_{m \in \mathbb{Z}} \left| \frac{\hat{\chi}(\theta + 2\pi m)}{(\theta + 2\pi m)^K} \right|^2 \sum_{m \in \mathbb{Z}} \left| (2^j(\theta + 2\pi m))^K \hat{f}(2^j(\theta + 2\pi m)) \right|^2 \quad (3.43)
\end{aligned}
$$

by the Cauchy-Schwarz inequality. To bound the left-hand side series, we split off the term for $m = 0$ and use the assumptions on $\hat{\chi}$ to bound, for $|\theta| \leq \pi$,

$$
\begin{aligned}
\sum_{m \in \mathbb{Z}} \left| \frac{\hat{\chi}(\theta + 2\pi m)}{(\theta + 2\pi m)^K} \right|^2 &= \left| \frac{\hat{\chi}(\theta)}{\theta^K} \right|^2 + \sum_{m \neq 0} \left| \frac{\hat{\chi}(\theta + 2\pi m)}{(\theta + 2\pi m)^K} \right|^2 \leq C^2 + \sum_{m \neq 0} \frac{|\hat{\chi}(\theta + 2\pi m)|^2}{|\theta + 2\pi m|^{2K}} \\
&\leq C^2 + \|\hat{\chi}\|_\infty^2 \sum_{m=1}^\infty \frac{2}{(\pi m)^{2K}} \leq C^2 + \frac{\|\hat{\chi}\|_\infty^2}{3} = C_\chi^2 \quad (3.44)
\end{aligned}
$$

If we plug this into Eq. (3.43) then we obtain

$$
|F(\theta)|^2 \leq 2^{-(2K-1)j} C_\chi^2 \sum_{m \in \mathbb{Z}} \left| (2^j(\theta + 2\pi m))^K \hat{f}(2^j(\theta + 2\pi m)) \right|^2
$$

and hence

$$
\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^\pi |F(\theta)|^2 d\theta &\leq 2^{-(2K-1)j} \frac{C_\chi^2}{2\pi} \int_{-\infty}^\infty \left| (2^j \omega)^K \hat{f}(2^j \omega) \right|^2 d\omega \\
&= 2^{-2Kj} \frac{C_\chi^2}{2\pi} \int_{-\infty}^\infty \left| \omega^K \hat{f}(\omega) \right|^2 d\omega = 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2,
\end{aligned}
$$

which is finite since $f \in H^K(\mathbb{R})$. This shows that $F \in L^2(\mathbb{R}/2\pi\mathbb{Z})$. By Parseval's theorem we can thus bound Eq. (3.42) by

$$
\sum_{k \in \mathbb{Z}} \left| \langle \chi_{j,k}, f \rangle \right|^2 \leq 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2 \leq 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2
$$

as desired.

The proof for $f \in H^K(\mathbb{S}^1)$ proceeds similarly. First note that $\widehat{g^{\text{per}}}(m) = \hat{g}(2\pi m)$ if we periodize a function $g \in L^2(\mathbb{R})$ by $g^{\text{per}}(x) := \sum_{n \in \mathbb{Z}} g(x + n)$, so

$$
\sum_{k=1}^{2^j} \left| \langle \chi_{j,k}^{\text{per}}, f \rangle \right|^2 = \sum_{k=1}^{2^j} \left| \langle \widehat{\chi_{j,k}^{\text{per}}}, \hat{f} \rangle \right|^2 = \sum_{k=1}^{2^j} \left| \sum_{m \in \mathbb{Z}} 2^{-j/2} e^{i2\pi m 2^{-j} k} \overline{\hat{\chi}(2^{-j} 2\pi m)} \hat{f}(m) \right|^2 \quad (3.45)
$$

which we recognize as squared norm of the inverse discrete Fourier transform of a vector v with 2^j components

$$
v_l := 2^{j/2} \sum_{m \in \mathbb{Z}} \overline{\hat{\chi}(2\pi m + 2\pi 2^{-j} l)} \hat{f}(2^j m + l),
$$

where it is useful to take $l \in \{-2^{j-1}+1, \ldots, 2^{j-1}\}$. To see that the components of this vector are well-defined and obtain a quantitative bound, we estimate

$$
|v_l|^2 = 2^j \left| \sum_{m \in \mathbb{Z}} \overline{\hat{\chi}(2\pi m + 2\pi 2^{-j} l)} \hat{f}(2^j m + l) \right|^2
$$

$$
\leq 2^j \sum_{m \in \mathbb{Z}} \left| \frac{\hat{\chi}(2\pi m + 2\pi 2^{-j} l)}{(2\pi m + 2\pi 2^{-j} l)^K} \right|^2 \sum_{m \in \mathbb{Z}} \left| (2\pi m + 2\pi 2^{-j} l)^K \hat{f}(2^j m + l) \right|^2
$$

$$
= 2^{-(2K-1)j} \sum_{m \in \mathbb{Z}} \left| \frac{\hat{\chi}(2\pi m + 2\pi 2^{-j} l)}{(2\pi m + 2\pi 2^{-j} l)^K} \right|^2 \sum_{m \in \mathbb{Z}} \left| (2\pi (2^j m + l))^K \hat{f}(2^j m + l) \right|^2.
$$

Since $|2\pi 2^{-j} l| \leq \pi$, we can upper-bound the left-hand side series as in Eq. (3.44),

$$
|v_l|^2 \leq 2^{-(2K-1)j} C_\chi^2 \sum_{m \in \mathbb{Z}} \left| (2\pi (2^j m + l))^K \hat{f}(2^j m + l) \right|^2,
$$

and obtain

$$
\|v\|_2^2 \leq 2^{-(2K-1)j} C_\chi^2 \sum_{n \in \mathbb{Z}} \left| (2\pi n)^K \hat{f}(n) \right|^2 = 2^{-(2K-1)j} C_\chi^2 \|f^{(K)}\|^2,
$$

which is finite since $f \in H^K(\mathbb{S}^1)$. As before we conclude by using the Plancherel formula in Eq. (3.45) and plugging in the upper bound.

$$
\sum_{k=1}^{2^j} \left| \langle \chi_{j,k}^{\mathrm{per}}, f \rangle \right|^2 = 2^{-j} \sum_{k=1}^{2^j} |v_k|^2 \leq 2^{-2Kj} C_\chi^2 \|f^{(K)}\|^2,
$$

which concludes the proof. ∎

We next use Lemma 3.10 to prove Lemma 3.3 and Lemma 3.4, which are wavelet approximation results for sufficiently smooth functions.

Proof of Lemma 3.3. For $f \in H^K(\mathbb{R})$ and $j \in \mathbb{Z}$, we have

$$
\|P_j f - f\|^2 = \sum_{l > j} \sum_{k \in \mathbb{Z}} |\langle \psi_{l,k}, f \rangle|^2.
$$

because the wavelets form an orthonormal basis. We would like to bound the inner series by using Lemma 3.10. For this, note that since $\hat{g}_s$ is a trigonometric polynomial with a zero of order K at $\theta = \pi$, there exists a constant C such that

$$
\frac{1}{\sqrt{2}} |\hat{g}_w(\theta)| = \frac{1}{\sqrt{2}} |\hat{g}_s(\theta + \pi)| \leq C|\theta|^K. \tag{3.46}
$$

Using Eq. (3.6) and $\|\hat{\phi}\|_\infty = 1$, it follows that

$$
|\hat{\psi}(\omega)| = |\frac{1}{\sqrt{2}} \hat{g}_w(\frac{\omega}{2}) \hat{\phi}(\frac{\omega}{2})| \leq \frac{C}{2^K} |\omega|^K.
$$

Since moreover $\|\hat{\psi}\|_\infty = 1$, we can invoke Lemma 3.10 with $\chi = \psi$ and obtain that

$$
\|P_j f - f\|^2 \leq \sum_{l > j} 2^{-2Kl} C_{\mathrm{UV}}^2 \|f^{(K)}\|^2 \leq 2^{-2Kj} C_{\mathrm{UV}}^2 \|f^{(K)}\|^2,
$$

where $C_{\mathrm{UV}}^2 = C^2/4^K + 1/3 \le C^2 + 1/3$.

In the same way we find that, for any $f \in H^K(\mathbb{S}^1)$ and $j \ge 0$,

$$\|P_j^{\mathrm{per}} f - f\|^2 = \sum_{l>j} \sum_{k=1}^{2^l} |\langle \psi_{l,k}^{\mathrm{per}}, f\rangle|^2 \le 2^{-2Kj} C_{\mathrm{UV}}^2 \|f^{(K)}\|^2,$$

again by Lemma 3.10.

For the last assertion, we use Lemma 3.9 to see that, for $K = 1$, Eq. (3.46) always holds with $C = N(N+1)/2$, hence we have $C_{\mathrm{UV}} \le 2N^2$. ∎

Proof of Lemma 3.4. The trigonometric polynomial $\hat{g}_s$ satisfies $\hat{g}_s(0) = \sqrt{2}$, so there is a constant $C > 0$ such that

$$|\frac{1}{\sqrt{2}} \hat{g}_s(\theta) - 1| \le C|\theta| \tag{3.47}$$

for $\theta \in [-\pi, \pi]$. Using the infinite product formula (3.8), it follows that, for all $|\omega| \le \pi$,

$$|\hat{\phi}(\omega) - 1| = |\prod_{k=1}^{\infty} \frac{1}{\sqrt{2}} \hat{g}_s(2^{-k}\omega) - 1| \le \sum_{k=1}^{\infty} |\frac{1}{\sqrt{2}} \hat{g}_s(2^{-k}\omega) - 1| \le \sum_{k=1}^{\infty} \frac{C}{\sqrt{2}} 2^{-k}|\omega| = \frac{C}{\sqrt{2}}|\omega| \tag{3.48}$$

using a telescoping sum and the fact that $|\hat{g}_s| \le \sqrt{2}$ (in fact, this holds for all $\omega \in \mathbb{R}$, but we will not need this). Now recall from Sobolev embedding theory that $\hat{f} \in L^1(\mathbb{R})$ for any $f \in H^1(\mathbb{R})$. Thus, the continuous representative of f can be computed by the inverse Fourier transform, i.e.,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega x} d\omega$$

for all $x \in \mathbb{R}$. As a consequence,

$$\|\alpha_j f - f_j\|^2 = \sum_{k \in \mathbb{Z}} \left|\langle \phi_{j,k}, f\rangle - 2^{-j/2} f(2^{-j}k)\right|^2 = \sum_{k \in \mathbb{Z}} \left|\langle \chi_{j,k}, f\rangle\right|^2$$

where $\chi := \phi - \delta_0$. Now, $\hat{\chi} = \hat{\phi} - \mathbf{1}$, hence $\|\hat{\chi}\|_\infty \le 2$. Together with the bound in Eq. (3.48) we obtain from Lemma 3.10 that

$$\|\alpha_j f - f_j\|^2 \le 2^{-2j} C_\phi \|f'\|^2,$$

where $C_\phi := C^2 + \frac{4}{3}$. The proof for $H^1(\mathbb{S}^1)$ proceeds completely analogously. Finally, Lemma 3.9 shows that if the scaling filter is supported in $\{0, \dots, N-1\}$ then Eqs. (3.47) and (3.48) always hold with $C = N^2/2$. Thus, $C_\phi \le 2N^2$. ∎

Finally, we prove Lemma 3.5, which is an approximation result for compactly supported functions, and which has been adapted from [Woj97].

Proof of Lemma 3.5. Let us denote by S the support of f. Since the scaling functions for fixed j form an orthonormal basis of V_j, and using Cauchy-Schwarz, we find that

$$\|P_j f\|^2 = \sum_{k \in \mathbb{Z}} |\langle \phi_{j,k}, f\rangle|^2 \le \|f\|^2 \sum_{k \in \mathbb{Z}} \int_S |\phi_{j,k}(x)|^2 dx = \|f\|^2 \int_{2^j S} \sum_{k \in \mathbb{Z}} |\phi(y-k)|^2 dy.$$

This allows us to conclude that

$$\|P_j f\|^2 \le \|f\|^2 2^j C_{\mathrm{IR}}^2 D(f),$$

which confirms the claim. If ϕ is bounded and supported on an interval of width N, we can bound $\sum_{k \in \mathbb{Z}} |\phi(y-k)|^2 \le N \|\phi\|_\infty^2$. ∎

Open questions

An interesting open problem is to determine the existence of a family of approximate Hilbert pair wavelets where the approximation error decreases sufficiently fast with the support of the wavelets, as discussed above [Sel02, ACGR20]. Similarly, it would be desirable to identify conditions under which the procedure outlined in Section 3.4 is rigorously guaranteed to find good ε-approximate E-dispersion pairs of biorthogonal filters.

Fermionic entanglement renormalization

We will first give a brief review of the fermionic formalism in Section 4.1. Then we describe the precise relation between Hilbert pair wavelets and entanglement renormalization, both on the level of lattice fermions and fermionic field theory. In Section 4.2 we describe the construction of fermionic entanglement circuits and how one can take a 'continuum limit' to compute correlation functions for a quantum field theory. In Section 4.3 we prove various results about the approximation accuracy of wavelet based entanglement renormalization, which are the main technical contribution of this chapter. The results in this chapter are an extension of results in [HSW⁺18] which discussed the case of lattice fermions. Our main contribution is a careful analysis of the relation to the corresponding quantum field theory.

4.1 Fermions and second quantization

In this section, we briefly review the second quantization formalism for fermions and quasi-free (or Gaussian) fermionic many-body states (see, e.g., [BR03] or [CR87] for further details), and we describe the vacuum state of one-dimensional massless free fermions on a lattice and in the continuum in terms of this formalism.

4.1.1 The CAR algebra and quasi-free states

If $\mathcal{H}$ is a complex Hilbert space (which is the *single particle space*), then let $\mathscr{A}_\wedge(\mathcal{H})$ be the algebra of canonical anti-commutation relations or *CAR algebra* on $\mathcal{H}$. It is the free unital C^*-algebra generated by elements $a(f)$ for $f \in \mathcal{H}$ such that $f \mapsto a(f)$ is anti-linear and subject to the relations

$$\{a(f), a(g)\} = 0,$$
$$\{a(f), a^*(g)\} = \langle f, g \rangle$$

where $\{x, y\} = xy + yx$ denotes the anti-commutator.

An important class of states on this algebra are the *gauge-invariant quasi-free* (or *Gaussian*) states. These states have the property that they are invariant under a global phase $f \mapsto e^{i\phi} f$ and that all correlation functions are determined by the two-point

functions. More precisely, for every operator Q on $\mathcal{H}$ such that $0 \le Q \le I$ there exists a unique gauge-invariant quasi-free state on $\mathscr{A}_\wedge(\mathcal{H})$, denoted ω_Q, such that we have the following version of Wick's rule:

$$\omega_Q(a^*(f_1)\dots a^*(f_n)a(g_1)\dots a(g_m)) = \delta_{n,m}\det[\langle g_i, Qf_j\rangle]$$

Thus, the state is fully specified by its two-point functions $\omega_Q(a^*(f)a(g)) = \langle g, Qf\rangle$. The operator Q is called the *symbol* of ω_Q. The state ω_Q is a pure state if and only if Q is a projection. In this case, Q can be interpreted as a projection onto a *Fermi sea* of negative energy modes. We will only be interested in this case, and henceforth we assume that Q is a projection.

To obtain a Hilbert space realization, we consider the fermionic Fock space

$$\mathscr{F}_\wedge(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\wedge n}$$

with the standard representation of $\mathscr{A}_\wedge(\mathcal{H})$, defined by $a(f) \mapsto a_0(f)$ where

$$a_0^*(f)v = f \wedge v.$$

Let $|\Omega\rangle$ denote the Fock vacuum vector $1 \in \mathcal{H}^{\wedge 0} = \mathbb{C}$. Then $|\Omega\rangle$ is the pure state corresponding to symbol $Q = 0$. Now let Q be an arbitrary orthogonal projection and choose a complex conjugation $\overline{(\cdot)}$ (that is, an anti-unitary involution) that commutes with Q. Then the map $a(f) \mapsto a_Q(f)$, where

$$a_Q(f) = a_0\big((I - Q)f\big) + a_0^*\big(\overline{Qf}\big), \tag{4.1}$$

defines a representation of the CAR algebra such that ω_Q corresponds to the Fock vacuum vector $|\Omega\rangle$.

4.1.2 Second-quantized operators

Next we recall the second quantization of operators on $\mathcal{H}$. If U is a unitary on $\mathcal{H}$ then U defines an automorphism of $\mathscr{A}_\wedge(\mathcal{H})$, known as a *Bogoliubov transformation*, through $a(f) \mapsto a(Uf)$. Provided that $[U, Q]$ is Hilbert-Schmidt, this automorphism can be implemented by a unitary operator $\Gamma_Q(U)$ on Fock space, which is unique up to an overall phase. It is defined by the property that for every $f \in \mathcal{H}$,

$$\Gamma_Q(U)a_Q(f)\Gamma_Q(U)^* = a_Q(Uf).$$

Now consider a unitary one-parameter subgroup $\{e^{itA}\}$ generated by a bounded Hermitian operator A on $\mathcal{H}$. We would like to know when e^{itA} can be unitarily implemented in the form

$$e^{it\,d\Gamma_Q(A)}a_Q(f)e^{-it\,d\Gamma_Q(A)} = a_Q(e^{itA}f) \tag{4.2}$$

for $t \in \mathbb{R}$ and $f \in \mathcal{H}$. For this, decompose A into blocks with respect to $\mathcal{H}_\pm$, which we define as the range of the projections $Q_+ = I - Q$ and $Q_- = Q$ (corresponding to positive and negative energy modes), respectively:

$$A = \begin{pmatrix} A_{++} & A_{+-} \\ A_{-+} & A_{--} \end{pmatrix}$$

In [CR87, Lun76] it is shown that, if A is bounded and the off-diagonal parts A_{+-}, A_{-+} are Hilbert-Schmidt, then there exists a self-adjoint generator $\mathrm{d}\Gamma_Q(A)$ on $\mathscr{F}_\wedge(\mathscr{H})$ such that (4.2) holds. We can moreover fix the undetermined additive constant by demanding that

$$\langle \Omega, \mathrm{d}\Gamma_Q(A)\Omega \rangle = 0,$$

which corresponds to *normal ordering* with respect to the state ω_Q.

If A is trace class then $\mathrm{d}\Gamma_Q(A)$ is bounded and in fact can be defined as an element of $\mathscr{A}_\wedge(\mathscr{H})$. In general, $\mathrm{d}\Gamma_Q(A)$ is unbounded, but we still have the bound [CR87, (2.53)]

$$\|\mathrm{d}\Gamma_Q(A)\Pi_n\| \le 4(n+2)\max\{\|A_{++}\|, \|A_{--}\|, \|A_{+-}\|_2, \|A_{-+}\|_2\}, \tag{4.3}$$

where Π_n denotes the orthogonal projection on the subspace of $\mathscr{F}_\wedge(\mathscr{H})$ spanned by states of no more than n particles $\bigoplus_{k=0}^n \mathscr{H}^{\wedge k}$. Combining [CR87, (2.14), (2.24), (2.25), (2.49)], one can similarly show that

$$\|(\mathrm{d}\Gamma_Q(A) - \mathrm{d}\Gamma_{Q'}(A))\Pi_n\| \le 4(n+2)\max_{\delta=\pm}\{\|Q_\delta A Q_\delta - Q'_\delta A Q'_\delta\|, \|Q_\delta A Q_{-\delta} - Q'_\delta A Q'_{-\delta}\|_2\} \tag{4.4}$$

for any two projections Q and Q'. This estimate will be useful in our error analysis in Section 4.3.2. Finally, if $h : \mathscr{H} \to \mathscr{H}$ is a bounded Hermitian operator, then we may interpret $H = \mathrm{d}\Gamma(h)$ as a Hamiltonian on the Fock space $\mathscr{F}_\wedge(\mathscr{H})$, and we can describe its ground states. By the spectral theorem we may write h as a direct sum of h_- and h_+ where $\langle f, h_- f \rangle \le 0$ and $\langle f, h_- f \rangle \ge 0$ for all $f \in \mathscr{H}$. If h has trivial kernel, the unique ground state is the Gaussian state with symbol the orthogonal projection Q onto the image of h_- (i.e. in the finite dimensional case Q is the projection onto all eigenspaces with negative eigenvalue); informally, the ground state has all negative energy modes filled. If h has nontrivial kernel, this gives rise to a ground state degeneracy (if P is any orthogonal projector onto a subspace of the kernel of h and Q is the orthogonal projection onto the image of h_-, then the state with symbol $Q + P$ is a ground state).

4.1.3   Massless free fermions

We will discuss two relevant models. The first is that of fermions on a lattice, and the second one is its continuum limit, the free Dirac fermionic field theory.

We consider a one-dimensional lattice $\mathbb{Z}$, and we let $\mathscr{H} = \ell^2(\mathbb{Z})$. As usual, we abbreviate $a_n = a[\delta_n]$ for the annihilation operator of a particle on site n. The Hamiltonian

$$H = -\sum_{n\in\mathbb{Z}} a_n^* a_{n+1} + a_{n+1}^* a_n. \tag{4.5}$$

can be thought of as describing particles 'hopping' between neighbouring sites. These particles are massless and the Hamiltonian is gapless. The Hamiltonian H can be written as $\mathrm{d}\Gamma(h)$ where $h : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ given by

$$hf[n] = -(f[n+1] + f[n-1]) \tag{4.6}$$

and correspondingly, as discussed in Section 4.1.2, its ground state is given by the quasi-free state which has as symbol the projection onto the space spanned by all negative eigenvectors of Eq. (4.6). In physics terminology, this is the state where we have filled the

'Dirac sea' consisting of negative energy solutions to the single particle Hamiltonian. We can split the lattice into even and odd sites $\mathbb{Z} = \mathbb{Z}_1 \oplus \mathbb{Z}_2$ and decompose $\mathcal{H} = \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$ accordingly. We may correspondingly redefine the fermionic operators $a_{1,n} = a_{2n+1}$ and $a_{2,n} = a_{2n}$, and finally apply a phase, writing $b_{i,n} = (-1)^n a_{i,n}$. Then the Hamiltonian in Eq. (4.5) is transformed to

$$H = -\sum_{n\in\mathbb{Z}} b_{2,n}^* b_{1,n} - b_{1,n}^* b_{2,n+1} + b_{1,n}^* b_{2,n} - b_{2,n+1}^* b_{1,n}. \tag{4.7}$$

This Hamiltonian can be written in Fourier space as

$$H = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \begin{pmatrix} b_1(\theta) \\ b_2(\theta) \end{pmatrix}^* \begin{pmatrix} 0 & e^{i\theta} - 1 \\ e^{-i\theta} - 1 & 0 \end{pmatrix} \begin{pmatrix} b_1(\theta) \\ b_2(\theta) \end{pmatrix}$$

To see what the symbol corresponding to the ground state is, we observe that for each θ the matrix

$$\begin{pmatrix} 0 & e^{i\theta} - 1 \\ e^{-i\theta} - 1 & 0 \end{pmatrix}$$

has a negative eigenvalue with eigenvector

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -\operatorname{sgn}(\theta) i e^{i\frac{\theta}{2}} \\ 1 \end{pmatrix}$$

which means that the space of negative energy modes consists of all functions $f = f_1 \oplus f_2$ in $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$ which are such that their Fourier transforms satisfy the phase relation

$$\hat{f}_1(\theta) = -\operatorname{sgn}(\theta) i e^{i\frac{\theta}{2}} \hat{f}_2(\theta),$$

or in other words, $f_1 = m(\mu_w) f_2$, with μ_w defined in Eq. (3.23). This implies that the symbol Q is the projection

$$Q = \frac{1}{2} \begin{pmatrix} I & m(\mu_w) \\ m(\mu_w)^* & I \end{pmatrix} \tag{4.8}$$

which projects onto the space of functions $f = f_1 \oplus f_2$ for which $f_1 = m(\mu_w) f_2$.

The continuum limit of this model is given by the free massless Dirac fermion. This theory has action

$$S(\Psi) = \frac{1}{2} \int \Psi^* \gamma^0 \gamma^\mu \partial_\mu \Psi \, dx dt$$

for a two-component complex fermionic field Ψ on the line (or on a circle). The fields have correlation function

$$\langle \Psi^*(x)\Psi(y)\rangle = \frac{1}{x - y}.$$

The stress energy tensor is a normal-ordered product of the fields and its derivatives. In complex coordinates $z = x + it$ and $\bar{z} = x - it$, the stress-energy tensor has a holomorphic component T_{zz} for which one may deduce that

$$\langle T_{zz}(x) T_{zz}(y)\rangle = \frac{1/2}{(x-y)^4}$$

and hence the theory has central charge $c = 1$. For details from the conformal field theory point of view, see [FMS12]. To be able to make rigorous statements about this theory we would like to describe this theory in the algebraic framework. In the algebraic approach described in Section 4.1.1 the fields $\Psi^*(x)$ are not bounded operators. In order to have well-behaved operators, one usually 'smears' the fields. That is, for some function f one informally defines

$$\Psi(f) = \int f(x)\Psi(x)\mathrm{d}x.$$

In this framework, the operator $\Psi(f)$ will be formally defined by the framework in Section 4.1.1 as a bounded operator (while $\Psi(x)$ can not be defined as a bounded operator).

We now describe the vacuum state of this quantum field theory in terms of the second quantization formalism. It will be convenient to consider the Dirac equation, derived from the action, in the form

$$i\gamma^\mu \partial_\mu \psi = 0,$$

with the Dirac matrices $\gamma^0 = i\sigma_z = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ and $\gamma^1 = -\sigma_x = \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right)$. In other words

$$\partial_t \psi_1 + i\partial_x \psi_2 = 0$$
$$\partial_t \psi_2 - i\partial_x \psi_2 = 0$$

The equation is easily seen to be solved by

$$\psi_1(x,t) = \chi_+(x+t) + \chi_-(x-t)$$
$$\psi_2(x,t) = i\left(\chi_+(x+t) - \chi_-(x-t)\right)$$

for arbitrary functions χ_+ and χ_-, which we take to be in $L^2(\mathbb{R})$ in order for the solutions to be normalizable. The energy of such a solution is given by

$$E = \int_{-\infty}^\infty \left(-\omega|\hat{\chi}_+(\omega)|^2 + \omega|\hat{\chi}_-(\omega)|^2\right)\mathrm{d}\omega.$$

Thus, the space of negative energy solutions is spanned by solutions for which χ_+ has a Fourier transform with support on the positive half-line (is analytic) and χ_- has a Fourier transform with support on the negative half-line (is anti-analytic).

We obtain a single-particle Hilbert space $\mathscr{H} = L^2(\mathbb{R}) \otimes \mathbb{C}^2$ corresponding to $\psi(x,t=0)$. Again, the symbol of the vacuum state is given by the projection onto the 'Dirac sea' of negative energy solutions. It can be expressed as

$$Q = \frac{1}{2}\begin{pmatrix} I & \mathscr{H} \\ -\mathscr{H} & I \end{pmatrix} \tag{4.9}$$

in terms of the *Hilbert transform*, which we recall is the unitary operator on $L^2(\mathbb{R})$ defined by

$$\widehat{\mathscr{H}f}(\omega) = -i\,\mathrm{sgn}(\omega)\hat{f}(\omega).$$

Indeed, it follows from $\mathscr{H}^* = -\mathscr{H}$ that Q is an orthogonal projection, and $Q\psi = \psi$ if ψ is the restriction to $t = 0$ of a negative-energy solution. We further note that the symbol Q

commutes with the component-wise complex conjugation on $\mathscr{H}$. We thus obtain a Fock space realization as described above in Section 4.1.1. The smeared Dirac field can be defined as $\Psi(f) := a_Q(f)$ for $f \in \mathscr{H}$.

We will also be interested in free Dirac fermions on the circle $\mathbb{S}^1$. In this case, we take $\mathscr{H} = L^2(\mathbb{S}^1) \otimes \mathbb{C}$. For periodic boundary conditions, the symbol Q^{per} has the same form as in (4.9), where we now let

$$\widehat{\mathscr{H}^{\mathrm{per}} f}[n] = -i \operatorname{sgn}(n) \hat{f}[n]$$

where there is some ambiguity in the sign function for $n = 0$ (reflecting a ground state degeneracy). For definiteness, we choose $\operatorname{sgn}(0) = 1$.

For anti-periodic boundary conditions, corresponding to the Dirac equation on the nontrivial spinor bundle over $\mathbb{S}^1$, we define a unitary operator T on $\mathscr{H}$ by

$$T f(x) = e^{-i\pi x} f(x) \tag{4.10}$$

for $x \in (0,1)$. Then the symbol is given by $T^* Q^{\mathrm{per}} T$.

4.1.4 Self-dual CAR algebra and Majorana fermions

Suppose that $\mathscr{H}_+ \cong \mathscr{H}_-$, as in the preceding section. Given an anti-unitary involution C on $\mathscr{H}$ such that $C Q_\delta = Q_{-\delta} C$ for $\delta = \pm$, we can also define the following operators on the Fock space $\mathscr{F}_\wedge(\mathscr{H}_+) \subset \mathscr{F}_\wedge(\mathscr{H})$,

$$c_Q(f) = a_0(Q_+ f) + a_0^*(C Q_- f) \tag{4.11}$$

These satisfy the relations of the *self-dual CAR algebra*, $\mathscr{A}_\wedge^{\mathrm{sd}}(\mathscr{H})$ [Ara71], which is generated by elements $c(f)$ for $f \in \mathscr{H}$ such that $f \mapsto c(f)$ is antilinear and

$$\{c(f), c^*(g)\} = \langle f, g \rangle,$$
$$c^*(f) = c(Cf).$$

for $f, g \in \mathscr{H}$. The second equation implies that a unitary U on $\mathscr{H}$ only defines an automorphism $\Gamma^c(U)$ of $\mathscr{A}_\wedge^{\mathrm{sd}}(\mathscr{H})$ by $c(f) \mapsto c(Uf)$ if $[U, C] = 0$ commutes with C. We can also second quantize generators as in Eq. (4.2). That is, if A is a bounded operator with Hilbert-Schmidt A_{+-}, A_{-+}, and if $A^* = -CAC$, we can define its second quantization $\mathrm{d}\Gamma_Q^c(A)$, such that

$$e^{it\mathrm{d}\Gamma_Q^c(A)} c_Q(f) e^{-it\mathrm{d}\Gamma_Q^c(A)} = c_Q(e^{itA} f). \tag{4.12}$$

We can apply this construction in the situation Section 4.1.3 to obtain a description of massless free Majorana fermions. Define the anti-unitary involution C as the following charge conjugation operator which exchanges positive and negative energy modes:

$$Cf = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bar{f} \tag{4.13}$$

Then it is clear from (4.9) that $CQ = (I - Q)C$, so the above construction applies. We denote by $\Phi(f) := c_Q(f)$ the smeared Majorana field.

Gaussian unitaries	$a_n \mapsto \sum_m U_{n,m} a_m$ for U unitary
Hamiltonian	$H = -\sum_n a_n^* a_{n+1} + a_{n+1}^* a_n$
Wavelet filters	g, h orthogonal wavelet filters
Filter relation	$\hat{g}_w(\theta) = -i \operatorname{sgn}(\theta) e^{i\frac{\theta}{2}} \hat{h}_w(\theta)$ for $\theta \in (-\pi, \pi)$
Application of wavelet transform	$b_{1,n} = (-1)^n a_{2n}$, $b_{2,n} = (-1)^n a_{2n+1}$, and apply W_h to the b_1 fermions and W_g to the b_2 fermions
Disentangling circuit	Apply wavelet decomposition, then H on wavelet modes.
Continuum theory	Free Dirac fermion
Wavelet functions	$\hat{\psi}^g(\omega) = -i \operatorname{sgn}(\omega)\hat{\psi}^h(\omega)$

Table 4.1: Overview of the construction of MERA from wavelets for fermions as described in [EW16, HSW$^+$18, WSSW22], to be compared with the corresponding results for bosonic systems in Table 5.1.

4.2 Entanglement renormalization circuits

In this section we will describe how an appropriate quantization of Hilbert pair wavelets yields an entanglement renormalization scheme for free fermions. For a discussion of Gaussian fermionic MERA, see [EV10b], and for fermionic MERA in general, see [CV09]. An overview of fermionic wavelet MERA is provided in Table 4.1.

4.2.1 Entanglement renormalization for lattice fermions

We start by reviewing the construction of entanglement renormalization circuits for massless free fermions, as worked out in [EW16, HSW$^+$18]. The exposition is based on [HSW$^+$18].

Recall that the hopping fermion Hamiltonian could be rewritten as Eq. (4.7)

$$H = -\sum_{n \in \mathbb{Z}} b_{1,n}^* b_{2,n} - b_{2,n}^* b_{1,n+1} + b_{2,n}^* b_{1,n} - b_{1,n+1}^* b_{2,n},$$

and that the ground state symbol was given by Eq. (4.8)

$$Q = \frac{1}{2}\begin{pmatrix} I & m(\mu_w) \\ m(\mu_w)^* & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & m(\mu_w)^* \end{pmatrix} (I \otimes |+\rangle \langle +|) \begin{pmatrix} I & 0 \\ 0 & m(\mu_w) \end{pmatrix}$$

as a projection on the single-particle Hilbert space $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$. Recall that $m(\mu_w)$ is a Fourier multiplication operator, where $\theta_w(\theta) = -i \operatorname{sgn}(\theta) e^{i\frac{\theta}{2}}$ for $|\theta| < \pi$. This symbol is straightforward to prepare in Fourier space. For instance, if one would like to prepare the ground state, one could apply (the second quantization of) the Fourier transform, which relates the ground state to a product state. However, we are interested in performing a circuit which is *local in real space* (recall that entanglement renormalization is a quantum circuit which implements a real-space renormalization). For this goal it is natural to replace a Fourier analysis of the ground state symbol by a wavelet analysis.

We now consider an approximate Hilbert pair as in Definition 3.1. As before, we denote by g_w, h_w, g_s, h_s the wavelet and scaling filters, and we let W_g and W_h denote the corresponding wavelet decomposition maps. Since W_g and W_h are unitary maps we can apply the fermionic second quantization of W_h to the b_1 fermions and W_g to the

b_2 fermions and we write

$$W = W_h \oplus W_g.$$

Since we have an approximate Hilbert pair, the filters are such that for $\theta \in (-\pi, \pi)$

$$h_w \approx m(\mu_w) g_w \tag{4.14}$$

This will allow us to renormalize the ground state. Suppose we had an *exact* Hilbert pair (we will see in detail in Section 4.3 how this derivation is influenced by the error in Eq. (4.14)). Then the second quantization of W will be such that if we let $|\Omega\rangle$ denote the ground state of Eq. (4.7) with symbol Q, (so $|\Omega\rangle$ is the Fock vacuum with representation a_Q as in Eq. (4.1)), then we may consider the state

$$\Gamma_Q(W) |\Omega\rangle \langle\Omega| \Gamma_Q(W),$$

which lives on the Fock space with single particle space $(\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2) \oplus (\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2)$, where the first factor in the direct sum collects the wavelet outputs of $W_g \oplus W_h$ and the second factor in the direct sum collects the scaling outputs of $W_g \oplus W_h$. This state will have symbol WQW^*. We note that for an exact Hilbert pair we have

$$(I \oplus m(\mu_w)) W_g = W_h m(\mu_w)$$

as follows from the observations in the proof of Lemma 3.6. Then it is easy to verify, using that W_g is unitary, that

$$\begin{aligned} WQW^* &= \begin{pmatrix} W_h & 0 \\ 0 & W_g m(\mu_w)^* \end{pmatrix} (I \otimes |+\rangle \langle+|) \begin{pmatrix} W_h^* & 0 \\ 0 & m(\theta_w) W_g^* \end{pmatrix} \\ &= (I \otimes |+\rangle \langle+|) \oplus Q. \end{aligned}$$

In other words, after applying a layer of the wavelet pair decomposition, the state is in a product state between the wavelet output and the scaling output, and moreover, the state on the scaling output is again the ground state of the Hamiltonian Eq. (4.7), while the state on the wavelet output is a product state. Thus we have implemented a layer of entanglement renormalization: we have separated the high frequency (wavelet) and low frequency (scaling modes), and moreover disentangled the high frequency modes, and done so in a scale invariant way. We can now repeat this construction on the scaling modes, and the $\mathscr{L}$-layer wavelet decomposition will give rise to an entanglement renormalization procedure of depth $\mathscr{L}$. If we have a finite number of layers, we obtain the following approximation of the symbol (where we insert the Fock vacuum state on the top scaling input):

Definition 4.1 (Approximate symbol for lattice fermion)**.** For any approximate Hilbert pair, $\mathscr{L} \in \mathbb{N}$, define the *approximate symbol* for the lattice fermion in Eq. (4.7) as the following projection on $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$:

$$\tilde{Q}_{\mathscr{L}}^{\text{latt}} := W^{(\mathscr{L}),*} \big(P_w \otimes |+\rangle \langle+| \big) W^{(\mathscr{L})}, \tag{4.15}$$

where $W^{(\mathscr{L})} := W_h^{(\mathscr{L})} \oplus W_g^{(\mathscr{L})}$ and P_w is the projection onto the wavelet output.

Note that $P_w \otimes |+\rangle\langle+|$ is the symbol of a state for which correlation functions can be straightforwardly evaluated. Indeed, we can intuitively think of $P_w \otimes |+\rangle\langle+|$ as the symbol of a 'Fermi sea' where on the part of the lattice corresponding to the wavelet output of the transform we have a have occupied a product mode. Equivalently, after applying the second quantization of a Hadamard gate and a Jordan-Wigner transformation this state corresponds to an 'infinite product state' where the wavelet qubits are in state $|101010\ldots\rangle$ and the scaling qubits in $|0000\ldots\rangle$.

We may thus conclude that the second quantization of an appropriate discrete wavelet transform defines an entanglement renormalization map. The relation in Eq. (4.14) will not be satisfied exactly by a pair of finite filters, but as we saw in Section 3.3 it can be approximated. In [EW16] it was shown that a construction using Daubechies D4 filters already gives a good result, and in [HSW$^+$18] it was shown that any approximate Hilbert pair leads to a good approximation of the ground state using the approximate symbol $Q_{\mathscr{L}}^{\text{latt}}$. This result will be extended in Theorem 4.7 to correlation functions for the corresponding quantum field theory.

Discrete wavelet transform and single-particle circuits

Let us now recall how discrete wavelet transforms can be written as *single-particle* ('first quantized' or 'classical') linear *circuits*, using the construction in Section 3.5, and show that the entanglement renormalization procedure described above indeed gives rise to an entanglement renormalization circuit. Recall that 'single-particle' means that the state space is a direct sum of local state spaces (such as $\ell^2(\mathbb{Z}) = \bigoplus_{n \in \mathbb{Z}} \mathbb{C}$). Thus let $W \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{C}^2)$ denote a single layer of a discrete wavelet transform, defined as in Eq. (3.10). Recall that by putting the scaling and wavelet outputs on the even and odd sublattice we obtain a unitary

$$W' \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}), \quad W' := \iota W,$$

where

$$\iota \colon \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \to \ell^2(\mathbb{Z})$$
$$\iota(f_w \oplus f_s)[2n] = f_w[n],$$
$$\iota(f_w \oplus f_s)[2n+1] = f_s[n]$$

for all $f_w, f_s \in \ell^2(\mathbb{Z})$. From Section 3.5, and as previously shown in [EW18], we know that if the scaling filters are real and have length N then W' can be decomposed into a product $W' = W_{N/2} \cdots W_1$, where each $W_k \colon \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ is a block-diagonal real-valued unitary of the form

$$W_k = \begin{cases} \bigoplus_{r \text{ odd}} u_{r,r+1}(\theta_k) & \text{if } k \text{ odd}, \\ \bigoplus_{r \text{ even}} u_{r,r+1}(\theta_k) & \text{if } k \text{ even}. \end{cases}$$

Here, the θ_k are suitable angles and $u_{r,r+1}(\theta_k)$ denotes the unitary which acts on the subspace $\ell^2(\{r, r+1\}) \subseteq \ell^2(\mathbb{Z})$ by the rotation matrix

$$u(\theta_k) = \begin{pmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{pmatrix}. \tag{4.16}$$

Thus, we obtain a decomposition of W^g into a single-particle linear circuit composed of 2-local unitaries (see Fig. 3.7, (a)). In the same way we can implement $\mathscr{L}$ layers of the discrete wavelet transform.

Given an approximate Hilbert pair (or any pair of wavelets) let $W := W_h \oplus W_g$, corresponding to performing both discrete wavelet transforms in parallel. If we apply the preceding construction to both wavelet transforms W_h and W_g we obtain two circuits, one for W'_h and one for W'_g, parametrized by angles θ_k^h and θ_k^g for $k = 1,\ldots,N/2$. These can be assembled into a single single-particle circuit for

$$W': \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2, \quad W' := W'_h \oplus W'_g$$

As shown in Fig. 4.1, (a), we take each site to carry two degrees of freedom (corresponding to the two components of the Dirac spinor). Instead we could also arrange the two wavelet transforms on the even and odd sublattices (by conjugating with ι). It is straightforward to see that the corresponding circuit can be implemented by 2-local unitaries and swap gates. Thus

$$Q_{\mathscr{L}}^{\text{latt}} = U_{\text{MERA}}^{(\mathscr{L}),*} P U_{\text{MERA}}^{(\mathscr{L})}, \tag{4.17}$$

where $P := P_w \otimes |0\rangle\langle 0|$ is a symbol on $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}+1} \otimes \mathbb{C}^2$ and

$$U_{\text{MERA}}^{(\mathscr{L})}: \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \to \ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}+1} \otimes \mathbb{C}^2$$

is the single-particle unitary defined by

$$U_{\text{MERA}}^{(\mathscr{L})} := (I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^{\mathscr{L}-1} \otimes \mathbb{C}^2} \oplus U_{\text{MERA}}) \cdots (I_{\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2} \oplus U_{\text{MERA}}) U_{\text{MERA}},$$

$$U_{\text{MERA}} := (I_{\ell^2(\mathbb{Z})} \otimes h) \oplus (I_{\ell^2(\mathbb{Z})} \otimes I_{\mathbb{C}^2})) W,$$

where h is the Hadamard matrix $h = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ which maps $h|0\rangle = |+\rangle$. Just like W, the unitary U_{MERA} can be implemented by a single-particle circuit of depth $N/2 + 1$, where N is the length of the filters, obtained by composing the circuit for W with an additional layer of Hadamard unitaries acting on the wavelet outputs (see Fig. 4.1). The unitary $U_{\text{MERA}}^{(\mathscr{L})}$ consists of $\mathscr{L}$ such circuit layers.

Quantum circuits from single-particle circuits

Since we seek to describe a quantum many-body state of fermions, the circuit that we will construct naturally arises as a quantum circuit that acts on a fermionic Fock space $\mathscr{F}_\wedge(\ell^2(\mathbb{Z}))$. We define a *Gaussian fermionic quantum circuit* to be the second quantization $\Gamma(U)$ of a single-particle circuit. Note that this is, a priori, different from the standard notion of a quantum circuit on a Hilbert space which is a tensor product of local qudit Hilbert spaces. However, in the one-dimensional setting, if one takes a so-called Jordan-Wigner transformation (see below, at Eq. (4.21)) of a Gaussian fermionic quantum circuit, one obtains a local quantum circuit on a chain of qubits. The resulting circuits are matchgate circuits [JM08] (which can be defined as the class of circuits which can be written as the Jordan-Wigner transformation of a one-dimensional nearest neighbour Gaussian fermionic circuit).

In our case, we can obtain a quantum circuit by second-quantizing the single-particle circuit $U_{\text{MERA}}^{(\mathscr{L})}$ for the wavelet transforms described above. Thus, $\Gamma(U_{\text{MERA}})$, applied to the Fock vacuum, prepares the state with symbol $Q_{\text{MERA}}^{\text{latt}}$, and hence we have constructed a fermionic entanglement renormalization circuit to approximate the ground state of a critical lattice fermion.

Entanglement renormalization for the Ising chain

In [EW16] it has been worked out how, for the particular case of Daubechies D4 wavelets, one obtains approximate entanglement renormalization circuits for the Ising model. It is well known that the one-dimensional quantum Ising model, with Hamiltonian

$$H = -\sum_n X_n X_{n+1} + Z_n \tag{4.18}$$

can be solved by relating it to free fermions. The Hamiltonian in Eq. (4.18) acts on a chain of qubits, and X_n and Z_n are Pauli operators which are given by an identity at every site except n, where they are given by X and Z respectively. In fact, under an appropriate Jordan-Wigner transformation, the Hamiltonian Eq. (4.7) can be transformed to two independent copies of Eq. (4.18). To this end, recall that we had fermions with annihilation operators $b_{1,n}$ and $b_{2,n}$. If we let

$$c_{1,2n} = b_{1,n}^* + b_{1,n}$$
$$c_{1,2n+1} = i(b_{2,n}^* - b_{2,n})$$
$$c_{2,2n} = i(b_{1,n}^* - b_{1,n})$$
$$c_{2,2n+1} = b_{2,n}^* + b_{2,n}$$

then the $c_{i,n}$ are Majorana fermions and we can rewrite Eq. (4.7) as $H_1 - H_2$ where

$$H_j = -\frac{i}{2} \sum_{n \in \mathbb{Z}} c_{j,n} c_{j,n+1}.$$

Thus, we see that the Hamiltonian decouples into two uncoupled Majorana Hamiltonians. If we take H_1, we may now map this to the Ising Hamiltonian in Eq. (4.18) by the Jordan-Wigner transformation

$$c_{1,2n} \longmapsto \left(\prod_{r<n} Z_r\right) X_n \tag{4.19}$$

$$c_{1,2n+1} \longmapsto \left(\prod_{r<n} Z_r\right) Y_n. \tag{4.20}$$

Because our single-particle circuit consists of orthogonal maps, it decouples into two separate circuits acting on the two decoupled Majorana modes. Denote by θ_k^g and θ_k^h the sequence of angles associated to an approximate Hilbert pair, as in Eq. (4.16), and let $\theta_k^\pm = \theta_k^g \pm \theta_k^h$. Then after application of the Jordan-Wigner transformation we obtain the following two-qubit gate for the k-th layer in a single layer of entanglement renormalization:

$$U_k = \begin{pmatrix} \cos(\theta_k^-) & 0 & 0 & -\sin(\theta_k^-) \\ 0 & \cos(\theta_k^+) & -\sin(\theta_k^+) & 0 \\ 0 & \sin(\theta_k^+) & \cos(\theta_k^+) & 0 \\ \sin(\theta_k^-) & 0 & 0 & \cos(\theta_k^-) \end{pmatrix} \tag{4.21}$$

This provides a generalization of the circuits in [EW16]. The resulting circuits are entanglement renormalization circuits with the structure shown Fig. 2.1 and in this case it is a (matchgate) circuit in the 'usual' sense, acting on qubits rather than fermionic modes.

4.2.2 Entanglement renormalization for the Dirac fermion

From Section 4.2.1 we have learned that the second quantization of a *discrete* wavelet transform for an approximate Hilbert pair corresponds to an entanglement renormalization procedure for a critical lattice fermionic Hamiltonian. It is now natural to wonder whether the second quantization of the wavelet decomposition of $L^2(\mathbb{R}) \otimes \mathbb{C}^2$ as defined in Section 3.1.1 using an (approximate) Hilbert pair of wavelets can be interpreted as entanglement renormalization of the continuum limit of Eq. (4.7) i.e. the massless Dirac fermion. This is indeed the case, as we will explain in this section. The way we approach this problem is by describing a procedure which approximates correlation functions of smeared operators. Informally, the procedure is that we first discretize the operators at some scale (i.e., we impose a UV cut-off), and then, in order to obtain the free fermion vacuum, we need to 'fill the Dirac sea' up to the relevant scale. So, the circuit, starting from the Fock vacuum, has to fill all the negative energy modes over the range of scales that are relevant for the inserted operators, directly analogous to a real-space renormalization procedure. The approximate Hilbert pair is exactly such that it allows one to construct such modes over a range of scales.

To make this more precise let $\{O_i\}$, $i = 1, \ldots, n$ be a set of smeared operators that are either linear in the fields (such as $\Psi(f)$, $\Psi^*(f)$ for a smearing function), or normal-ordered quadratic operators (such as smeared components of the stress-energy tensor), and which are compactly supported. We denote the correlation functions by

$$G(\{O_i\}) = \langle O_1 \cdots O_n \rangle. \tag{4.22}$$

We would like to approximate such correlation functions using an entanglement renormalization circuit.

4.2.3 Quantum circuits for correlation functions

We now explain how we the construction of entanglement renormalization circuits can be used to compute correlation functions for free Dirac and Majorana fermions.

We now explain how the latter can be computed by a fermionic quantum circuit of MERA type. Loosely speaking, we will do the following, in order to compute some correlation function $G(\{O_i\})$:

(i) We discretize the operators using the scaling functions. This gives a set of operators $\{O_i^{\mathrm{MERA}}\}$ on the lattice. One can also consider this procedure the other way around: this embeds a discrete theory in the continuous theory by smearing all operators with appropriate scaling functions.

(ii) We compute the correlation function of the operators $\{O_i^{\mathrm{MERA}}\}$ using the state obtained by applying the second quantization of $U_{\mathrm{MERA}}^{(\mathscr{L})}$ (illustrated in Fig. 4.1) to the Fock vacuum.

Let us first discuss the case of the free Dirac fermion on the real line in more detail.

Recall from Eq. (4.9) that the symbol of the vacuum state of the free Dirac fermion on the real line is given by the following operator on $L^2(\mathbb{R}) \otimes \mathbb{C}^2$, which is the single-particle Hilbert space:

$$Q = \frac{1}{2} \begin{pmatrix} I & \mathscr{H} \\ -\mathscr{H} & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \mathscr{H}^* \end{pmatrix} \left(I_{L^2(\mathbb{R})} \otimes |+\rangle \langle +| \right) \begin{pmatrix} I & 0 \\ 0 & \mathscr{H} \end{pmatrix}, \tag{4.23}$$

where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

To obtain a suitable approximation, consider an approximate Hilbert pair as in Definition 3.1. As before, we denote by g_w, h_w, g_s, h_s the wavelet and scaling filters, by α_j^g and α_j^h discretization maps (defined as in Eq. (3.9)) and by $W_g^{(\mathscr{L})}$ and $W_h^{(\mathscr{L})}$ the $\mathscr{L}$-layer discrete wavelet transforms (defined in Eq. (3.11)). We now approximate Eq. (4.23) by first truncating to a finite number of scales, using one of the two wavelet transforms, and then by replacing the Hilbert transform of the one wavelet basis by the other wavelet basis. Schematically,

$$I_{L^2(\mathbb{R})} \rightsquigarrow \alpha_j^{h,*} W_h^{(\mathscr{L}),*} P_w W_h^{(\mathscr{L})} \alpha_j^h, \qquad P_w W_h^{(\mathscr{L})} \alpha_j^h \mathscr{H} \rightsquigarrow P_w W_g^{(\mathscr{L})} \alpha_j^g,$$

where $P_w = I_{\ell^2(\mathbb{Z})} \otimes \sum_{k=0}^{\mathscr{L}-1} |k\rangle \langle k|$ denotes the orthogonal projection onto the wavelet coefficients.

Definition 4.2 (Approximate symbol for the Dirac fermion). For any approximate Hilbert pair, $j \in \mathbb{Z}$, and $\mathscr{L} \in \mathbb{N}$, define the *approximate symbol* as the following projection on $L^2(\mathbb{R}) \otimes \mathbb{C}^2$:

$$\tilde{Q}_{j,\mathscr{L}} := \alpha_j^* W^{(\mathscr{L}),*} \big(P_w \otimes |+\rangle \langle+| \big) W^{(\mathscr{L})} \alpha_j, \tag{4.24}$$

where $\alpha_j := \alpha_j^h \oplus \alpha_j^g$ and $W^{(\mathscr{L})} := W_h^{(\mathscr{L})} \oplus W_g^{(\mathscr{L})}$.

In other words, $\tilde{Q}_{j,\mathscr{L}} = \alpha_j^* Q_{\mathscr{L}}^{\text{latt}} \alpha_j$. The symbol $\tilde{Q}_{j,\mathscr{L}}$ should be seen as an approximation of the true symbol at scales ranging from 2^{-j+1} to $2^{-j+\mathscr{L}}$. In Section 4.3 we will derive that this is indeed a good approximation to the symbol, and bound the error that derives from taking only a finite number of scales and from the inaccuracy in the Hilbert pair relation.

The key point is that we can now compute an approximation to the correlation function $G(\{O_i\})$ as follows.

Definition 4.3 (MERA correlation functions). Consider an approximate Hilbert pair with filters g, h. Given a correlation function (4.36), $j \in \mathbb{Z}$, and $\mathscr{L} \geq 0$, we define the corresponding *MERA correlation function* by

$$G_{j,\mathscr{L}}^{\text{MERA}}(\{O_i\}) := \langle \Omega | O_1^{\text{MERA}} \cdots O_n^{\text{MERA}} | \Omega \rangle, \tag{4.25}$$

where O_i^{MERA} is obtained from O_i by replacing $\Psi(f)$ by $\Psi_{\text{MERA}}(f) := a_P(U_{\text{MERA}}^{(\mathscr{L})} \alpha_j f)$ and $d\Gamma_Q(A)$ by $d\Gamma_{\text{MERA}}(A) := d\Gamma_P(U_{\text{MERA}}^{(\mathscr{L})} \alpha_j A \alpha_j^* U_{\text{MERA}}^{(\mathscr{L}),*})$. Here, $P := P_w \otimes |0\rangle \langle 0|$.

More precisely, using Eq. (4.1), we find that

$$\Psi_{\text{MERA}}(f) = a_0 \big((I - P) U_{\text{MERA}}^{(\mathscr{L})} \alpha_j f \big) + a_0^* \big(\overline{P U_{\text{MERA}}^{(\mathscr{L})} \alpha_j f} \big), \tag{4.26}$$

where $a_0^{(*)}$ are the ordinary creation and annihilation operators on Fock space. If f is a smearing function then in order to find $\Psi_{\text{MERA}}(f)$ we first compute $\alpha_j f$ either by expanding the scaling basis or simply by sampling on a (dyadic) grid (Lemma 3.4), then we apply $\mathscr{L}$ layers of the local circuit U_{MERA} (Fig. 4.1), and finally we apply the projections P and $I - P$. One can proceed similarly for $d\Gamma_{\text{MERA}}(A)$. This shows that the correlation functions (4.37) can be efficiently calculated in the single-particle picture.
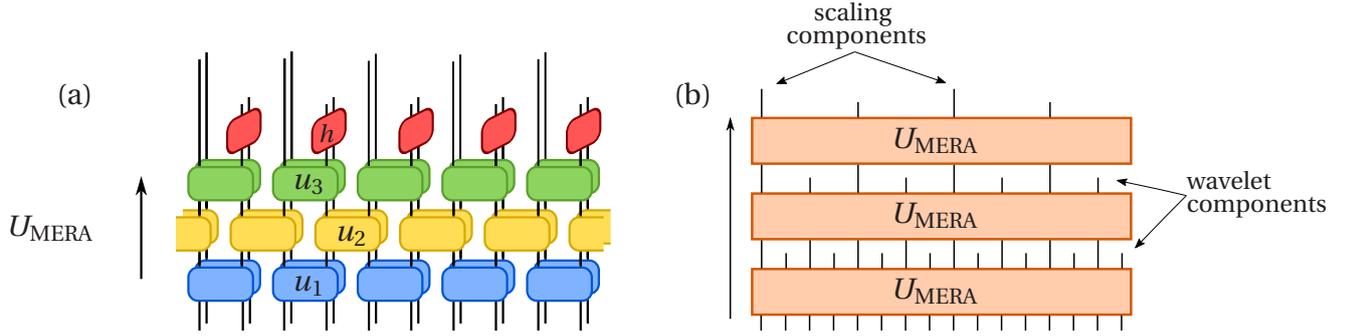
Figure 4.1: (a) A single MERA layer U_{MERA}, acting in the single-particle picture as a wavelet decomposition. The Hadamard unitary h (dis)entangles the modes of the two wavelet transforms that make up the Hilbert pair. We abbreviate $u_k := u_k^h \oplus u_k^g$. (b) Illustration of the unitary $U_{\mathrm{MERA}}^{(\mathscr{L})}$ corresponding to $\mathscr{L}$ MERA layers before second quantization. Each layer is a local circuit of depth $N/2 + 1$, as in (a).

Finally, we need to take into account that in concrete computations we have finite systems. Let us assume that we would like to approximate a correlation function involving $\Psi^{(*)}(f_i)$ and $\mathrm{d}\Gamma_Q(A_i)$, where the smearing functions f_i and the kernel of A_i are compactly supported. In this case, it is easy to see that Eq. (4.25) will involve creation and annihilation operators that act only on finitely many sites $S \subseteq \mathbb{Z}$ (which can be computed from the supports as well as the parameters j, $\mathscr{L}$, and N). In this case, we can replace $\ell^2(\mathbb{Z})$ by $\ell^2(S)$, P by its restriction P_S onto $\mathscr{H}_S := \ell^2(S) \otimes \mathbb{C}^{\mathscr{L}+1} \otimes \mathbb{C}^2$, and the infinitely wide layers U_{MERA} by finitely many local unitaries. Let us denote by $|P_S\rangle$ the corresponding Gaussian state in the fermionic Fock space $\mathscr{F}_\wedge(\mathscr{H}_S)$ and we denote by $\Gamma_0(U_{\mathrm{MERA}}^{(\mathscr{L})}) := \bigoplus_{k=0}^\infty (U_{\mathrm{MERA}}^{(\mathscr{L})})^{\wedge k}$ the second quantizations of the single-particle unitaries $U_{\mathrm{MERA}}^{(\mathscr{L})}$. Since second quantization commutes with convolution, this can be written as a fermionic quantum circuit composed of $\mathscr{L}$ many identical layers, each of depth $N/2 + 1$ (which structurally looks like Fig. 4.1, (b)). Thus, we recognize that $|\mathrm{MERA}_{\mathscr{L}}\rangle := \Gamma_0(U_{\mathrm{MERA}}^{(\mathscr{L})})^* |P_S\rangle$ is precisely the quantum state prepared by a fermionic MERA. Moreover, we can compute the MERA correlation functions by

$$G_{j,\mathscr{L}}^{\mathrm{MERA}}(\{O_i\}) = \langle \mathrm{MERA}_{\mathscr{L}} | O_1' \cdots O_n' | \mathrm{MERA}_{\mathscr{L}} \rangle, \qquad (4.27)$$

where O_i' is obtained from O_i by replacing

$$\Psi(f) \mapsto \Psi'(f) := a_0(\alpha_j f)$$
$$\mathrm{d}\Gamma_Q(A) \mapsto \mathrm{d}\Gamma_0(\alpha_j A \alpha_j^*) - \langle \mathrm{MERA}_{\mathscr{L}} | \mathrm{d}\Gamma_0(\alpha_j A \alpha_j^*) | \mathrm{MERA}_{\mathscr{L}} \rangle.$$

Note that $\langle \mathrm{MERA}_{\mathscr{L}} | \mathrm{d}\Gamma_0(\alpha_j A \alpha_j^*) | \mathrm{MERA}_{\mathscr{L}} \rangle$ is actually *finite* because we truncated the range of wavelet scales, so this normal ordering is well-defined (even if the original operator A was not trace class). Thus, Eq. (4.27) can be interpreted as an ordinary correlation function in a fermionic MERA. This at last justifies our notation.

4.2.4   Circle, boundary conditions, Majorana fermions

On the circle $\mathbb{S}^1$ we proceed similarly, except that there is now a natural largest scale. We will only discuss the continuous case. For periodic boundary conditions, we use the following symbol, which intuitively approximates the true symbol at scales above $2^{-\mathscr{L}}$:

Definition 4.4 (Approximate symbol Dirac fermion, periodic case)**.** For any approximate Hilbert pair and $\mathscr{L} \in \mathbb{N}$, define the *approximate periodic symbol* as the following projection on $L^2(\mathbb{S}^1) \otimes \mathbb{C}^2$:

$$\tilde{Q}^{\mathrm{per}}_{\mathscr{L}} := \alpha^{\mathrm{per},*}_{\mathscr{L}} W^{(\mathscr{L}),\mathrm{per},*} \big(P_w \otimes |+\rangle \langle +| + P_s \otimes |L\rangle \langle L| \big) W^{(\mathscr{L}),\mathrm{per}} \alpha^{\mathrm{per}}_{\mathscr{L}}, \qquad (4.28)$$

where $\alpha^{\mathrm{per}}_{\mathscr{L}} := \alpha^{h,\mathrm{per}}_{\mathscr{L}} \oplus \alpha^{g,\mathrm{per}}_{\mathscr{L}}$ and $W^{(\mathscr{L}),\mathrm{per}} := W^{(\mathscr{L}),\mathrm{per}}_h \oplus W^{(\mathscr{L}),\mathrm{per}}_g$ refer to the periodic versions as defined in Section 3.1.2; P_s projects onto the single scaling coefficient and $|L\rangle := \frac{1}{\sqrt{2}}(|0\rangle - i\,|1\rangle)$ to ensure compatibility with our choice for the Hilbert transform on constant functions. Had we made a different choice for the value of the Hilbert transform on constant functions this would only change the top level state (it is a well-known fact that the Dirac fermion with periodic boundary conditions has a two-fold ground state degeneracy). We also observe that $|L\rangle \langle L| = \frac{1}{2}(I + \sigma_2) = \frac{1}{2}(I + \gamma_{\mathrm{chir}})$ is the chiral projector, where $\gamma_{\mathrm{chir}} = \gamma_0 \gamma_1 = \sigma_2$. Had we chosen the convention $\mathrm{sgn}(0) = -1$ we would take the state $|R\rangle := \frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle)$, and $|R\rangle \langle R| = \frac{1}{2}(I - \gamma_{\mathrm{chir}})$.

Given Eq. (4.28), we start with

$$W^{(\mathscr{L}),\mathrm{per},*} \big(P_w \otimes |+\rangle \langle +| + P_s \otimes |L\rangle \langle L| \big) W^{(\mathscr{L}),\mathrm{per}} = U^{(\mathscr{L}),\mathrm{per},*}_{\mathrm{MERA}} P_{\mathrm{per}} U^{(\mathscr{L}),\mathrm{per}}_{\mathrm{MERA}},$$

for a suitably defined unitary $U^{(\mathscr{L}),\mathrm{per}}_{\mathrm{MERA}}$ and $P_{\mathrm{per}} = P_w \otimes |0\rangle \langle 0| + P_s \otimes |L\rangle \langle L|$. This is already a symbol on a finite-dimensional Hilbert space $\mathbb{C}^{2^{\mathscr{L}}} \otimes \mathbb{C}^2$. As before, $U^{(\mathscr{L}),\mathrm{per}}_{\mathrm{MERA}}$ is a product of unitaries, one for each layer, but now these unitaries will depend on the scale $j = 0, \dots, \mathscr{L} - 1$ (cf. Section 3.1.2). Since taking the periodization of composition of convolutions is the same as convolving their periodizations, we can obtain the unitary $U^{\mathrm{per},j}_{\mathrm{MERA}}$ for the j-th layer simply by 'periodizing' the two-local unitaries U_{MERA} and analogously construct the circuit. Just like the filters, the MERA layers become identical for sufficiently large j.

This leads to an approximation of the exact correlation functions $G^{\mathrm{per}}(\{O_i\})$ for periodic boundary conditions

$$G^{\mathrm{MERA},\mathrm{per}}_{\mathscr{L}}(\{O_i\}) := \langle \Omega | O^{\mathrm{MERA}}_{\mathrm{per},1} \cdots O^{\mathrm{MERA}}_{\mathrm{per},n} | \Omega \rangle,$$

where $O^{\mathrm{MERA}}_{\mathrm{per},i}$ is obtained from O_i by replacing $\Psi(f)$ by $\Psi^{\mathrm{per}}_{\mathrm{MERA}}(f) := a_{P_{\mathrm{per}}}(U^{(\mathscr{L}),\mathrm{per}}_{\mathrm{MERA}} \alpha^{\mathrm{per}}_{\mathscr{L}} f)$ and $\mathrm{d}\Gamma_Q(A)$ by $\mathrm{d}\Gamma^{\mathrm{per}}_{\mathrm{MERA}}(A) := \mathrm{d}\Gamma_{P_{\mathrm{per}}}(U^{(\mathscr{L}),\mathrm{per}}_{\mathrm{MERA}} \alpha^{\mathrm{per}}_{\mathscr{L}} A \alpha^{\mathrm{per},*}_{\mathscr{L}} U^{(\mathscr{L}),\mathrm{per},*}_{\mathrm{MERA}})$. As before, this can be interpreted as a correlation function of local operators in a fermionic MERA on a circle[1], and we will show in Theorem 4.7 that this gives an accurate approximation.

For anti-periodic boundary conditions on the circle, recall that the symbol was given by $T^* Q^{\mathrm{per}} T$ (see Eq. (4.10)). This means that we can compute correlation functions for anti-periodic boundary conditions with the same circuit as for the periodic fermion, but replacing α_j by $\alpha_j T$. We note that the smearing functions f in this case are naturally anti-periodic (they are sections of a nontrivial bundle), so Tf is periodic and our results apply.

---

[1] Given a 2-local circuit for a wavelet transform, as constructed in Section 3.5 it is not hard to see that taking the corresponding circuit with periodic boundary conditions will give the periodized version of the wavelet transform.

Finally we discuss the case of Majorana fermions. For simplicity, we only consider the case of the line (cf. Section 4.1.4). Suppose that we want to approximate a correlation function of the form

$$G^{\mathrm{maj}}(\{f_i\}) = \langle\Omega|\Phi(f_1)\dots\Phi(f_n)|\Omega\rangle,\tag{4.29}$$

where the smeared Majorana field is given by $\Phi(f) = a_0((I-Q)f) + a_0^*(CQf)$ in terms of the symbol Q of the free Dirac fermion, and the charge conjugation operator C defined in Eq. (4.13). Consider the self-dual CAR algebra on the range of $P' = P_w \otimes I_{\mathbb{C}^2}$ which is a subspace $\mathcal{H}'$ of $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2 \otimes \mathbb{C}^{\mathcal{L}+1}$ (that is, the subspace corresponding to the wavelet coefficients) with charge conjugation C' given by the anti-unitary operator on $\mathcal{H}'$ which acts by $x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in the second tensor factor and componentwise complex conjugation in the standard basis. Similarly to Eq. (4.26), define

$$\Phi_{\mathrm{maj}}^{\mathrm{MERA}}(f) = a_0\big((P'-P)U_{\mathrm{MERA}}^{(\mathcal{L})}\alpha_j f\big) + a_0^*\big(C'PU_{\mathrm{MERA}}^{(\mathcal{L})}\alpha_j f\big).$$

We note that the above formula defines a representation of the self-dual CAR algebra $\mathcal{A}_{\wedge}^{\mathrm{sd}}(\mathcal{H}')$ since, clearly, $C'P = (P'-P)C'$. As before, we can approximate the correlation function (4.29) by

$$G_{j,\mathcal{L}}^{\mathrm{MERA,maj}}(\{f_i\}) = \langle\Omega|\Phi_{\mathrm{maj}}^{\mathrm{MERA}}(f_1)\dots\Phi_{\mathrm{maj}}^{\mathrm{MERA}}(f_n)|\Omega\rangle,$$

which for compactly supported f_i can be computed by an ordinary fermionic MERA. Note that

$$C'U_{\mathrm{MERA}}^{(\mathcal{L})}\alpha_j(f) = U_{\mathrm{MERA}}^{(\mathcal{L})}C\alpha_j(f) = U_{\mathrm{MERA}}^{(\mathcal{L})}\alpha_j(Cf)$$

where, with a slight abuse of notation, also write C for the similarly defined operator on $\ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$. Thus, we can also implement $\Gamma^c(U_{MERA}^{(\mathcal{L})})$ as a circuit of Majorana fermions, mapping the state on $\mathcal{A}_{\wedge}^{\mathrm{sd}}(\mathcal{H}')$ corresponding to P to the state on $\mathcal{A}_{\wedge}^{\mathrm{sd}}(U_{\mathrm{MERA}}^{(\mathcal{L})}(\mathcal{H}'))$ with symbol $U_{\mathrm{MERA}}^{(\mathcal{L}),*}PU_{\mathrm{MERA}}^{(\mathcal{L})}$.

4.2.5   Scaling dimensions

For MERA tensor networks, it has been observed that the (local and global) symmetries of the underlying theory can be approximately implemented in terms of the tensor network itself [MV18b]. In particular, a single layer of the MERA should always correspond to a rescaling by a factor two, and as in Fig. 2.3, its eigenvalues should be related to scaling dimensions of the CFT. In the wavelet construction, the relation between a single MERA layer and rescaling is very explicit.

In fact, we can easily show that the operator corresponding to a fermionic field has exact scaling dimension $\frac{1}{2}$, as was already observed (but not proven) in [EW16]. For this, consider (formally) the Dirac fermion field $\Psi_i(x)$, where δ_x is a delta function centered at x and $i \in \{1,2\}$. Its MERA realization at scale $j \in \mathbb{Z}$ is given by

$$\Psi_i^{\mathrm{MERA}}(x) = a^*(\alpha_j(\delta_x \otimes |i\rangle)) = \sum_{k \in \mathbb{Z}} \bar{\phi}_{j,k}(x)a^*(|k\rangle \otimes |i\rangle).\tag{4.30}$$

Since the scaling functions are compactly supported, the right-hand side expression is well-defined and we take it as the definition of $\Psi_i^{\mathrm{MERA}}(x)$. Now note that the scaling

superoperator for a single MERA layer consists of a conjugation by the second quantization of U_{MERA} and a contraction with the quasi-free state with symbol $I_{\ell^2(\mathbb{Z})} \otimes |+\rangle\langle+|$ on the wavelet output. Thus, any creation operator $a^*(f)$ gets mapped to $a^*(P_s W f)$, where P_s denotes the projection onto the scaling modes. Using Eq. (4.30), it follows that the scaling superoperator maps

$$
\begin{aligned}
\Psi_1^{\text{MERA}}(x) &\mapsto \sum_{k\in\mathbb{Z}} \bar{\phi}_{j,k}(x) a^*(P_s W^h |k\rangle \otimes |1\rangle) = \sum_{k\in\mathbb{Z}} \bar{\phi}_{j,k}(x) a^*((\downarrow m(\overline{\hat{h}_s}) |k\rangle) \otimes |1\rangle) \\
&= \sum_{k\in\mathbb{Z}} \bar{\phi}_{j,k}(x) \sum_{n\in\mathbb{Z}} \bar{h}_s[k-2n] a^*(|n\rangle \otimes |1\rangle) = \sum_{n\in\mathbb{Z}} \bar{\phi}_{j-1,n}(x) a^*(|n\rangle \otimes |1\rangle) \\
&= \sum_{n\in\mathbb{Z}} 2^{-\frac{1}{2}} \bar{\phi}_{j,n}(\tfrac{x}{2}) a^*(|n\rangle \otimes |1\rangle) = 2^{-\frac{1}{2}} \Psi_1^{\text{MERA}}(\tfrac{x}{2}),
\end{aligned}
$$

where we used Eqs. (3.5) and (3.10). We can argue similarly for the other component, as well as for the adjoints. Thus, we conclude that a single MERA layer coarse-grains $\Psi^{\text{MERA}}(x) \mapsto 2^{-\frac{1}{2}} \Psi^{\text{MERA}}(\tfrac{x}{2})$. The interpretation is that a single layer of the MERA corresponds to a rescaling of the fields by a factor two (as it should) and that it *exactly* reproduces the correct scaling dimension of $\frac{1}{2}$ for the fermionic fields.

If the scaling function is differentiable, we see that by differentiating we get that if we let $\partial_x \Psi_i^{\text{MERA}}(x) := \sum_{k\in\mathbb{Z}} (\partial_x \bar{\phi})(2^j x - k) a^*(|k\rangle \otimes |i\rangle)$ it holds that

$$
\partial_x \Psi_i^{\text{MERA}}(x) \mapsto 2^{-\frac{1}{2}-1} \partial_x \Psi_i^{\text{MERA}}(\tfrac{x}{2})
$$

which should be interpreted as a *descendent*, in CFT language, with a scaling dimension of $\frac{3}{2}$. In fact, if the wavelet function ψ^g has K vanishing moments (or equivalently, a factor $(1 + e^{i\theta})^K$ in the scaling filters g_s [Mal08]), then there exist a vector $\phi^{g,l} \in \ell^2(\mathbb{Z})$ for $l = 1, \dots, K$ with the same support as ψ^g such that

$$
\frac{1}{2^l \sqrt{2}} \phi^{g,l}[m] = \sum_n g_s[n] \phi^{g,l}[2m - n]
$$

even if ϕ^g is not l time differentiable (note that $\phi^{g,l}$ is only defined at integer values), see Theorem 7.1 in [SN96], and similarly for $\phi^{h,l}$. This implies that if the wavelets have K vanishing moments, the MERA captures K descendents of the fermion fields exactly. At this point we observe that a wavelet filter leading to K vanishing moments must have support at least $2K$, so one needs (as expected) a larger circuit depth to capture more descendent scaling dimensions. In general the other scaling dimension of the theory are only approximately reproduced and it would be interesting to prove quantitative bounds (for example, using our Theorem 4.7).

4.2.6 Numerical examples

Since the quantum circuits we obtain are the second quantization of a single-particle circuit we can simulate them classically for high circuit depth (bond dimension). In Fig. 4.2, (a) and (b) we show approximations to the smeared two-point functions for the fermionic fields and for the stress-energy tensor for $K = L = 1$ and $K = L = 3$ in the wavelet construction, corresponding to MERA tensor networks with bond dimensions $\chi = 4$ and $\chi = 64$ respectively. Another statistic is the entanglement entropy of an interval. In order to define this one needs a cut-off, for which we use the wavelet
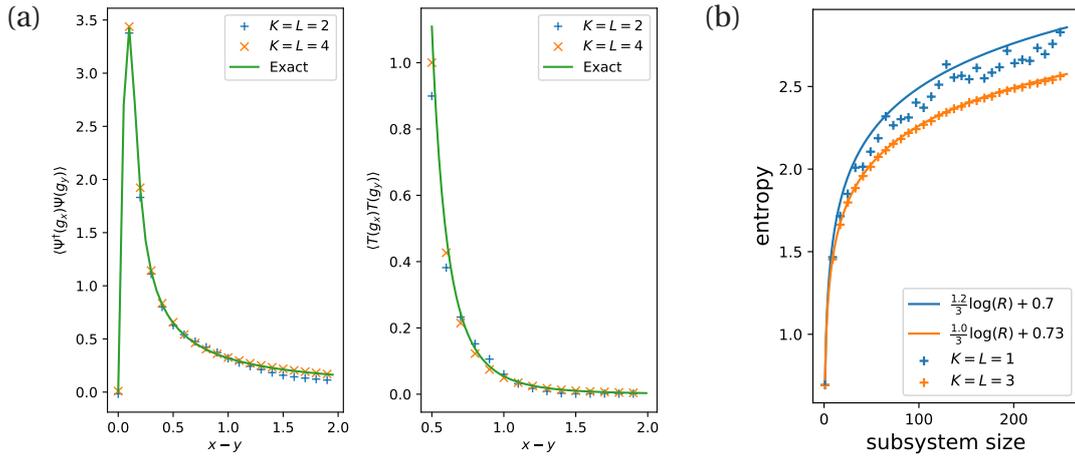
Figure 4.2: (a) Correlation function $\langle \Psi^*(g_x)\Psi(g_y)\rangle$ evaluated using our approximate quantum circuits where the smearing functions g_x, g_y are Gaussians with standard deviation $\sigma = 0.05$ peaked at x and y, respectively; correlation functions $\langle T(g_x)T(g_y)\rangle$ evaluated using our approximation quantum circuits. The stress-energy tensor is smeared in both space and time; see Section 4.3.2 for details. (c) Subsystem entropies for the corresponding quantum states. The logarithmic fits show that we obtain excellent agreement with the Cardy formula for central charge $c = 1$ already for $K + L = 6$.

discretization. The Cardy formula [Car86] predicts that for a conformal field theory the entanglement entropy of an interval scales as $S_E = \frac{c}{3}\log(L) + c'$ where c is the central charge, L is the size of the interval, and c' a non-universal constant depending on the cut-off. In Fig. 4.2, (c) we have plotted the entanglement entropies obtained from our construction (for the same wavelets). For $K = L = 3$ the agreement with the Cardy formula for $c = 1$ is already very accurate. As another numerical illustration of the accuracy of approximation, one may compute eigenvalues of the entanglement renormalization superoperator and extract scaling dimensions of the conformal field theory from its eigenvalues [PEV09]. One way to do so is by applying a Jordan-Wigner transformation to the circuit for the Majorana fermion to obtain a (matchgate) circuit for the Ising model. The results are illustrated in Table 4.2.

4.3   Approximation of correlation functions

In this section we prove our main technical result on the approximation of correlation functions of the Dirac fermion. To achieve this we adapt the approach pioneered in [HSW$^+$18], which proves the accuracy of the construction in Section 4.2.1, to the continuum setting. In Section 4.3.1 we discuss the approximation of the symbol by a wavelet construction. The key difference to [HSW$^+$18] is that we consider the symbol in the continuum setting, and we argue that the discretization maps α_j^g and α_j^h allow us to relate the continous and discrete symbols. We determine how the wavelet approximation results in Section 3.6 can be used to bound the corresponding errors. Moreover, we extend the results to the periodic case, and we provide a slightly improved scaling of the error bounds. In Section 4.3.2 we discuss how the approximation of the symbol leads to approximation of correlation functions. One main difference to [HSW$^+$18] is

	χ	E	$\Delta E / E$	Δ_σ	Δ_μ
Exact		$-\frac{4}{\pi}$		0.125	0.125
$K = 1, L = 1$	2	-1.2560	0.0135	0.0968	0.1696
$K = 1, L = 2$	4	-1.2705	0.0021	0.1360	0.1173
$K = 2, L = 1$	4	-1.2630	0.0081	0.1031	0.1563
$K = 1, L = 3$	8	-1.2727	0.0005	0.1226	0.1283
$K = 2, L = 2$	8	-1.2722	0.0008	0.1310	0.1204
$K = 3, L = 1$	8	-1.2655	0.0061	0.1052	0.1522
$K = 1, L = 4$	16	-1.2731	0.0001	0.1261	0.1242
$K = 2, L = 3$	16	-1.2731	0.0001	0.1238	0.1264

Table 4.2: Values of the Majorana fermion energy density E, the relative error in energy density $(E + \frac{\pi}{2})/E$ and the scaling dimensions Δ_σ and Δ_μ for the Majorana CFT (or, equivalently, of the Ising CFT). Some other scaling dimensions, in particular those of the fermion fields themselves, are exactly reproduced because of the structure of the wavelet transform, as discussed in Section 4.2.5. These values were computed by first decoupling the circuit obtained from an approximate Hilbert pair of wavelets for the Dirac fermion into two circuits for Majorana fermions, and then taking a Jordan-Wigner transform. This yields an entanglement renormalization circuit for the Ising model, in which the spin and disorder fields σ and μ are local. For details on this procedure, see [EW16].

that we need to consider the dependence on the smoothness of operators we insert, which follows naturally from our approximation of the symbol. The result we prove is also more general than the discrete result [HSW⁺18] since we also allow normal ordered quadratic operators (such as smearings of the stress-energy tensor) in the correlators.

4.3.1 Symbol approximations from Hilbert pairs

We will first show that Eq. (4.24) and Eq. (4.28) provide accurate approximations of free fermion symbols. We will restrict to compactly supported smearing functions $f \in L^2(\mathbb{R}) \otimes \mathbb{C}^2$; and we denote by $D(f)$ the size of a minimal interval which contains the support of f.

Lemma 4.5. *The following relation holds:* $\alpha_j^g m(\lambda_{s,j})^* \mathcal{H} = m(\mu_w)\alpha_j^g$. *Similarly, in the periodic case it holds for all* $f \in L^2(\mathbb{S}^1)$ *with zero mean that* $\alpha_j^{g,\text{per}} m(\lambda_{s,j}^{\text{per}})^* \mathcal{H} f = m(\mu_{w,j}^{\text{per}})\alpha_j^{g,\text{per}} f$.

Proof. We want to show that $\alpha_j^g(m(\lambda_{s,j})^* \mathcal{H} f) = m(\mu_w)\alpha_j^g(f)$ for $f \in L^2(\mathbb{R})$. By rescaling f it is easy to see that it suffices to show the result for $j = 0$. We know that by Eq. (3.25), $\mathcal{H} = m(\lambda_s)m(\mu_w)$, so $\alpha_0^g(m(\lambda_s)^* \mathcal{H} f) = \alpha_0^g(m(\mu_w)f)$. Next we take a Fourier transform and observe that

$$\widehat{\alpha_0^g(f)}(\theta) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \overline{\hat{\phi}^g(\theta + 2\pi n)} \hat{f}(\theta + 2\pi n).$$

Since μ_w is 2π-periodic the result follows. In the periodic case it holds that

$$\overline{\alpha_j^{g,\mathrm{per}}(f)}[n] = \sum_{m\in\mathbb{Z}} \overline{\hat{\phi}^{g,\mathrm{per}}[n+2^j m]}\, \hat{f}[n+2^j m].$$

which similarly implies the desired result. Note that the ambiguity in our choice of $\mathrm{sgn}(0)$ in the definition of $\mathcal{H}$ is not relevant if we assume that f has mean zero. $\blacksquare$

The following result shows that the symbols in Eq. (4.24) and Eq. (4.28) indeed yield reasonable approximations when restricted to appropriate functions.

Proposition 4.6. *Consider an ε-approximate Hilbert pair with scaling filters supported in $\{0,\dots,N-1\}$.*

(i) *Let $f \in H^1(\mathbb{R}) \otimes \mathbb{C}^2$ with compact support. Then, for all $j \in \mathbb{Z}$, $\mathcal{L} \in \mathbb{N}$, and $\mathcal{L}' = 0,\dots,\mathcal{L}$,*

$$\left\|\left(Q - \tilde{Q}_{j,\mathcal{L}}\right)f\right\| \le 3\varepsilon\mathcal{L}'\|f\| + 2^{(j-\mathcal{L}')/2}7\sqrt{N D(f)}B\|f\| + 2^{-j}5N^2\|f'\|,$$

where $B := \max\{\|\phi_g\|_\infty, \|\phi_h\|_\infty\}$.

(ii) *Let $f \in H^1(\mathbb{S}^1) \otimes \mathbb{C}^2$. Then, for all $\mathcal{L} \in \mathbb{N}$ and $\mathcal{L}' = 0,\dots,\mathcal{L}$,*

$$\left\|\left(Q^{\mathrm{per}} - \tilde{Q}_{\mathcal{L}}^{\mathrm{per}}\right)f\right\| \le 2\varepsilon\mathcal{L}'\|f\| + 2^{-\mathcal{L}'}9N^2\|f'\|.$$

In Theorem 4.7, we will describe how to choose j and $\mathcal{L}'$ optimally for a given number of layers $\mathcal{L}$.

Proof. (i) Let

$$\begin{aligned}
Q_j &:= \begin{pmatrix} I & 0 \\ 0 & \mathcal{H}^* m(\lambda_{s,j}) \end{pmatrix} \alpha_j^* \left(I_{\ell^2(\mathbb{Z})} \otimes |+\rangle\langle+|\right) \alpha_j \begin{pmatrix} I & 0 \\ 0 & m(\lambda_{s,j})^* \mathcal{H} \end{pmatrix} \\
&= \alpha_j^* \begin{pmatrix} I & 0 \\ 0 & m(\mu_w)^* \end{pmatrix} \left(I_{\ell^2(\mathbb{Z})} \otimes |+\rangle\langle+|\right) \begin{pmatrix} I & 0 \\ 0 & m(\mu_w) \end{pmatrix} \alpha_j,
\end{aligned} \tag{4.31}$$

where we used Lemma 4.5. Then, using the first formula,

$$\begin{aligned}
\|(Q - Q_j)f\| &\le \frac{1}{2}\Big(\|(I - \alpha_j^{h,*}\alpha_j^h)f_1\| + \|(I - m(\lambda_{s,j})\alpha_j^{g,*}\alpha_j^g m(\lambda_{s,j})^*)\mathcal{H}f_2\| \\
&\quad + \|(I - m(\lambda_{s,j})\alpha_j^{g,*}\alpha_j^h)f_1\| + \|(I - \alpha_j^{h,*}\alpha_j^g m(\lambda_{s,j})^*)\mathcal{H}f_2\| \Big) \\
&= \frac{1}{2}\Big(\|(I - P_j^h)f_1\| + \|(I - P_j^g)m(\lambda_{s,j})^*\mathcal{H}f_2\| \\
&\quad + \|(I - m(\lambda_{s,j})\alpha_j^{g,*}\alpha_j^h)f_1\| + \|(I - \alpha_j^{h,*}\alpha_j^g m(\lambda_{s,j})^*)\mathcal{H}f_2\| \Big) \\
&\le \frac{1}{2}\Big(\|(I - P_j^h)f_1\| + \|(I - P_j^h)\mathcal{H}f_2\| + \|(I - P_j^g)m(\lambda_{s,j})^* f_1\| \\
&\quad + \|(I - P_j^g)m(\lambda_{s,j})^*\mathcal{H}f_2\| + \|(\alpha_j^g m(\lambda_{s,j})^* - \alpha_j^h)f_1\| \\
&\quad + \|(\alpha_j^g m(\lambda_{s,j})^* - \alpha_j^h)\mathcal{H}f_2\| \Big).
\end{aligned}$$

The norms in the first line can be upper-bounded by using Lemma 3.3 (for the second, note that $\|(m(\lambda_{s,j})^* f_i)'\| = \|f_i'\|$ for $i = 1, 2$). For the norms in the second line we use Lemma 3.8. Together, we find that

$$
\begin{aligned}
\|(Q - Q_j) f\| &\leq \frac{1}{2} \left(2^{-j} C_{\mathrm{UV}} \left(2 \|f_1'\| + 2 \|\mathcal{H} f_2'\| \right) + 2^{-j} C_\chi \left(\|f_1'\| + \|\mathcal{H} f_2'\| \right) \right) \\
&\leq \frac{1}{2} 2^{-j} \left(2 C_{\mathrm{UV}} + C_\chi \right) \sqrt{2} \|f'\| \leq 2^{-j} 5 N^2 \|f'\|
\end{aligned}
\tag{4.32}
$$

where we used that the Hilbert transform preserves the norm of the derivative ($\|\mathcal{H} f_2'\| = \|f_2'\|$).

Next, we define

$$
Q_{j,\mathscr{L}} := \alpha_j^* \begin{pmatrix} I & 0 \\ 0 & m(\mu_w)^* \end{pmatrix} \left(W_h^{(\mathscr{L}),*} P_w W_h^{(\mathscr{L})} \otimes |+\rangle\langle+| \right) \begin{pmatrix} I & 0 \\ 0 & m(\mu_w) \end{pmatrix} \alpha_j.
$$

Using the second expression in Eq. (4.31), we can then split the remaining error as

$$
\|(Q_j - \tilde{Q}_{j,\mathscr{L}}) f\| \leq \|(Q_j - Q_{j,\mathscr{L}'}) f\| + \|(Q_{j,\mathscr{L}'} - \tilde{Q}_{j,\mathscr{L}'}) f\| + \|(\tilde{Q}_{j,\mathscr{L}'} - \tilde{Q}_{j,\mathscr{L}}) f\|
\tag{4.33}
$$

The third term in Eq. (4.33) can be estimated using Lemma 3.5:

$$
\begin{aligned}
\|(\tilde{Q}_{j,\mathscr{L}'} - \tilde{Q}_{j,\mathscr{L}}) f\| &\leq \|\alpha_j^* W^{(\mathscr{L}'),*} (P_s \otimes |+\rangle\langle+|) W^{(\mathscr{L})} \alpha_j f\| \leq \|P_{j-\mathscr{L}'} f\| \\
&\leq 2^{(j-\mathscr{L}')/2} \sqrt{ND(f)} \max\{\|\phi_g\|_\infty, \|\phi_h\|_\infty\} (\|f_1\| + \|f_2\|) \\
&\leq 2^{(j-\mathscr{L}')/2} \sqrt{2} \sqrt{ND(f)} B \|f\|.
\end{aligned}
$$

For the second term in Eq. (4.33), we use Eq. (3.34) in Lemma 3.6:

$$
\|Q_{j,\mathscr{L}'} - \tilde{Q}_{j,\mathscr{L}'}\| \leq \|P_w (W_h^{(\mathscr{L})} m(\mu_w) - W_g^{(\mathscr{L})})\| + \|m(\mu_w)^* W_h^{(\mathscr{L}),*} - W_g^{(\mathscr{L}),*}\| \leq 2\varepsilon \mathscr{L}'
$$

Finally, for the first term in Eq. (4.33), we would like to apply Lemma 3.5, but we need to be careful because $m(\mu_w)$ does not preserve compact support. So we first use Eq. (3.35) in Lemma 3.6 to get rid of $m(\mu_w)$, and then apply Lemma 3.5:

$$
\begin{aligned}
\|(Q_j - Q_{j,\mathscr{L}'}) f\| &= \|(P_s \otimes |+\rangle\langle+|) W_h^{(\mathscr{L}')} \begin{pmatrix} I & 0 \\ 0 & m(\mu_w) \end{pmatrix} \alpha_j f\| \\
&\leq \|P_s (W_h^{(\mathscr{L}')} m(\mu_w) - W_g^{(\mathscr{L}')}) \alpha_j^g f\| + \|(P_s \otimes I) W^{(\mathscr{L}')} \alpha_j f\| \\
&\leq \varepsilon \mathscr{L}' \|\alpha_j^g f_2\| + 2 \|P_s W_g^{(\mathscr{L}')} \alpha_j^g f\| + \|(P_s \otimes I) W^{(\mathscr{L}')} \alpha_j f\| \\
&\leq \varepsilon \mathscr{L}' \|f\| + 3 \left(\|P_{j-\mathscr{L}'}^g f_2\| + \|P_{j-\mathscr{L}'}^h f_1\| \right) \\
&\leq \varepsilon \mathscr{L}' \|f\| + 2^{(j-\mathscr{L}')/2} 5 \sqrt{ND(f)} B \|f\|.
\end{aligned}
$$

Thus, we can upper bound Eq. (4.33) by

$$
\|(Q_j - \tilde{Q}_{j,\mathscr{L}}) f\| \leq 3\varepsilon \mathscr{L}' \|f\| + 2^{(j-\mathscr{L}')/2} 7 \sqrt{ND(f)} B \|f\|.
\tag{4.34}
$$

Combining Eqs. (4.32) and (4.34) we obtain the desired bound.

(ii) Using $\phi_{0,1}^{g,\text{per}} = \phi_{0,1}^{h,\text{per}} = \mathbf{1}$, it is easy to see that our choice of input to the scaling layer ensures that

$$Q^{\text{per}}\mathbf{1} = \tilde{Q}_{\mathscr{L}}^{\text{per}}\mathbf{1},$$

so we can assume without loss of generality that f has zero mean or, equivalently, that $P_0 f = 0$ and we may apply Lemma 4.5. Similarly as before (but without having to worry about an IR cut-off), we introduce

$$Q_{\mathscr{L}}^{\text{per}} := \alpha_{\mathscr{L}}^{\text{per},*} \begin{pmatrix} I & 0 \\ 0 & m(\mu_w)^* \end{pmatrix} \left(W_h^{(\mathscr{L}),\text{per},*} P_w W_h^{(\mathscr{L}),\text{per}} \otimes |+\rangle\langle+| \right) \begin{pmatrix} I & 0 \\ 0 & m(\mu_w) \end{pmatrix} \alpha_{\mathscr{L}}^{\text{per}}$$

and use a triangle inequality

$$\left\|\left(Q^{\text{per}} - \tilde{Q}_{\mathscr{L}}^{\text{per}}\right)f\right\| \leq \left\|\left(Q^{\text{per}} - Q_{\mathscr{L}'}^{\text{per}}\right)f\right\| + \left\|\left(Q_{\mathscr{L}'}^{\text{per}} - \tilde{Q}_{\mathscr{L}'}^{\text{per}}\right)f\right\| + \left\|\left(\tilde{Q}_{\mathscr{L}'}^{\text{per}} - \tilde{Q}_{\mathscr{L}}^{\text{per}}\right)f\right\|.$$

For the first term, we use Lemmas 3.3 and 3.8 and obtain

$$\left\|\left(Q^{\text{per}} - Q_{\mathscr{L}'}^{\text{per}}\right)f\right\| \leq 2^{-\mathscr{L}'} 5N^2 \|f'\|,$$

in complete analogy to Eq. (4.32). For the second term, note that we can ignore the scaling part in Eq. (4.28) since we assumed that $P_0 f = 0$. Thus, we can use Eq. (3.38) in Lemma 3.7 and find

$$\|Q_{\mathscr{L}'}^{\text{per}} - \tilde{Q}_{\mathscr{L}'}^{\text{per}}\| \leq 2\varepsilon\mathscr{L}'.$$

Finally, the third term can be upper bounded by using Lemma 3.3,

$$\left\|\left(\tilde{Q}_{\mathscr{L}'}^{\text{per}} - \tilde{Q}_{\mathscr{L}}^{\text{per}}\right)f\right\| \leq \left\|(I - P_{\mathscr{L}'}^{\text{per}})f\right\| \leq 2^{-\mathscr{L}'}\sqrt{2}C_{\text{UV}}\|f'\| \leq 2^{-\mathscr{L}'} 4N^2 \|f'\|$$

(note that here we are comparing different UV cut-offs, in contrast to before). By combining these bounds we obtain the desired result. ∎

If we keep track of all the wavelet constants in the proof of Proposition 4.6 rather than bounding them in terms of N then the proof shows in fact the bound

$$\left\|\left(Q - \tilde{Q}_{j,\mathscr{L}}\right)f\right\| \leq 3\varepsilon\mathscr{L}'\|f\| + 2^{(j-\mathscr{L}')/2} 7\sqrt{D(f)}C_{\text{IR}}\|f\| + 2^{-j}\frac{1}{\sqrt{2}}(2C_{\text{UV}} + C_\chi)\|f'\|, \quad (4.35)$$

which will be useful if we want to investigate numerically how fast our error bounds converge.

We note that the error bounds in Proposition 4.6 are closely related to the quantum error correcting properties or MERA, see [KK17] which argues that if one encodes a quantum state by applying an entanglement renormalization circuit to it (so the information is encoded in the 'IR degrees of freedom') this is insensitive to localized perturbations.

4.3.2 Approximation bounds for correlation functions

The bounds on the approximate symbol from Proposition 4.6 can be used to estimate the approximation error for correlation functions. We start with the Dirac fermion on the line, whose vacuum state is the quasi-free state ω_Q with symbol Q defined in Eq. (4.9). We are interested in correlation functions of the form involving the smeared Dirac field $\Psi(f)$ and normal-ordered quadratic operators. In the Fock representation, the two-component Dirac field is implemented by the operators $\Psi(f) := a_Q(f)$, defined as in Eq. (4.1), and the normal-ordered quadratic operators the $\mathrm{d}\Gamma_Q(A)$ defined in Section 4.1.2. Thus, we wish to approximate correlation functions of the form

$$G(\{O_i\}) := \langle \Omega | O_1 \cdots O_n | \Omega \rangle, \tag{4.36}$$

where each O_i is either a component of $\Psi(f)$ or its adjoint $\Psi^*(f)$, or a normal-ordered operator $\mathrm{d}\Gamma_Q(A)$.

We would like to approximate such correlation functions by using the symbol $\tilde{Q}_{j,\mathscr{L}}$ defined in Eq. (4.24). Thus we fix an approximate Hilbert pair, $j \in \mathbb{Z}$, and $\mathscr{L} > 0$, and consider

$$\tilde{G}_{j,\mathscr{L}}(\{O_i\}) := \langle \Omega | \tilde{O}_1 \cdots \tilde{O}_n | \Omega \rangle, \tag{4.37}$$

where the $\tilde{O}_i$ are obtained from the O_i by replacing $\Psi(f)$ by $\check{\Psi}(f) := a_{\tilde{Q}_{j,\mathscr{L}}}(P_j f)$ and by replacing $\mathrm{d}\Gamma_Q(A)$ by $\mathrm{d}\Gamma_{\tilde{Q}_{j,\mathscr{L}}}(P_j A P_j)$.

On the circle, we denote the corresponding correlation functions for periodic boundary conditions by $G^{\mathrm{per}}(\{O_i\})$ and $\tilde{G}_{\mathscr{L}}^{\mathrm{per}}(\{O_i\})$, respectively. They are defined in terms of the symbol Q^{per} and its approximation $\tilde{Q}_{\mathscr{L}}^{\mathrm{per}}$ defined in Eq. (4.28).

The following theorem is our main technical result. It states that $G(\{O_i\}) \approx \tilde{G}_{j,\mathscr{L}}(\{O_i\})$ under appropriate conditions (and similarly in the periodic case), which shows that entanglement renormalization circuits can accurately compute correlation functions, justifying the construction in Section 4.2.3.

Theorem 4.7. *Consider an ε-approximate Hilbert pair with scaling filters supported in $\{0, \ldots, N-1\}$, scaling functions bounded by B, and $\varepsilon \in (0,1)$.*

(i) *Let $f_1, \ldots, f_n$ be compactly supported functions in $H^1(\mathbb{R}) \otimes \mathbb{C}^2$ and let $A_1, \ldots, A_m$ be Hilbert-Schmidt integral operators with compactly supported kernels in $H^1(\mathbb{R}^2) \otimes M_2(\mathbb{C})$, all with L^2-norm at most 1. Let $O_i = \Psi(f_i)$ or $\Psi^*(f_i)$ for $i = 1, \ldots, n$ and $O_{n+i} = \mathrm{d}\Gamma_Q(A_i)$ for $i = 1, \ldots, m$. Then we can find, for every $\mathscr{L} > 0$, a scale $j \in \mathbb{Z}$ such that*

$$\left| G(\{O_i\}) - \tilde{G}_{j,\mathscr{L}}(\{O_i\}) \right| \leq 8^m m! (n+m) \left(6\varepsilon \log_2 \frac{3C^3 D}{\varepsilon} + C D^{1/3} 2^{-\frac{\mathscr{L}}{3}} \right).$$

The constant $C := 14(\sqrt{2N}B + N^2)$ depends only on the Hilbert pair, and the constant $D := \max\{1, d(f, A) D(f, A)\}$ depends only on the smoothness and support of the smearing functions, where $d(f, A) := \max\{\|f_i'\|, \|\nabla A_i\|\}$ and $D(f, A) := \max\{D(f_i), D(A_i)\}$; ∇A_i denotes the gradient of the kernel of A_i and $D(A_i)$ denotes the side length of the smallest square supporting the kernel.

(ii) *Let $f_1, \ldots, f_n$ be functions in $H^1(\mathbb{S}^1) \otimes \mathbb{C}^2$ and let $A_1, \ldots, A_m$ be Hilbert-Schmidt integral operators with kernels in $H^1(\mathbb{S}^1) \otimes M_2(\mathbb{C})$, all with L^2-norm at most 1. Then we have, for every $\mathscr{L} > 0$, that*

$$\left| G^{\mathrm{per}}(\{O_i\}) - \tilde{G}_{\mathscr{L}}^{\mathrm{per}}(\{O_i\}) \right| \leq 8^m \, m! (n+m) \left(6\varepsilon \log_2 \frac{59 N^2 D}{\varepsilon} + 26 N^2 D 2^{-\mathscr{L}} \right).$$

The constant D is defined as $D := \max\{1, \|f_i'\|, \|\nabla A_i\|\}$, with ∇A_i the gradient of the kernel of A_i.

Before giving the proof, we comment on some aspects of the theorem. The main idea behind the theorem and its proof is that the approximation of the correlation functions is accurate as long as the approximation to the symbol is accurate on the scales at which the system is probed. Quite intuitively, large support requires us to accurately approximate large scales, and strong fluctuations (large derivatives) require accuracy at small scales. The constant $D = \max\{1, d(f, A)D(f, A)\}$ reflects the number of scales needed for accurate approximation for given smearing functions f_i and kernels A_i. Intuitively, D is invariant under dilatations, reflecting the scale invariance of the theory. On the circle $\mathbb{S}^1$, there is a natural largest scale, allowing for a slightly simpler formulation. While we state the theorem for the Dirac fermion, Proposition 4.6 readily implies a similar result for correlation functions of the Majorana fermion (Section 4.1.4).

Our assumptions on the operators A_i imply that they are in fact trace class. Thus, the operators $\mathrm{d}\Gamma(A_i)$ and $\mathrm{d}\Gamma_Q(A_i)$ can be directly defined in the CAR algebra, so we could work directly with the state ω_Q on the algebra rather than in the Fock space representation. Such an approach could improve the dependence on m of the bounds, since one can estimate $\|\mathrm{d}\Gamma_Q(A_i)\| = \|\mathrm{d}\Gamma(A_i) - \omega_Q(\mathrm{d}\Gamma(A_i))\| \leq 2\|A_i\|_1$.

While in Theorem 4.7 we order the insertions in $G(\{O_i\})$ in a particular way, other orderings are also possible. This follows either from using the commutation relations (leading to terms depending on $A_k f_l$) or by directly adjusting the proof (leading to a change in the dependence on n and m, since in the proof we would insert the particle-number projections Π_{2k} in different places).

Theorem 4.7 takes an approximate Hilbert pair as input. While, as discussed in Section 3.4, we are not aware of a rigorous proof for the existence of approximate Hilbert pair with a superpolynomial scaling between the filter error and the wavelet support, the Selesnick construction does provide a family of wavelet pairs which numerically have all the desirable properties. For this reason, we kept careful track of the constants in our error bounds.

We note that in the proofs of both Proposition 4.6 and Theorem 4.7 we bound the wavelet parameters C_{UV}, C_{IR}, and C_χ from Lemmas 3.3, 3.5 and 3.8 in terms of the support N to arrive at simpler expressions. Sharper numerical bounds can be obtained by using C_{UV}, C_{IR}, and C_χ directly (see Table 3.1 for the Selesnick construction). If one tracks these constants throughout the proof, using Eq. (4.35) rather than Proposition 4.6, one sees that C can be taken to be

$$C = 2(4C_{\mathrm{UV}} + C_\chi) + 20 C_{\mathrm{IR}}. \tag{4.38}$$

The precise numerical constants are not very important, but we can use this to illustrate Theorem 4.7 numerically for two-point functions (using Table 3.1 to evaluate Eq. (4.38)),
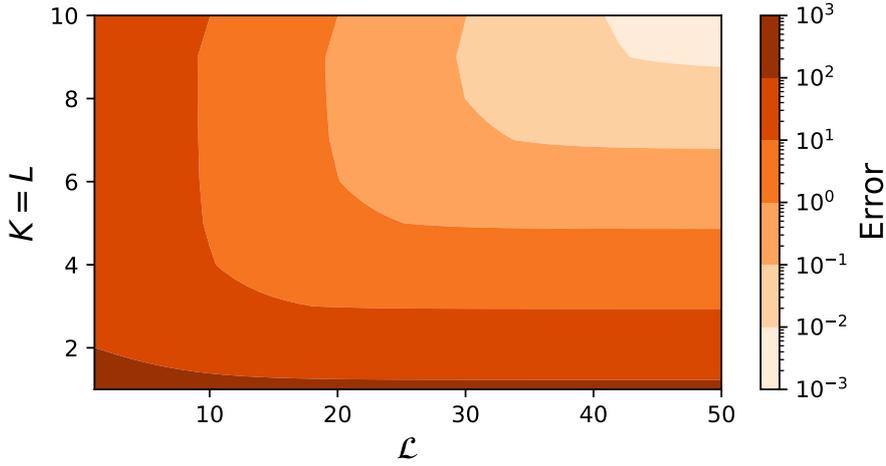
Figure 4.3: The error bound from Theorem 4.7 illustrated for a two-point function. It is obtained by evaluating Eq. (4.38) using Table 3.1 for an approximate Hilbert pair with parameters $K = L$. The smearing functions are taken to be translates of a function f with $\|f\| = 1$ and optimal trade-off between smoothness and support (that is, $D = \sqrt{2}$ in the formulation of Theorem 4.7).

see Fig. 4.3. We see that, even for relatively small circuit depth, our Theorem 4.7 combined with numerical results of Table 3.1 yields a reasonably small upper bound on the approximation error.

Proof of Theorem 4.7. (i) We first estimate the error in the correlation functions in terms of the corresponding symbols for fixed $j \in \mathbb{Z}$ and $\mathscr{L}' \in \{0, \dots, \mathscr{L}\}$. We define $Q_- := Q$, $Q_+ := I - Q$, $\tilde{Q}_- := \tilde{Q}_{j,\mathscr{L}}$, and $\tilde{Q}_+ := P_j - \tilde{Q}_{j,\mathscr{L}}$ (!). For $i = 1, \dots, n$,

$$\|O_i - \tilde{O}_i\| = \|a_Q(f_i) - a_{\tilde{Q}_{j,\mathscr{L}}}(P_j f_i)\| \le \|(Q_+ - \tilde{Q}_+) f_i\| + \|(Q_- - \tilde{Q}_-) f_i\|,$$

where we used the definition of the operators $\tilde{O}_i$ described above, Eq. (4.1) and the fact that $\tilde{Q}_{j,\mathscr{L}} P_j = P_j \tilde{Q}_{j,\mathscr{L}} = \tilde{Q}_{j,\mathscr{L}}$. By Proposition 4.6, we have the estimate

$$\|(Q_- - \tilde{Q}_-) f_i\| \le 3\varepsilon \mathscr{L}' \|f_i\| + 2^{(j-\mathscr{L}')/2} 7 \sqrt{ND(f_i)} B \|f_i\| + 2^{-j} 5N^2 \|f_i'\|.$$

Moreover, using Lemma 3.3,

$$\|(Q_+ - \tilde{Q}_+) f_i\| \le \|P_j f_i - f_i\| + \|(Q_- - \tilde{Q}_-) f_i\| \le 2^{-j} 4N^2 \|f_i'\| + \|(Q_- - \tilde{Q}_-) f_i\|.$$

Thus we find that

$$\|O_i - \tilde{O}_i\| \le 6\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 14 \sqrt{ND(f_i)} B + 2^{-j} 14N^2 \|f_i'\| \tag{4.39}$$

using $\|f_i\| \le 1$. For $i = n+1, \dots, n+m$, if we let Π_n denote the projection onto the n-particle subspace of the Fock space then by Eq. (4.4) we have the bound

$$\|(O_i - \tilde{O}_i) \Pi_{2k}\| \le 4(2k+2) \max_{\delta=\pm} \{\|Q_\delta A_i Q_\delta - \tilde{Q}_\delta A_i \tilde{Q}_\delta\|, \|Q_\delta A_i Q_{-\delta} - \tilde{Q}_\delta A_i \tilde{Q}_{-\delta}\|_2\}$$

$$\le 4(2k+2) \max_{\delta=\pm} \{\|(Q_\delta - \tilde{Q}_\delta) A_i\|_2 + \|A_i (Q_\delta - \tilde{Q}_\delta)\|_2\}.$$

To estimate $\|(Q_\delta - \tilde{Q}_\delta) A_i\|_2$, let $\{e_n\}$ be an orthonormal basis of $L^2(\mathbb{R}) \otimes \mathbb{C}^2$, so

$$\|(Q_\delta - \tilde{Q}_\delta) A_i\|_2^2 = \sum_n \|(Q_\delta - \tilde{Q}_\delta) A_i e_n\|^2$$

$$\leq \sum_n \left(\left(3\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 7\sqrt{ND(A_i)} B \right) \|A_i e_n\| + 2^{-j} 9N^2 \|(A_i e_n)'\| \right)^2$$

using Proposition 4.6 and Lemma 3.3 (for $\delta = +$) and the fact that, by our assumption on the support of the kernel of A_i, the support of $A_i e_n$ is contained in an interval of size $D(A_i)$. Since A_i has a kernel h_i in $H^1(\mathbb{R}) \otimes M_2(\mathbb{C})$, it holds that $(A_i e_n)' = (\partial_x A_i) e_n$, where $\partial_x A_i$ denotes the integral operator with kernel $\partial_x h_i$. Thus, we conclude using Cauchy-Schwarz

$$\|(Q_\delta - \tilde{Q}_\delta) A_i\|_2 \leq 3\varepsilon \mathscr{L}' \|A_i\|_2 + 2^{(j-\mathscr{L}')/2} 7\sqrt{ND(A_i)} B \|A_i\|_2 + 2^{-j} 9N^2 \|\partial_x A_i\|_2.$$

Since the adjoint of an integral operator has the transposed and conjugated kernel, we obtain the same bound on $\|A_i(Q_\delta - \tilde{Q}_\delta)\|_2 = \|(Q_\delta - \tilde{Q}_\delta) A_i^*\|_2$ but with $\|\partial_y A_i\|$ in place of $\|\partial_x A_i\|$, and hence

$$\left\| \left(O_i - \tilde{O}_i \right) \Pi_{2k} \right\| \leq 4(2k+2) \left(6\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 7\sqrt{ND(A_i)} B + 2^{-j} 14N^2 \|\nabla A_i\|_2 \right) \quad (4.40)$$

using $\|A_i\|_2 = \|h_i\| \leq 1$, and where we have written ∇A_i for the operator which has the gradient of h_i as kernel. To estimate the error in the correlation functions, we use a telescoping sum

$$\left| G(\{O_i\}) - \tilde{G}_{j,\mathscr{L}}(\{O_i\}) \right| \leq \sum_{i=1}^{n+m} \delta_i, \quad (4.41)$$

where

$$\delta_i = |\langle \Omega | O_1 \cdots O_{i-1} (O_i - \tilde{O}_i) \tilde{O}_{i+1} \cdots \tilde{O}_{n+m} | \Omega \rangle|.$$

Now, $\|O_i\| \leq 1$ for $i = 1, \ldots, n$ by $\|f_i\| \leq 1$. For $i = 1, \ldots, m$, we can replace O_{n+i} by $O_{n+i} \Pi_{2(m-i)}$, and similarly for $\tilde{O}_{n+i}$. Since $\|O_{n+i} \Pi_{2(m-i)}\| \leq 8(m-i+1)$ by Eq. (4.3) and $\|A_{n+i}\|_2 \leq 1$, we find that, for $i = 1, \ldots, n$,

$$\delta_i \leq 8^m m! \left(6\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 14\sqrt{ND(f_i)} B + 2^{-j} 14N^2 \|f_i'\| \right)$$

by Eq. (4.39) and, for $i = 1, \ldots, m$,

$$\delta_{n+i} \leq 8^m m! \left(6\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 14\sqrt{ND(A_i)} B + 2^{-j} 14N^2 \|\nabla A_i\| \right)$$

by Eq. (4.40). If we plug these bounds into Eq. (4.41) we obtain

$$\left| G(\{O_i\}) - \tilde{G}_{j,\mathscr{L}}(\{O_i\}) \right|$$
$$\leq 8^m m!(n+m) \left(6\varepsilon \mathscr{L}' + 2^{(j-\mathscr{L}')/2} 14\sqrt{ND(f, A)} B + 2^{-j} 14N^2 d(f, A) \right), \quad (4.42)$$

where we used the definitions of $D(f, A)$ and $d(f, A)$. We have thus obtained a bound on the approximation error which holds for all $j \in \mathbb{Z}$ and $\mathscr{L}' = 0, \ldots, \mathscr{L}$.

We now choose j and $\mathscr{L}'$ to obtain that vanishes as the number of layers $\mathscr{L}$ increases and ε goes to zero. We first choose $j = \lceil \frac{\mathscr{L}'}{3} + \frac{1}{3}\log_2\frac{d(f,A)^2}{D(f,A)}\rceil$ and obtain

$$\left|G(\{O_i\}) - \tilde{G}_{j,\mathscr{L}}(\{O_i\})\right| \leq 8^m m!(n+m)\Big(6\varepsilon\mathscr{L}'$$
$$+ 14(\sqrt{2N}B + N^2)d(f,A)^{1/3}D(f,A)^{1/3}2^{-\frac{\mathscr{L}'}{3}}\Big)$$
$$= 8^m m!(n+m)\Big(6\varepsilon\mathscr{L}' + CD^{1/3}2^{-\frac{\mathscr{L}'}{3}}\Big),$$

using the definitions of C and D. We now choose $\mathscr{L}' = \min\{\mathscr{L}, \lceil \log_2(C^3D/\varepsilon)\rceil\}$, which is always nonnegative, and obtain

$$\left|G(\{O_i\}) - \tilde{G}_{j,\mathscr{L}}(\{O_i\})\right| \leq 8^m m!(n+m)\left(6\varepsilon\left(\log_2\frac{C^3D}{\varepsilon} + 1\right) + \max\{CD^{1/3}2^{-\frac{\mathscr{L}}{3}}, \varepsilon\}\right)$$
$$\leq 8^m m!(n+m)\left(6\varepsilon\log_2\frac{3C^3D}{\varepsilon} + CD^{1/3}2^{-\frac{\mathscr{L}}{3}}\right),$$

which proves the desired bound.

(ii) The proof for the circle goes along the same lines using the corresponding bound from Proposition 4.6 and $j = \mathscr{L}$. Instead of Eqs. (4.39) and (4.40), we find that, for all $\mathscr{L}' \in \{0,\ldots,\mathscr{L}\}$ and for $i = 1,\ldots,n$,

$$\|O_i - \tilde{O}_i\| \leq 4\varepsilon\mathscr{L}' + 2^{-\mathscr{L}'}18N^2\|f_i'\| + 2^{-\mathscr{L}}4N^2\|f_i'\| \leq 4\varepsilon\mathscr{L}' + 2^{-\mathscr{L}'}22N^2\|f_i'\|,$$

while for $i = n+1,\ldots n+m$,

$$\|(O_i - \tilde{O}_i)\Pi_{2k}\| \leq 8(2k+2)\left(6\varepsilon\mathscr{L}' + 2^{-\mathscr{L}'}26N^2\|\nabla A_i\|\right).$$

Thus we obtain

$$\left|G^{\text{per}}(\{O_i\}) - \tilde{G}^{\text{per}}_{j,\mathscr{L}}(\{O_i\})\right| \leq 8^m m!(n+m)\left(6\varepsilon\mathscr{L}' + 26N^2D2^{-\mathscr{L}'}\right)$$

in place of Eq. (4.42). Finally, we choose $\mathscr{L}' = \min\{\mathscr{L}, \lceil\log_2\frac{26N^2D}{\varepsilon}\rceil\}$, which is always nonnegative, and arrive at

$$\left|G^{\text{per}}(\{O_i\}) - \tilde{G}^{\text{per}}_{\mathscr{L}}(\{O_i\})\right| \leq 8^m m!(n+m)\left(6\varepsilon\log_2\frac{59N^2D}{\varepsilon} + 26N^2D2^{-\mathscr{L}}\right).$$

This is the desired bound. ∎

To illustrate Theorem 4.7 and to show that the class of operators considered is an interesting class, we now describe how to compute correlation functions involving smeared stress-energy tensors. The stress-energy tensor is a fundamental object in conformal field theory. Its mode decomposition form two copies of the Virasoro algebra, encoding the conformal symmetry of the theory [FMS12]. It is convenient to choose a different basis and write the Dirac action in the form

$$S(\Psi) = \frac{1}{2}\int \Psi^*\begin{pmatrix} \bar{\partial} & 0 \\ 0 & \partial \end{pmatrix}\Psi\,\mathrm{d}x\mathrm{d}t$$

where $\partial = \partial_x + \partial_t$ and $\bar{\partial} = \partial_x - \partial_t$. Then, formally, the holomorphic component $T = T_{zz}$ of the stress-energy tensor, is the normal ordering of $\Psi_1^*\partial\Psi_1$. Solutions of the Dirac

equation in this basis are of the form $\chi(x, t) = \chi_+(x + t) \oplus \chi_-(x - t)$. The *unsmeared* stress energy tensor $T(x)$ (which is only a formal expression in the algebraic formalism) is given by $T(x) = \mathrm{d}\Gamma_Q(D_x)$ where

$$D_x \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} \delta_x f_1' \\ 0 \end{pmatrix}$$

where δ_x is a δ-function centered at x. To smear this operator, consider two smearing functions h_x and h_t. The h_t should be thought of as a smearing in the time direction and we use the Dirac equation to interpret this on our Hilbert space corresponding to $t = 0$. Thus, we define

$$D(h) \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} h_x(h_t \star f_1)' \\ 0 \end{pmatrix}$$

where $\star$ denotes convolution. We then define the *smeared* stress-energy tensor by the normal-ordered second quantization: $T(h) = \mathrm{d}\Gamma_Q(D(h))$. If h_x and h_t are compactly supported functions in $H^1(\mathbb{R})$, then the operator $T(h)$ satisfies the conditions of Theorem 4.7. In Fig. 4.2, (b) we show the numerical result of computing two-point functions $\langle T(h_1) T(h_2) \rangle$ using our quantum circuits, where the h_i are taken to be Gaussian smearing functions. In agreement with our theorem, we find that the two-point functions are approximated accurately for approximate Hilbert pairs of suitably good quality. (Strictly speaking, the Gaussians need to be approximated by compactly supported functions so that Theorem 4.7 applies.)

Discussion and open questions

One question which immediately arises is to extend the wavelet construction to general free models (not just the critical model described in the current chapter), such as massive fermions or bosonic models. In the next chapter we will see that this is straightforwardly possible for a general class of models in the bosonic case; in the fermionic case it is less clear how to do so. One challenge for more general free fermionic models is how to deal with ground states which are not at half-filling.

In future work we hope to construct entanglement renormalization circuits for more general classes of conformal field theories. A challenging open problem is to extend the relation between wavelet analysis and quantum circuits for conformal field theories to interacting models. It is not at all clear that this is possible, but a natural starting point could be Wess-Zumino-Witten theories, as many of these can be constructed algebraically as symmetries on a finite number of free massless fermions [Fuc95]. The algebraic construction of these theories is closely related to the representation theory of loop groups [Was98], and one starting point could be to revisit this analysis of the loop group in terms of wavelet theory. See [OS21a, OS21b] for a related perspective.

Another direction would be to investigate entanglement renormalization from the perspective of vertex algebras. A recent attempt to discretize vertex algebras to a spin chain model, with a view towards quantum computer simulation of conformal field theories can be found in [ZW18]. For MPS tensor networks it has been shown that they are sufficiently expressive to compute correlation functions for a very general class of conformal field theories using vertex algebra techniques [KS16, KS17b].

From a computational point of view it would be interesting to investigate whether a wavelet circuit can serve as a starting point for perturbation theory, and get faster convergence of MERA optimization algorithms.

Bosonic entanglement renormalization

In Section 5.1 we will briefly review the bosonic formalism. We will then explain the relation between entanglement renormalization and biorthogonal wavelet filters. In particular, in Section 5.2 we derive a relation the filters have to satisfy to disentangle the ground state of a given Hamiltonian and explain how this gives rise to a circuit. We also explain how this extends to quantum field theory correlation functions. Finally, we prove an approximation theorem for correlation functions for bosonic entanglement renormalization in Section 5.3. In this chapter we will occasionally be slightly less formal than in Chapter 4, on the one hand for the prosaic reason that [WW21c] was written with a physics audience in mind, but also in order to make sure that the emphasis is on concepts rather than formalism. With this motivation in mind we have also chosen to work out in detail the error bounds for approximation of lattice observables, to complement the fermionic setting, where we have focussed on the field theory.

5.1 Bosons and second quantization

In this section we will review second quantization for bosons and quasi-free (or Gaussian) many-body states. Further details may be found for instance in [BR03] or [Pet90]. We describe ground states of quadratic bosonic Hamiltonians on a one-dimensional lattice, and the vacuum state of a free bosonic field theory in this formalism. In the context of quantum information theory the bosonic formalism is also known as *continuous variable* quantum information.

5.1.1 The CCR algebra and Gaussian states

If $\mathcal{H}$ is a complex Hilbert space (which is the *single particle space*), let $\sigma(f,g) = \mathrm{Im}\langle f,g\rangle$ be the canonical symplectic form on the corresponding real vector space. Let $\mathscr{A}_\vee(\mathcal{H})$ be the algebra of canonical commutation relations or *CCR algebra* on $\mathcal{H}$ which is the free unital C^*-algebra generated by elements $W(f)$ for $f \in \mathcal{H}$ subject to the relations

$$W(f)W(g) = e^{-\frac{i}{2}\sigma(f,g)}W(g)W(f)$$
$$W(f)^* = W(-f).$$

To obtain a Hilbert space realization, we consider the Fock state which is defined by

$$\omega(W(f)) = e^{-\frac{\|f\|^2}{2}}$$

and we consider the corresponding GNS representation. This defines a semi-group $t \mapsto e^{tB(f)}$ for $t \in \mathbb{R}_{\geq 0}$, and by Stone's theorem there must exist a generator $B(f)$ such that $W(tf) = e^{tB(f)}$. The operators $B(f)$ will be unbounded and allows us to define the creation and annihilation operators

$$a(f) = \frac{1}{\sqrt{2}}(B(f) + iB(if))$$
$$a^*(f) = \frac{1}{\sqrt{2}}(B(f) - iB(if))$$

These satisfy the more familiar form of the canonical commutation relations

$$[a(f), a^*(g)] = \langle f, g \rangle.$$

With this interpretation, the GNS Hilbert space from the Fock state is in fact given by

$$\mathscr{F}_\vee(\mathscr{H}) = \bigoplus_{n=0}^{\infty} \mathscr{H}^{\vee n}.$$

with $a(f)$ and $a^*(f)$ acting as the usual annihilation and creation operators, acting as $a^*(f)v = f \vee v$. We let $|\Omega\rangle$ denote the Fock vacuum vector $1 \in \mathscr{H}^{\vee 0}$.

The most basic example is when $\mathscr{H} = \mathbb{C}$ with the standard inner product. Interpreting this as a real vector space, this is isomorphic to $\mathbb{R}^2$ and σ is the standard symplectic form. In this case, we may write $q = B(1)$ and $p = B(i)$, and $a = \frac{1}{\sqrt{2}}(q + ip)$ and $a^* = \frac{1}{\sqrt{2}}(q - ip)$, corresponding to the usual harmonic oscillator, satisfying $[q, p] = i$ and the commutation relation $[a, a^*] = 1$.

An important class of states on this algebra are the *gauge-invariant quasi-free* (or *Gaussian*) states. In continuous-variable terminology, the gauge-invariance condition corresponds to zero displacement Gaussian states. These states have the property that all correlation functions are determined by the two-point functions.

More precisely, let γ be a positive symmetric bilinear form defined on $\mathscr{H}$ as real Hilbert space, which is moreover such that $\gamma - i\sigma$ is positive semidefinite. Then we may define a state ω_γ which is such that

$$\omega_\gamma(W(f)) = e^{-\frac{1}{2}\gamma(f,f)}.$$

Then

$$\omega_\gamma(B(f)B(g)) = \gamma(f,g) + i\sigma(f,g) \tag{5.1}$$

and in particular

$$\gamma(f,g) = \frac{1}{2}(\omega_\gamma(B(f)B(g)) + B(g)B(f)).$$

Both the bilinear form γ and the two-point functions completely determine the state. We will typically identify γ with a real linear operator on $\mathscr{H}$ so the form is given by $\langle f, \gamma g \rangle$ and refer to it as the *covariance matrix*.

5.1.2 Second-quantized operators

Next we recall the second quantization of operators on $\mathcal{H}$. If S is a symplectic real linear map on $\mathcal{H}$ (now interpreting $\mathcal{H}$ as a real vector space), i.e.

$$\sigma(Sf, Sg) = \sigma(f, g)$$

for all $f, g \in \mathcal{H}$, then S gives rise to an automorphism of $\mathcal{A}_\vee(\mathcal{H})$, known as a *Bogoliubov transformation* or *Gaussian map*, through $W(f) \mapsto W(Sf)$. Gaussian states transform as $\omega_\gamma \mapsto \omega_{S^*\gamma}$, where $S^*\gamma = S\gamma S^\mathsf{T}$, under this transformation. We will in fact only need the setting where we write $\mathcal{H} = H \oplus iH$, with H a real vector space, and where we let $S = A \oplus (A^\mathsf{T})^{-1}$, which is symplectic for any invertible real linear map $A : H \to H$.

5.1.3 Free bosons

We will discuss two relevant models. The first is that of translation invariant chains of harmonic oscillators. We let $\mathcal{H} = \ell^2(\mathbb{Z})$, and we abbreviate $q_n = B(\delta_n)$ and $p_n = B(i\delta_n)$. We then consider quadratic Hamiltonians of the form

$$H = \frac{1}{2}\Big(\sum_{n\in\mathbb{Z}} p_n^2 + \sum_{n,m\in\mathbb{Z}} q_n V[n-m] q_m\Big), \tag{5.2}$$

where $V \in \ell^2(\mathbb{Z}, \mathbb{R})$ is such that $\hat{V}$ is a positive function (so the Hamiltonian is bounded from below). The ground state of such a quadratic Hamiltonian is a quasi-free state and can be written as ω_γ with

$$\gamma = \begin{pmatrix} \gamma^q & 0 \\ 0 & \gamma^p \end{pmatrix}$$

where we decompose $\mathcal{H} = \ell^2(\mathbb{Z}) = \ell^2(\mathbb{Z}, \mathbb{R}) \oplus i\ell^2(\mathbb{Z}, \mathbb{R})$, and γ^q and γ^p have matrix entries $\gamma^p_{nm} = \omega_\gamma(p_n p_m)$ and $\gamma^q_{nm} = \omega_\gamma(q_n q_m)$. Upon taking a Fourier transform we find that $\hat{\gamma}^p$ and $\hat{\gamma}^q$ are multiplication operators in the Fourier domain $\gamma^p = m(\frac{1}{2}E)$ and $\gamma^q = m(\frac{1}{2E})$ where $E(\theta) = \hat{V}(\theta)$ is the dispersion relation of the Hamiltonian.

A paradigmatic example is the harmonic chain with mass M,

$$H = \frac{1}{2}\Big(\sum_{n\in\mathbb{Z}} p_n^2 + M^2 q_n^2 + \frac{1}{4}(q_n - q_{n+1})^2\Big), \tag{5.3}$$

which has dispersion relation $E(\theta) = \sqrt{M^2 + \sin^2(\frac{\theta}{2})}$. In particular, the *massless* harmonic chain is gapless and has dispersion relation $E(\theta) = |\sin(\frac{\theta}{2})|$. For details about quadratic bosonic Hamiltonians and Gaussian states from the perspective of quantum information and computation, see for instance, [AEPW02, PEDC05, WM07, KLM01].

The second relevant model is the continuum limit of the harmonic chain with mass M, which yields a free bosonic field theory with mass M. We denote by $\mathcal{H}$ the single-particle space, which is an space of functions on the real line. Let the dispersion relation be given by $E(\omega) = \sqrt{\omega^2 + M^2}$. In order for the theory to be well-defined we need to impose some decay bounds on $\mathcal{H}$. We may take $\mathcal{H} = H_{-\frac{1}{2}, M}(\mathbb{R})$, which is a Sobolev space with inner product

$$\langle f, g \rangle_{-\frac{1}{2}, M} := \langle \hat{f}, \frac{1}{E}\hat{g} \rangle_{L^2(\mathbb{R})}.$$

Gaussian unitaries	$q_n \mapsto A_{n,m} q_m$ and $p_n \mapsto B_{n,m} p_m$ with $A^{-1} = B^{\mathsf{T}}$
Hamiltonian	$H = \frac{1}{2}\left(\sum_{n \in \mathbb{Z}} p_n^2 + \sum_{n,m \in \mathbb{Z}} q_n V_{n-m} q_m\right)$ with dispersion relation $E(\theta)$
Wavelet filters	(g, h) pair of biorthogonal wavelet filters
Filter relation	$\hat{g}_w(\theta) = E(\theta)\hat{h}_w(k)$
Application of wavelet transform	$A = W_g$, $B = W_h$ Apply squeezing to normalize dispersion relation, then apply the wavelet decomposition.
Continuum theory	Free bosonic scalar field (for the harmonic chain)
Wavelet functions	$\hat{\psi}^g(\omega) = \frac{\lvert\omega\rvert}{4}\hat{\psi}^h(\omega)$

Table 5.1: Overview of the construction of MERA from wavelets for bosons as described in the current chapter, to be compared with the corresponding results for fermionic systems in Table 4.1.

As we will not prove any rigorous approximation bounds for field theory correlation functions in this chapter we will further ignore technical subtleties relating to this construction and choice of Hilbert spaces, such as those occurring for $M = 0$, and refer to [GJ12] for technical details on the rigorous construction of bosonic quantum field theories. Given a real valued $f \in \mathcal{H}$ we let $\Phi(f) = B(f)$ and $\Pi(f) = B(if)$ (which are unbounded operators on the Fock space $\mathcal{F}_\vee(\mathcal{H})$). Formally, we may write $\Phi(x) = \Phi(\delta_x)$ and $\Pi(x) = \Pi(\delta_x)$ where δ_x is a δ-function centered at x. The free boson with mass M, which is the continuum limit of a harmonic chain with mass M, has Hamiltonian

$$H = \frac{1}{2}\int \Pi(x)^2 + M^2 \Phi(x)^2 + \left(\partial\Phi(x)\right)^2 \mathrm{d}x.$$

which has E as its dispersion relation. The covariance matrix is given by $\gamma^p = m(\frac{1}{2}E)$ and $\gamma^p = m(\frac{1}{2E})$ where now $E(\omega) = \sqrt{\omega^2 + M^2}$ (and note that these operators are now well-defined by our choice of $\mathcal{H}$). We are particularly interested in the massless case, which gives rise to a conformal field theory.

5.2   Entanglement renormalization circuits

In this section we investigate how second quantization of an an appropriate biorthogonal wavelet transform gives rise to entanglement renormalization circuits for free bosons. A general discussion of entanglement renormalization for free bosons can be found in [EV10a], which introduced the notion of free bosonic entanglement renormalization and provides variational algorithms for finding such circuits. Table 5.1 provides an overview of the bosonic wavelet MERA, also compare Table 4.1.

5.2.1   Entanglement renormalization for lattice bosons

As in the previous chapter, we define a Gaussian circuit to be the second quantization of single-particle symplectic circuit. Here, as we consider systems in one spatial dimension, a single-particle (symplectic) circuit, of depth N is a symplectic map $S: \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ which can be written as $S = S_N \circ \ldots S_1$ where each S_i is such that acts strictly locally on nearest neighbour sites, as in Section 3.5 In contrast to the fermionic case, the involved

Hilbert spaces are infinite dimensional, even when restricting to a finite number of sites and there is no exact mapping to a finite number of qubits such as the Jordan-Wigner transformation. On the other hand, these circuits are such that they could be implemented directly on a continuous-variable based quantum computer, for instance using linear optics [KLM01].

Given a quadratic Hamiltonian as in Eq. (5.2) we would like to construct a Gaussian circuit which maps the ground state of Eq. (5.2) to an unentangled state on the odd sublattice and to the ground state of a new Hamiltonian on the even sublattice. If the original Hamiltonian was at a critical point we expect this new Hamiltonian to be the same Hamiltonian, otherwise it will be some new *renormalized* Hamiltonian.

If we were to perform renormalization in Fourier space this would be straightforward, by just 'squeezing' each Fourier mode separately. However, again we are interested in a procedure that is local in real space, and we will use a wavelet transform.

Consider a biorthogonal pair of wavelet filters (g_s, g_w) and (h_s, h_w) as in Section 3.2, with the corresponding discrete wavelet transforms W_g and W_h. Since the filters which implement W_g and W_h are real, it is possible to consider the wavelet decomposition maps as $W_g, W_h : \ell^2(\mathbb{Z}, \mathbb{R}) \to \ell^2(\mathbb{Z}, \mathbb{R})$, and we will do so throughout this chapter. Then, as $W_g^{-1} = W_h^\mathsf{T}$ the map $W = W_g \oplus W_h$ defines a symplectic map on $\ell^2(\mathbb{Z})$ for any pair of biorthogonal wavelet filters. Reinterpreting this map as mapping the scaling and wavelet output to the even and odd sublattice (i.e. by considering the map $W_g' \oplus W_h'$ as in Section 3.5) we see that its second quantization will have the structure of a layer of entanglement renormalization, and it has the right interpretation as it 'splits' the high and low frequency modes. However, just as in the fermionic case, we need to choose the filters such that W actually disentangles the state, and the wavelet output is unentangled. We consider a general Hamiltonian as in Eq. (5.2) with dispersion relation E. We normalize the dispersion relation such that $E(\pi) = 1$, which can be implemented by the symplectic (squeezing) map $(\sqrt{E(\pi)}I) \oplus (1/\sqrt{E(\pi)}I)$ (note that this map can be seen as a product of maps acting on single sites). Then the condition for the wavelet output to be disentangled is that the Fourier transforms of the filters satisfy

$$\hat{g}_w(\theta) = E(\theta)\hat{h}_w(\theta). \tag{5.4}$$

Intuitively, what happens is that W separates the bosonic modes in high frequency and low frequency modes, and Eq. (5.4) makes sure that the high frequency modes are not entangled to the low frequency modes in the ground state. To derive Eq. (5.4) recall that the ground state of the Hamiltonian in Eq. (5.2) is determined by its covariance matrix $\gamma = \gamma^q \oplus \gamma^p$, which are such that

$$
\begin{aligned}
\gamma^q &= m(\frac{1}{2E}) \\
\gamma^p &= m(\frac{1}{2}E).
\end{aligned}
\tag{5.5}
$$

The covariance matrix of an unentangled (uncorrelated) product state (i.e. the Fock vacuum) is $\frac{1}{2}I$. Under a map of the form $A \oplus (A^\mathsf{T})^{-1}$ the covariance matrix transforms as

$$
\begin{aligned}
\gamma^q &\mapsto A\gamma^q A^\mathsf{T} \\
\gamma^p &\mapsto (A^\mathsf{T})^{-1}\gamma^p A^{-1}.
\end{aligned}
$$

Again, we normalize such that $E(\pi) = 1$, by the squeezing map $(\sqrt{E(\pi)}I) \oplus (1/\sqrt{E(\pi)}I)$. Suppose we have biorthogonal filters (g_s, g_w) and (h_s, h_w) satisfying Eq. (5.4), then $W = W_g \oplus W_h$ disentangles the ground state. To see that this is indeed true, we compute the result of applying the wavelet decomposition map to the ground state covariance matrix $\gamma = \gamma^q \oplus \gamma^p$ given in terms of the dispersion relation by Eq. (5.5). For this, we remark that from $\hat{g}_w(\theta) = E(\theta)\hat{h}_w(\theta)$ it follows that $\hat{h}_s(\theta) = E(\theta + \pi)\hat{g}_s(\theta)$. Then,

$$E(\theta)\hat{h}_w(\theta)\hat{f}(2\theta) = \hat{g}_w(\theta)\hat{f}(2\theta),$$
$$E(\theta)\hat{h}_s(\theta)f(2\theta) = \hat{g}_s(\theta)E^{(1)}(2k)f(2\theta),$$

where $E^{(1)}$ is the renormalized dispersion relation on the scaling output defined By

$$E^{(1)}(\theta) = E\left(\tfrac{\theta}{2}\right)E\left(\tfrac{\theta}{2} + \pi\right), \tag{5.6}$$

This shows that $m(E)W_h^{\mathsf{T}} = W_g^{\mathsf{T}} \circ (m(E^{(1)}) \oplus I)$ and hence

$$W_h\gamma^p W_h^{\mathsf{T}} = W_h W_g^{\mathsf{T}}(\gamma^{p,(1)} \oplus \tfrac{1}{2}I) = \gamma^{p,(1)} \oplus \tfrac{1}{2}I$$
$$\gamma^{p,(1)} = m(\tfrac{1}{2}E^{(1)}).$$

Similarly, it holds that

$$W_g\gamma^q W_g^{\mathsf{T}} = \gamma^{q,(1)} \oplus \tfrac{1}{2}I$$
$$\gamma^{q,(1)} = m(\tfrac{1}{2E^{(1)}}).$$

We thus see that W has unentangled the high-frequency modes to a product state, and the low frequency modes are renormalized to a Gaussian state with a new dispersion relation $E^{(1)}$ given by Eq. (5.6).

We can now recursively apply the same construction to the scaling output, as in Fig. 2.1, now with the renormalized dispersion relation. To introduce some notation, we let $E^{(l)}$ be the dispersion relation after l layers of renormalization, recursively defined by (cf. Eq. (5.6), note that we first normalize the dispersion relation by a factor $E^{(l)}(\pi)$)

$$E^{(l+1)}(\theta) = \frac{E^{(l)}(\tfrac{\theta}{2})}{E^{(l)}(\pi)}\frac{E^{(l)}(\tfrac{\theta}{2} + \pi)}{E^{(l)}(\pi)}. \tag{5.7}$$

The normalization by $E^{(l)}(\pi)$ could also be absorbed in the filters, but we would like the filters to be such that $\hat{g}_s(0) = \hat{h}_s(0) = \sqrt{2}$, as is standard in the signal processing literature and convenient for the analysis. Then at the l-th layer we need filters $g^{(l)}, h^{(l)}$ satisfying $\hat{g}_w^{(l)}(\theta) = \frac{E^{(l)}(\theta)}{E^{(l)}(\pi)}\hat{h}_w^{(l)}(\theta)$ (cf. Eq. (5.4)), and we let

$$R_{g^{(l)}} = W_{g^{(l)}}\sqrt{E^{(l)}(\pi)},$$
$$R_{h^{(l)}} = W_{h^{(l)}}\frac{1}{\sqrt{E^{(l)}(\pi)}}. \tag{5.8}$$

Finally, we define the $\mathcal{L}$-layer renormalization map as $R^{(\mathcal{L})} = R_g^{(\mathcal{L})} \oplus R_h^{(\mathcal{L})}$, where

$$R_a^{(\mathcal{L})} = (R_{a^{(\mathcal{L}-1)}} \oplus I^{\oplus(\mathcal{L}-1)}) \circ \ldots \circ (R_{a^{(1)}} \oplus I) \circ R_{a^{(0)}}$$

for $a = g, h$. Then, $R^{(\mathscr{L})}$ maps the state with dispersion relation ω to a product state with covariance matrix $\frac{1}{2}I$ on the $\mathscr{L}$ high frequency levels, and a state with dispersion relation $E^{(\mathscr{L})}$ on the remaining low frequency level. While we motivated the procedure from the perspective of disentangling a given entangled state, the resulting transformation can also be used in the opposite direction, to prepare the ground state by applying the circuit to a product state, thus realizing the state as a bosonic MERA state.

Finally we note that if the wavelet filters have compact support (i.e. they are FIR filters) of size $2N$, then by Section 3.5 W (reinterpreted as W') gives rise to a Gaussian circuit of depth N that maps the low-frequency modes to the odd sublattice and the high-frequency modes to the even sublattice as shown in Fig. 5.1. We conclude that this is exactly the structure of an entanglement renormalization circuit. The converse to this construction is also true: any Gaussian entanglement renormalization circuit which is of the form $A \oplus (A^\mathsf{T})^{-1}$ gives rise to a biorthogonal pair of wavelets, as follows from the discussion in Section 3.5.

In general, Eq. (5.4) can not be satisfied exactly for compactly supported wavelet filters (in particular if $E(\theta)$ is not a ratio of trigonometric polynomials). In this case we may still *approximate* this relation, just as for the approximate Hilbert pairs in the fermionic case. Suppose we are given a family of filter pairs $(g_s^{(l)}, g_w^{(l)})$ and $(h_s^{(l)}, h_w^{(l)})$ for $l = 1, \ldots, \mathscr{L}$, where the l-th pair represents the l-th layer such that

$$\left| \hat{g}_w^{(l)}(\theta) - \frac{E^{(l)}(\theta)}{E^{(l)}(\pi)} \hat{h}_w^{(l)}(\theta) \right| \le \varepsilon \qquad \forall l = 1, \ldots, \mathscr{L}, \tag{5.9}$$

so they approximately reproduce the dispersion relation at each layer (up to normalization). Then we define

Definition 5.1 (Approximate covariance matrix lattice boson)**.** For any family of pairs of biorthogonal wavelet filters $(g_s^{(l)}, g_w^{(l)})$ and $(h_s^{(l)}, h_w^{(l)})$, for $l = 1, \ldots, \mathscr{L}$, define the *approximate covariance matrix* $\gamma_{\text{MERA}}^{(\mathscr{L})}$ as the following operator on $\ell^2(\mathbb{Z})$:

$$\gamma_{\text{MERA}}^{(\mathscr{L})} = \gamma_{\text{MERA}}^{q,(\mathscr{L})} \oplus \gamma_{\text{MERA}}^{p,(\mathscr{L})} \tag{5.10}$$

where

$$\gamma_{\text{MERA}}^{q,(\mathscr{L})} := \frac{1}{2} R_h^{(\mathscr{L}),\mathsf{T}} R_h^{(\mathscr{L})},$$

$$\gamma_{\text{MERA}}^{p,(\mathscr{L})} := \frac{1}{2} R_g^{(\mathscr{L}),\mathsf{T}} R_g^{(\mathscr{L})},$$

and $R_h^{(\mathscr{L})}$ and $R_g^{(\mathscr{L})}$ are defined as in Eq. (5.8).

This definition raises two interesting questions:

(i) Do there exist filters which approximately satisfy Eq. (5.9)? We have provided a partial answer to this question by giving a general heuristic procedure for constructing such filters in Section 3.4.

(ii) Secondly, one can wonder whether a good approximation of the dispersion relation at the level of a single layer as in Eq. (5.9) will indeed give rise to a good approximation of the ground state. We answer this question in the affirmative in Section 5.3, given certain conditions on the wavelet filters.
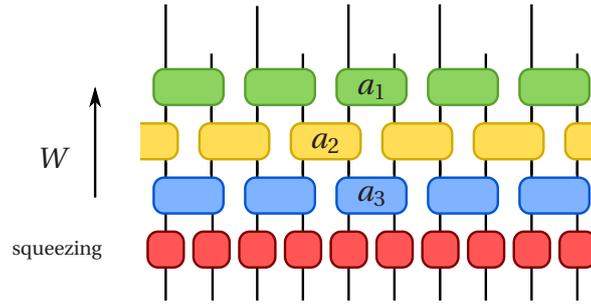
Figure 5.1: Decomposition of a single layer of the entanglement renormalization map $R^{(1)}$ as a circuit. The wavelet transform $W = W_g \oplus W_h$ is decomposed as a circuit with two-local gates a_i, and follows the bottom layer which squeezes by $E(\pi)^{\frac{1}{2}} \oplus E(\pi)^{-\frac{1}{2}}$ to normalize the dispersion relation. This figure can be interpreted both as a linear circuit implementing a symplectic transformation, and as its second quantization, which is a bosonic Gaussian circuit.

Finally, we make the observation that if we are only interested in an approximation, and the theory flows to either a critical theory or a trivial theory under renormalization, then we only need a small number of 'transition layers' and can pick fixed filters $(g_s^{(l)}, g_w^{(l)})$ and $(h_s^{(l)}, h_w^{(l)})$ for sufficiently large l.

Entanglement renormalization for the harmonic chain

For the massless harmonic chain, Eq. (5.6) amounts to $E(\theta) \mapsto |\sin(\frac{\theta}{4})\cos(\frac{\theta}{4})| = \frac{1}{2}\sin(\frac{\theta}{2})$, so the dispersion relation is invariant under the renormalization step if we include the subsequent normalization. Hence the state on the scaling output of the entanglement renormalization will be the same after any number of layers. This implies that we can keep iterating the same entanglement renormalization layer with identical filters at each layer, giving a *scale-invariant* bosonic entanglement renormalization procedure for the massless harmonic chain.

In the massive case, the mass renormalizes as

$$M \mapsto 2\sqrt{M^2 + M^4}. \tag{5.11}$$

This is a *relevant* perturbation to the massless chain [EV10a], and with increasing number of layers the dispersion relation becomes flat; correspondingly we can let the filters at the deeper layers approach orthogonal wavelet filters.

5.2.2   Entanglement renormalization for the free bosonic field

In Chapter 4 we saw that the continuous wavelet basis corresponding to an approximate Hilbert pair had a natural interpretation as an entanglement renormalization scheme for the Dirac fermionic field theory. We recall that the key point was that the scaling functions are a natural UV cut-off that is compatible with the entanglement renormalization circuits. In this section we will argue that there is a similar natural connection in the bosonic case.

Entanglement renormalization for the massless bosonic field

We first suppose that the biorthogonal wavelet filters are *exactly* related by the dispersion relation of the massless harmonic chain, that is,

$$\hat{g}_w(\theta) = \left|\sin\left(\frac{\theta}{2}\right)\right| \hat{h}_w(\theta). \tag{5.12}$$

We claim that, in this case, the scaling functions defined in Eq. (3.16) are related as

$$\hat{\phi}^g(\omega) = \frac{|\omega|}{2|\sin(\frac{\omega}{2})|} \hat{\phi}^h(\omega). \tag{5.13}$$

To verify this claim, we note that as a consequence of Eq. (5.12) and the relation in Eq. (3.15) we have $\hat{h}_s(\theta) = |\cos(\frac{\theta}{2})|\hat{g}_s(\theta)$. Next, from Eq. (3.16) it follows that the scaling functions are related as $\hat{\phi}^h(\omega) = \gamma(\omega)\hat{\phi}^g(\omega)$, where

$$\gamma(\omega) = \prod_{n=1}^{\infty} |\cos(2^{-n-1}\omega)|.$$

This expression implies that $\gamma(\omega)$ has to satisfy $\gamma(\omega) = |\cos(\frac{\omega}{4})|\gamma(\frac{\omega}{2})$, and we can easily verify that $\gamma(\omega) = \frac{2|\sin(\frac{\omega}{2})|}{|\omega|}$, which has the right normalization $\gamma(0) = 1$. This proves Eq. (5.13), which in turn, using Eqs. (3.17) and (5.12), also implies that

$$
\begin{aligned}
\hat{\psi}^g(\omega) &= \frac{1}{\sqrt{2}} \hat{g}_w\left(\frac{\omega}{2}\right) \hat{\phi}^g\left(\frac{\omega}{2}\right) \\
&= \frac{1}{\sqrt{2}} \left|\sin\left(\frac{\omega}{4}\right)\right| \hat{h}_w\left(\frac{\omega}{2}\right) \frac{|\omega|}{4|\sin(\frac{\omega}{4})|} \hat{\phi}^h\left(\frac{\omega}{2}\right) \\
&= \frac{|\omega|}{4} \hat{\psi}^h(\omega).
\end{aligned}
\tag{5.14}
$$

Equation (5.14) shows that the wavelet functions are related precisely by the linear dispersion relation of the massless free boson. This shows that if we take a discrete (biorthogonal) wavelet transform, with dispersion relation for a bosonic lattice model (in this case the massless harmonic chain), then the associated continuous wavelet transform is closely related to the continuum limit of the lattice model (in this case the massless free boson).

We now assume that we have filters which (approximately) satisfy Eq. (5.12). We would like to compute correlation functions of smeared fields $\Phi(f)$ and $\Pi(f)$ for real valued single particle functions f. Analogous to the fermionic case we will do the following, in order to compute some correlation function:

(i) We discretize the operators using the scaling functions at some appropriate scale, or equivalently, we embed the discrete theory into the continuous one using the scaling functions.

(ii) We compute the correlation function of the discretized operators using the state prepared by the (discrete) entanglement renormalization circuit.

Let us denote by $\alpha_j^h, \alpha_j^g : L^2(\mathbb{R}, \mathbb{R}) \to \ell^2(\mathbb{Z}, \mathbb{R})$ the discretization maps $(\alpha_j^h f)[k] = \langle \psi_{j,k}^h, f \rangle$ and $(\alpha_j^g f)[k] = \langle \psi_{j,k}^g, f \rangle$. Then in the above procedure the discretization procedure is given by

$$\Phi(f) \mapsto \sum_n \alpha_j^h(f)[n] q_n = \sum_n \langle \phi_{l,n}^h, f \rangle q_n,$$

$$\Pi(f) \mapsto \sum_n \alpha_j^g(f)[n] p_n = \sum_n \langle \phi_{l,n}^g, f \rangle p_n. \tag{5.15}$$

(the inner product here is the usual $L^2(\mathbb{R})$ inner product). Using $\mathscr{L}$ layers of entanglement renormalization gives the following approximate covariance:

Definition 5.2 (Approximate covariance massless bosonic field)**.** For any pair of filters (g_s, g_w) and (h_s, h_w) which approximately satisfy Eq. (5.12), $j \in \mathbb{Z}$, and $\mathscr{L} \in \mathbb{N}$, define the *approximate covariance matrix* by $\gamma_{\mathrm{MERA}}^{(\mathscr{L})}$ as the following bilinear form:

$$\gamma_{\mathrm{MERA}}^{(j,\mathscr{L})} = \gamma_{\mathrm{MERA}}^{q,(j,\mathscr{L})} \oplus \gamma_{\mathrm{MERA}}^{p,(j,\mathscr{L})} \tag{5.16}$$

where

$$\gamma_{\mathrm{MERA}}^{q,(j,\mathscr{L})}(f_1, f_2) := \frac{1}{2} \langle R_h^{(\mathscr{L})} \alpha_j^h f_1, R_h^{(\mathscr{L})} \alpha_j^h f_2 \rangle,$$

$$\gamma_{\mathrm{MERA}}^{p,(j,\mathscr{L})}(f_1, f_2) := \frac{1}{2} \langle R_g^{(\mathscr{L})} \alpha_j^g f_1, R_g^{(\mathscr{L})} \alpha_j^g f_2 \rangle$$

for any real-valued functions f_1, f_2.

To confirm that this a reasonable approximation, consider the case where we have smeared fields $\Phi(f)$ with f of the form $f = \sum_n s[n] \phi_{l,n}^g$ and $\Pi(\tilde{f})$ with $\tilde{f}$ of the form

$$\tilde{f} = \sum_n \tilde{s}[n] \phi_{l,n}^h,$$

and with Eq. (5.12) satisfied exactly. Because the wavelet functions are precisely related by the correct dispersion relation, in order to compute correlation functions, it suffices to express the functions f and $\tilde{f}$ in the wavelet bases $\{\psi_{l',n'}^h\}$ and $\{\psi_{l',n'}^g\}$. To see this it suffices to look at two-point functions, and suppose that we want to compute the correlation $\langle \Pi(f_1) \Pi(f_2) \rangle$, where $f_i = \sum_n s_i[n] \phi_{l,n}^h$. We rewrite $f_i = \sum_{l,n} w_i[l,n] \psi_{l,n}^h$ and we note that $m(\frac{1}{2E}) \psi^h = \psi^g$ and hence $2^l m(\frac{1}{2E}) \psi_{l,n}^h = \psi_{l,n}^g$, so

$$\sum_{l,n} 2^{-l} w_1[l,n] w_2[l,n] = \langle \sum_{l,n} w_1[l,n] \psi_{l,n}^h, \sum_{l',n'} 2^{-l'} w_2[l',n'] \psi_{l',n'}^g \rangle$$

$$= \langle \sum_{l,n} w_1[l,n] \psi_{l,n}^h, \sum_{l',n'} w_2[l',n'] m(\frac{1}{2E}) \psi_{l',n'}^h \rangle$$

$$= \langle f_1, m(\frac{1}{2E}) f_2 \rangle$$

which is indeed the correct correlation function. A similar computation holds for correlation functions involving the field Φ. By Eq. (3.18) the $w_i[l,n]$ are computed from $s_i[n]$ precisely by applying the discrete wavelet transform, and the factor of 2^{-l} derives from our normalization of the dispersion relation (the 'squeezing layer' in Fig. 5.1). In other words, the correlation functions will be given precisely by applying the entanglement renormalization circuit to the operators $\sum_n s[n] q_n$ and $\sum_n \tilde{s}[n] p_n$. In this chapter we do not prove that Eq. (5.16) provides an accurate approximation, but it should be relatively straightforward to adapt the arguments in Theorem 4.7 and Theorem 5.3.

Scaling dimensions

As an application, we can consider the entanglement renormalization superoperator, which coarse-grains operators by conjugating with a single layer of the renormalization circuit. Recall that for critical lattice models, entanglement renormalization superoperator has been proposed to approximately encode the conformal data of the continuum limit of the theory [EV13] and Fig. 2.3. We will now verify that, similar to the fermionic case, the entanglement renormalization superoperator reproduces *exactly* the scaling dimensions of the Φ and Π fields in the massless case, as well as the scaling dimension of a number of descendants (equal to the number of vanishing moments of the wavelet filters), similar as for the fermionic wavelet MERA as we saw in Chapter 4. In this case we consider the operators $\Phi_{\mathrm{MERA}}(x) = \sum_n \phi^h(x-n)q_n$ and $\Pi_{\mathrm{MERA}}(x) = \sum_n \phi^g(x-n)p_n$ for any $x \in \mathbb{R}$, which are the discretizations of the operators $\Phi(x)$ and $\Pi(x)$. It can be easily seen that the entanglement renormalization superoperator maps these operators

$$\Phi_{\mathrm{MERA}}(x) \mapsto \sum_{n,l} \sqrt{2}h_s[l]\phi^h(x-2n-l)q_n$$

$$= \sum_n \phi^h\left(\frac{x}{2}-n\right)q_n = \Phi_{\mathrm{MERA}}\left(\frac{x}{2}\right)$$

where we use that for the scaling function Eq. (3.16) it holds that [Mal08]

$$\frac{1}{\sqrt{2}}\phi^h(\frac{x}{2}) = \sum_n h_s[n]\phi^h(x-n). \tag{5.17}$$

Similarly one finds $\Pi_{\mathrm{MERA}}(x) \mapsto \frac{1}{2}\Pi_{\mathrm{MERA}}(\frac{x}{2})$. This corresponds, as expected, to scaling dimensions 0 and 1. As in Section 4.2.5 if the wavelets have K vanishing moments, the superoperator reproduces K descendents exactly.

The massive bosonic field

Since in the bosonic case the construction is not limited to a critical model, we may approach the free massive boson with mass M in a similar manner. In this case, we suppose we have two families of filters $(g_s^{(l)}, g_w^{(l)})$ and $(h_s^{(l)}, h_w^{(l)})$, now with $l \in \mathbb{Z}$ and such that $\sqrt{(M^{(l)})^2 + 1}\,\hat{g}_w^{(l)}(\theta) \approx \sqrt{(M^{(l)})^2 + \sin^2(\frac{\theta}{2})}\,\hat{h}_w^{(l)}(\theta)$ where $M^{(0)} = m$ and $M^{(l)}$ is the mass after l layers of renormalization, as defined by Eq. (5.11). If these filters are chosen in a way that they converge to a fixed *orthonormal* filter as l goes to infinity, and to a fixed pair of biorthogonal filters as in the massless case for l to $-\infty$, it makes sense to define a new type of scaling and wavelet functions which are different at each level l as a generalization of of the scaling and wavelet functions:

$$\hat{\phi}_l^a(\omega) = \prod_{j=1}^\infty \frac{\hat{a}^{(l+j)}(2^{-j}\omega)}{\sqrt{2}}$$

$$\hat{\psi}_l^a(\omega) = \frac{1}{\sqrt{2}}\hat{a}^{(l+1)}\left(\frac{\omega}{2}\right)\hat{\phi}_{l+1}^a\left(\frac{\omega}{2}\right) \tag{5.18}$$

for $a = g, h$ as a generalization of of the scaling and wavelet functions defined in Eq. (3.16) and Eq. (3.17). Again, the wavelet functions $\psi_{l,n}^a(x) = 2^{-\frac{l}{2}}\psi_l(2^{-l}x - n)$ for filters $a = g, h$ form a dual basis (provided they exist). The behaviour for $l \to \pm\infty$ is

consistent with the fact that the mass term is a relevant perturbation of the conformal field theory and the theory flows from a critical massless boson to a trivial theory. As before, we can now discretize the theory using the scaling functions at some given scale and use the discrete circuit to compute correlation functions.

5.2.3   Numerical examples

In Section 3.4 we have provided a construction of filters approximately satisfying Eq. (5.4) for the massless harmonic chain. This construction depends on two integer parameters K and L, where K controls the number of vanishing moments of the filters and L controls the accuracy of the approximation of the dispersion relation. This corresponds to a circuit depth of $N = K + 2L$ for a single layer. In Fig. 5.2 we illustrate our approximation result by numerically computing correlation functions of the massless harmonic chain using these filters [WW21b] (using that we can easily simulate the circuits in the single-particle picture).

5.3   Approximation of correlation functions

In this section we formulate and prove a general approximation result for translation-invariant quadratic Hamiltonians of the form in Eq. (5.2). Our proof strategy is inspired by the techniques in [HSW+18], with the technical complications that the wavelet transforms are not unitary and are allowed to vary layer by layer.

Recall that if $g_s, h_s \in \ell^2(\mathbb{Z}, \mathbb{R})$ are a pair of scaling filters that satisfy the perfect reconstruction condition in Eq. (3.14), then we can define corresponding wavelet filters $g_w, h_w \in \ell^2(\mathbb{Z})$ and single-layer decomposition maps $W_g, W_h : \ell^2(\mathbb{Z}, \mathbb{R}) \to \ell^2(\mathbb{Z}, \mathbb{R})$ such that $W_h^T W_g = W_g^T W_h = I$.

Now suppose that we are given a sequence of filters $g_s^{(l)}, h_s^{(l)}$ as above. Here, $l \in \mathbb{N}$ for convenience of notation. In practice, one is usually interested in a finite number of layers; in this case we may choose the sequence of filters to eventually become constant. For $a = g, h$ and $\mathscr{L} \in \mathbb{N}$, we define the $\mathscr{L}$-layer decomposition maps

$$W_a^{(\mathscr{L})} : \ell^2(\mathbb{Z}, \mathbb{R}) \to \ell^2(\mathbb{Z}, \mathbb{R})^{\otimes(1+\mathscr{L})},$$
$$W_a^{(\mathscr{L})} := \left(W_{a^{(\mathscr{L}-1)}} \oplus I^{\oplus(\mathscr{L}-1)}\right) \circ \ldots \circ \left(W_{a^{(1)}} \oplus I\right) \circ W_{a^{(0)}},$$

and write $W^{(\mathscr{L})} = W_h^{(\mathscr{L})} \oplus W_g^{(\mathscr{L})}$. We assume that the family is *stable* in the sense that the corresponding (generalized) scaling functions ϕ_l^a defined in Eq. (5.18) exist, are square integrable, and bounded in L^∞-norm. We can also define the wavelet decomposition maps starting at layer $\mathscr{L}' \geq 0$, that is,

$$W_a^{(\mathscr{L}', \mathscr{L})} : \ell^2(\mathbb{Z}, \mathbb{R}) \to \ell^2(\mathbb{Z}, \mathbb{R})^{\otimes(1+\mathscr{L}-\mathscr{L}')},$$
$$W_a^{(\mathscr{L}', \mathscr{L})} := \left(W_{a^{(\mathscr{L})}} \oplus I^{\oplus(\mathscr{L}-\mathscr{L}'-1)}\right) \circ \ldots \circ \left(W_{a^{(\mathscr{L}'+1)}} \oplus I\right) \circ W_{a^{(\mathscr{L}')}}.$$

For $\mathscr{L}' = 0$ we recover $W_a^{(\mathscr{L})}$ as defined earlier. We assume that the wavelet decomposition maps are bounded. Finally, we shall assume that the filters have finite support. Then the same is true for the scaling functions. In the case that the filters are independent of l, the above notion of stability is equivalent to the familiar notion from wavelet
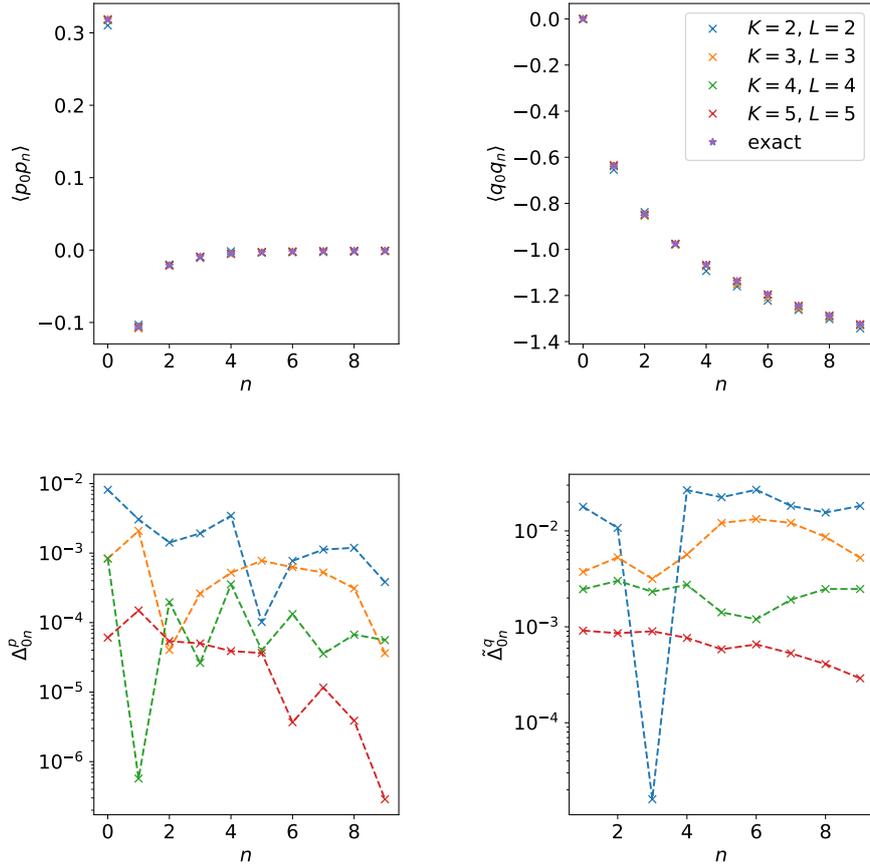
Figure 5.2: Approximation of correlation functions for the massless harmonic chain by the MERA. We used the filter construction of Section 3.4 and $\mathscr{L} = 20$ layers of renormalization. The former depends on parameters K and L which are explained in the main text. We show the correlation functions $\langle p_0 p_n \rangle$ and $\langle p_0 p_n \rangle$, as well as their approximation errors $\Delta_{0n}^p := |\langle p_0 p_n \rangle_{\text{exact}} - \langle p_0 p_n \rangle_{\text{MERA}}|$ and $\tilde{\Delta}_{0n}^q := |\langle q_0 q_n \rangle_{\text{exact}} - \langle q_0 q_n \rangle_{\text{MERA}}|$.

theory. For finitely supported filters there exists an easy criterion to determine this, see [CDF92].

Recall that for the entanglement renormalization circuit we also insert a squeezing operation between each wavelet decomposition layer, defining $R_{a^{(l)}}$, $R_a^{(\mathcal{L})}$ and $R^{\mathcal{L}}$ for $a = g, h$ as in Eq. (5.8). Our approximation to the covariance matrix is then given by

$$
\begin{aligned}
\gamma_{\mathrm{MERA}}^{q,(\mathcal{L})} &= \frac{1}{2} R_h^{(\mathcal{L}),\mathsf{T}} R_h^{(\mathcal{L})}, \\
\gamma_{\mathrm{MERA}}^{p,(\mathcal{L})} &= \frac{1}{2} R_g^{(\mathcal{L}),\mathsf{T}} R_g^{(\mathcal{L})}.
\end{aligned}
\tag{5.19}
$$

Suppose the filter pairs $g^{(l)}$, $h^{(l)}$ form an ε-approximate $E^{(l)}$-dispersion pair at each level, that is

$$
|\hat{g}_w^{(l)}(\theta) - \frac{E^{(l)}(\theta)}{E^{(l)}(\pi)} \hat{h}_w^{(l)}(\theta)| = |\hat{g}_w^{(l)}(\theta) - \hat{\tilde{g}}_w^{(l)}(\theta)| \le \varepsilon,
\tag{5.20}
$$

where we have introduced the filter

$$
\hat{\tilde{g}}_w^{(l)}(\theta) := \frac{E^{(l)}(\theta)}{E^{(l)}(\pi)} \hat{h}_w^{(l)}(\theta).
\tag{5.21}
$$

This filter, together with $\hat{\tilde{h}}_w^{(l)}(\theta) := E^{(l)}(\pi)/E^{(l)}(\theta) \times \hat{g}_w^{(l)}(\theta)$, forms a pair of biorthogonal wavelet filters, with corresponding scaling filters $\tilde{g}_s^{(l)}$, $\tilde{h}_s^{(l)}$ that satisfy Eq. (3.14). However, these filters are almost never finitely supported. By construction, the pair $\tilde{g}^{(l)}$, $h^{(l)}$ satisfies the dispersion relation Eq. (5.4) exactly.

We now state an approximation theorem for general dispersion relations, which is the main result of this chapter. We measure the approximation error in terms of quantities

$$
\begin{aligned}
\Delta_{nm}^p &:= |\gamma_{nm}^p - (\gamma_{\mathrm{MERA}}^{p,(\mathcal{L})})_{nm}|, \\
\Delta_{nm}^q &:= |\gamma_{nm}^q - (\gamma_{\mathrm{MERA}}^{q,(\mathcal{L})})_{nm}|.
\end{aligned}
\tag{5.22}
$$

If $E(0) = 0$, then it is also interesting we have to regulate the covariance matrix γ^q. Since $E(0) = 0$, the value of γ_{nm}^q is not unambiguously defined, and if one computes $\langle f_1, \gamma f_2 \rangle$ one has to restrict at least one of the test functions to a subspace of functions f_i for which $\hat{f}_i(0) = 0$ (so f_i sums to zero). With this in mind we define

$$
\tilde{\gamma}_{nm}^q := \gamma_{nm}^q - \gamma_{nn}^q
\tag{5.23}
$$

which equals $\tilde{\gamma}_{nm}^q = \langle \delta_n, \gamma^q(\delta_m - \delta_n) \rangle$ and consider

$$
\tilde{\Delta}_{nm}^q := |\tilde{\gamma}_{nm}^q - (\tilde{\gamma}_{\mathrm{MERA}}^{q,(\mathcal{L})})_{nm}|.
\tag{5.24}
$$

Theorem 5.3. *Consider a translation-invariant Hamiltonian of the form of Eq. (5.2), with dispersion relation $E(\theta)$ such that $E^{(l)}(\pi) \le 1$ and $E^{(l)}(\theta) \le \Omega$ for $l = 1, \ldots, \mathcal{L}$ for $\Omega \ge 1$. Suppose we have a sequence of filters such that Eq. (5.20) holds for $\varepsilon \le 1$, with finite support of size at most N and scaling functions that are uniformly bounded by $\|\phi_l^a\|_\infty \le B$ for $a = g, h$ and $l = 1, \ldots, \mathcal{L}$. Assume moreover that the wavelet decomposition maps are uniformly bounded by $\|W_a^{(l',l)}\| \le \kappa$ for all $a = g, h, \tilde{g}$ and $1 \le l \le l' \le \mathcal{L}$, where $D \ge 1$.*

Then the approximation error of the covariance matrices can be bounded as follows:

$$\Delta_{nm}^p \le \kappa^2 \left(C 2^{-\frac{\mathscr{L}}{2}} + 3\varepsilon\kappa \log_2 \frac{C}{\varepsilon} \right),$$

$$\Delta_{nm}^q \le 2\kappa^2 \left(C 2^{-\frac{\mathscr{L}}{2}} + 3\varepsilon\kappa \log_2 \frac{C}{\varepsilon} \right) \|\gamma^q \delta_0\|,$$

$$\tilde{\Delta}_{nm}^q \le 2\kappa^2 \left(C 2^{-\frac{\mathscr{L}}{2}} + 3\varepsilon\kappa \log_2 \frac{C}{\varepsilon} \right) \|\gamma^q (\delta_n - \delta_m)\|,$$

where $C := 4B^2 N^{\frac{3}{2}} \Omega$.

To interpret the error bounds, we note that

$$\|\gamma^q \delta_0\|^2 = \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{E(\theta)^2}, \tag{5.25}$$

$$\|\gamma^q (\delta_n - \delta_m)\|^2 = \int_{-\pi}^{\pi} \frac{\sin^2(\frac{1}{2}(n-m)k)}{E(\theta)^2} \mathrm{d}\theta. \tag{5.26}$$

We will first bound the error that arises from only taking a finite number of layers. Let $p_s^{(\mathscr{L})}$ denote the projection onto the first tensor factor of $\ell^2(\mathbb{Z}, \mathbb{R})^{\otimes(\mathscr{L}+1)}$ and $p_w^{(\mathscr{L})} = I - p_s^{(\mathscr{L})}$ the projection onto the remaining tensor factors. Thus, $p_s^{(\mathscr{L})} W_a^{(\mathscr{L})} f$ is the scaling component of the decomposed signal and $p_w^{(\mathscr{L})} W_a^{(\mathscr{L})}$ its wavelet component. The following lemma is a straightforward adaptation of Lemma 3.5 and Lemma 1 in [HSW+18], which confirms the intuition that for finitely supported signals, lower-frequency wavelet modes contribute less.

Lemma 5.4. *Suppose we have sequence of filters as above, with finite support of size at most N and scaling functions that are uniformly bounded by $\|\phi_l^a\|_\infty \le B$ for $a = g, h$ and $l = 1, \dots, \mathscr{L}$. Then,*

$$\|p_s^{(\mathscr{L})} W_a^{(\mathscr{L})} \delta_n\| \le 2^{-\frac{\mathscr{L}-1}{2}} B^2 N^{\frac{3}{2}} \tag{5.27}$$

where δ_n *is the unit signal concentrated at* n.

Proof. Let b denote the filters dual to a (i.e., $b = h$ if $a = g$, and vice versa). We note that $p_s^{(\mathscr{L})} W_a^{(\mathscr{L})} \delta_n[m] = \langle \phi_{0,n}^b, \phi_{\mathscr{L},m}^a \rangle$, where $\phi_{\mathscr{L},m}^a(x) := 2^{-\mathscr{L}/2} \phi_{\mathscr{L}}^a(2^{-\mathscr{L}} x - m)$ are the translated and shifted scaling functions. This follows from the fact that

$$\langle \phi_{0,n}^b, \phi_{0,m}^a \rangle = \delta_{nm}$$

and by applying inductively the fact that by definition of the scaling functions $\phi_{l+1,m}^a = \sum_n a_s^{(l+1)}(2m - n)\phi_{l,n}^a$. Now we can proceed as in the proof of Lemma 3.5 and estimate

$$\|p_s^{(\mathscr{L})} W_a^{(\mathscr{L})} \delta_n\|^2 = \sum_m \left| \int_{-\infty}^{\infty} \mathrm{d}x \, \phi_0^b(x-n) 2^{-\frac{\mathscr{L}}{2}} \phi_{\mathscr{L}}^a(2^{-\mathscr{L}} x - m) \right|^2$$

$$= \sum_m \left| \int_{x_0+n}^{x_0+n+N-1} \mathrm{d}x \, \phi_0^b(x-n) 2^{-\frac{\mathscr{L}}{2}} \phi_{\mathscr{L}}^a(2^{-\mathscr{L}} x - m) \right|^2$$

$$\le \sum_m \|\phi_0^b\|^2 \int_{x_0+n}^{x_0+n+N-1} \mathrm{d}x \, |2^{-\frac{\mathscr{L}}{2}} \phi_{\mathscr{L}}^a(2^{-\mathscr{L}} x - m)|^2$$

$$= 2^{-\mathscr{L}} \|\phi_0^b\|^2 \sum_m \int_{x_0+n}^{x_0+n+N-1} \mathrm{d}x \, |\phi_{\mathscr{L}}^a(2^{-\mathscr{L}} x - m)|^2$$

$$\le 2^{-\mathscr{L}+1} N^2 \|\phi_0^b\|^2 \|\phi_{\mathscr{L}}^a\|_\infty^2,$$

where in the second line we use that ϕ^b is compactly supported on $[x_0, x_0 + N - 1]$ for some x_0, in the third inequality we use Cauchy-Schwarz, and for the final inequality we use that at most $2N$ terms in the sum have nonzero overlap. Finally we may estimate $\|\phi_0^b\|^2 \le NB^2$ and $\|\phi_{\mathscr{L}}^a\|_\infty^2 \le B^2$, which yields Eq. (5.27).   ∎

The following lemma bounds the approximation error for $\mathscr{L}$ layers as a function of an intermediate layer $\mathscr{L}'$ that will later be chosen appropriately.

Lemma 5.5. *Suppose we have a sequence of filters such that Eq.* (5.20) *holds, with finite support of size at most N and scaling functions that are uniformly bounded by $\|\phi_l^a\|_\infty \le B$ for $a = g, h$ and $l = 1, \dots, \mathscr{L}$. Assume moreover that the wavelet decomposition maps are uniformly bounded by $\|W_a^{(l', l)}\| \le \kappa$ for all $a = g, h, \tilde{g}$ and $1 \le l \le l' \le \mathscr{L}$, where $D \ge 1$. Finally, let $\mathscr{L}' \in \{1, \dots, \mathscr{L}\}$. Then we have the following bounds:*

(i) For all $f \in \ell^2(\mathbb{Z}, \mathbb{R})$ and $n \in \mathbb{N}$,

$$|\langle \delta_n | \gamma^q - \gamma_{\text{MERA}}^{q, (\mathscr{L})} | f \rangle| \le 2\kappa^2 \left(\varepsilon \mathscr{L}' \kappa + 2^{-\frac{\mathscr{L}'-1}{2}} B^2 N^{\frac{3}{2}} \max\{2\|\gamma^{p, (\mathscr{L})}\|, 1\} \right) \|\gamma^q f\|. \tag{5.28}$$

(ii) Assuming $E^{(l)}(\pi) \le 1$ for all $l = 0, \dots, \mathscr{L} - 1$, we have the following bound for all $n \in \mathbb{N}$:

$$\|(\gamma^p - \gamma_{\text{MERA}}^{p, (\mathscr{L})}) \delta_n\| \le \kappa^2 \left(\varepsilon \mathscr{L}' \kappa + 2^{-\frac{\mathscr{L}'-1}{2}} B^2 N^{\frac{3}{2}} \max\{2\|\gamma^{p, (\mathscr{L}')}\|, 1\} \right). \tag{5.29}$$

Here, we recall that $\gamma^q(\theta) = m(\frac{1}{2E})$ and $\gamma^{p, (l)}(\theta) = m(\frac{1}{2} E^{(l)})$.

To interpret these bounds, we note that $\|\gamma^{p, (\mathscr{L})}\| = \max_k \frac{E^{(\mathscr{L})}(\theta)}{2}$, which is typically $\mathcal{O}(1)$. As a remark, for the critical harmonic chain we that $E^{(l)}(\pi) = \frac{1}{2}$, in which case it is not hard to see that the scaling of Eq. (5.29) can be improved to $2^{-\frac{3}{2}\mathscr{L}'}$.

Proof of Lemma 5.5. (i) To prove Eq. (5.28), we first observe that by definition of $\tilde{g}$ it holds that

$$R_{h^{(l)}} E^{(l)}(\theta) = (E^{(l+1)} \oplus I) R_{\tilde{g}^{(l)}}$$

and hence

$$R_h^{(\mathscr{L})} \gamma^p = (\gamma^{p, (\mathscr{L})} \oplus \frac{1}{2} I) R_{\tilde{g}}^{(\mathscr{L})}, \tag{5.30}$$

where $\hat{\gamma}^{p, (\mathscr{L})}(\theta) = \frac{1}{2} E^{(\mathscr{L})}(\theta)$ denotes the covariance matrix defined using the renormalized dispersion relation. We use this, together with the fact that $4\gamma^p \gamma^q = I$ on the domain of γ^q to write

$$\begin{aligned}
\gamma^q - \gamma_{\text{MERA}}^{q, (\mathscr{L})} f &= (I - \gamma_{\text{MERA}}^{q, (\mathscr{L})} 4\gamma^p) \gamma^q f \\
&= (W_h^{(\mathscr{L}'), \mathsf{T}} W_g^{(\mathscr{L}')} - R_h^{(\mathscr{L}), \mathsf{T}} R_h^{(\mathscr{L})} 2\gamma^p) \gamma^q f \\
&= (W_h^{(\mathscr{L}'), \mathsf{T}} W_g^{(\mathscr{L}')} - R_h^{(\mathscr{L}), \mathsf{T}} (2\gamma^{p, (\mathscr{L})} \oplus I) R_{\tilde{g}}^{(\mathscr{L})}) \gamma^q f \\
&= (W_h^{(\mathscr{L}'), \mathsf{T}} W_g^{(\mathscr{L}')} - W_h^{(\mathscr{L}), \mathsf{T}} (2\gamma^{p, (\mathscr{L})} \oplus I) W_{\tilde{g}}^{(\mathscr{L})}) \gamma^q f.
\end{aligned}$$

for f in the domain of γ^q. Thus,

$$
\begin{aligned}
|\langle\delta_n|\gamma^q - \gamma_{\mathrm{MERA}}^{q,(\mathscr{L})}|f\rangle| &\leq |\langle\delta_n|W_h^{(\mathscr{L}'),\mathsf{T}}p_s^{(\mathscr{L}')}W_g^{(\mathscr{L}')}|\gamma^q f\rangle| \\
&+ |\langle\delta_n|W_h^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}(W_g^{(\mathscr{L}')} - W_{\tilde{g}}^{(\mathscr{L}')})|\gamma^q f\rangle| \\
&+ |\langle\delta_n|(W_h^{(\mathscr{L}),\mathsf{T}}(2\gamma^{p,(\mathscr{L})}\oplus I)W_{\tilde{g}}^{(\mathscr{L})} - W_h^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}W_{\tilde{g}}^{(\mathscr{L}')})|\gamma^q f\rangle|.
\end{aligned}
\tag{5.31}
$$

We will bound the three terms separately, starting with the second term. By our assumption on the filters (Eq. (5.20)), $\|W_{g^{(l)}} - W_{\tilde{g}^{(l)}}\| \leq 2\varepsilon$. Hence, using a telescoping sum,

$$
\|W_g^{(\mathscr{L}')} - W_{\tilde{g}}^{(\mathscr{L}')}\| \leq \sum_{l=0}^{\mathscr{L}'-1}\|W_g^{(l+1,\mathscr{L}')}\|\,\|W_{g^{(l)}} - W_{\tilde{g}^{(l)}}\|\,\|W_{\tilde{g}}^{(l)}\| \leq 2\varepsilon\mathscr{L}'\kappa^2,
\tag{5.32}
$$

so we obtain the estimate

$$
\begin{aligned}
|\langle\delta_n|W_h^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}(W_g^{(\mathscr{L}')} - W_{\tilde{g}}^{(\mathscr{L}')})|\gamma^q f\rangle| &\leq \|W_h^{(\mathscr{L}')}\|\,\|W_g^{(\mathscr{L}')} - W_{\tilde{g}}^{(\mathscr{L}')}\|\,\|\gamma^q f\| \\
&\leq 2\varepsilon\mathscr{L}'\kappa^3\|\gamma^q f\|.
\end{aligned}
$$

The first term in Eq. (5.31) can be bounded directly using Lemma 5.4,

$$
|\langle\delta_n|W_h^{(\mathscr{L}'),\mathsf{T}}p_s^{(\mathscr{L}')}W_g^{(\mathscr{L}')}|\gamma^q f\rangle| \leq \|p_s^{(\mathscr{L}')}W_h^{(\mathscr{L}')}\delta_n\|\,\|W_g^{(\mathscr{L}')}\gamma^q f\| \leq 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\kappa\|\gamma^q f\|,
$$

and the third term may be similarly bounded as

$$
\begin{aligned}
&|\langle\delta_n|(W_h^{(\mathscr{L}),\mathsf{T}}(2\gamma^{p,(\mathscr{L})}\oplus I)W_{\tilde{g}}^{(\mathscr{L})} - W_h^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}W_{\tilde{g}}^{(\mathscr{L}')})|\gamma^q f\rangle| \\
&= |\langle\delta_n|(W_h^{(\mathscr{L}'),\mathsf{T}}(W_h^{(\mathscr{L}',\mathscr{L}),\mathsf{T}}\oplus I^{\oplus\mathscr{L}'})(2\gamma^{p,(\mathscr{L})}\oplus I^{\oplus(\mathscr{L}-\mathscr{L}')}\oplus 0^{\oplus\mathscr{L}'})W_{\tilde{g}}^{(\mathscr{L})}|\gamma^q f\rangle| \\
&= |\langle\delta_n|(W_h^{(\mathscr{L}'),\mathsf{T}}p_s^{(\mathscr{L}')}(W_h^{(\mathscr{L}',\mathscr{L}),\mathsf{T}}\oplus 0^{\oplus\mathscr{L}'})(2\gamma^{p,(\mathscr{L})}\oplus I^{\oplus\mathscr{L}})W_{\tilde{g}}^{(\mathscr{L})}|\gamma^q f\rangle| \\
&\leq \|p_s^{(\mathscr{L}')}W_h^{(\mathscr{L}')}\delta_n\|\,\|W_h^{(\mathscr{L}',\mathscr{L})}\|\,\|2\gamma^{p,(\mathscr{L})}\oplus I^{\oplus\mathscr{L}}\|\,\|W_{\tilde{g}}^{(\mathscr{L})}\|\,\|\gamma^q f\| \\
&\leq 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\kappa^2\max\{2\|\gamma^{p,(\mathscr{L})}\|,1\}\|\gamma^q f\|.
\end{aligned}
$$

By combining the three estimates we obtain Eq. (5.28).

(ii) To prove Eq. (5.29) we use Eqs. (5.19) and (5.30) to write

$$
\begin{aligned}
\gamma^p - \gamma_{\mathrm{MERA}}^{p,(\mathscr{L})} &= \gamma^p - \frac{1}{2}R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L})} = \gamma^p R_h^{(\mathscr{L}'),\mathsf{T}}R_g^{(\mathscr{L}')} - \frac{1}{2}R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L})} \\
&= R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}}(\gamma^{p,(\mathscr{L}')}\oplus\frac{1}{2}I)R_g^{(\mathscr{L}')} - \frac{1}{2}R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L})}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\|(\gamma^p - \gamma_{\mathrm{MERA}}^{p,(\mathscr{L})})\delta_n\| &\leq \|R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}}\gamma^{p,(\mathscr{L}')}p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\| \\
&+ \frac{1}{2}\|(R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}} - R_g^{(\mathscr{L}'),\mathsf{T}})p_w^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\| \\
&+ \frac{1}{2}\|(R_g^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}R_g^{(\mathscr{L}')} - R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L})})\delta_n\|
\end{aligned}
$$

As before we bound the three terms separately, starting with the second term. Since for each l we have $E^{(l)}(\pi) \le 1$, we may estimate $\|R_g^{(\mathscr{L}')}\| \le \|W_g^{(\mathscr{L}')}\| \le D$ and, by a telescoping sum as in Eq. (5.32), $\|R_{\tilde{g}}^{(\mathscr{L}')} - R_g^{(\mathscr{L}')}\| \le 2\varepsilon\mathscr{L}'\kappa^2$. Thus:

$$\frac{1}{2}\|(R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}} - R_g^{(\mathscr{L}'),\mathsf{T}})p_w^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\| \le \frac{1}{2}\|(R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}} - R_g^{(\mathscr{L}'),\mathsf{T}})\|\|R_g^{(\mathscr{L}')}\| \le \varepsilon\mathscr{L}'\kappa^3.$$

For the remaining terms, we note that $E^{(l)}(\pi) \le 1$ also implies that $\|p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\| \le \|p_s^{(\mathscr{L}')}W_g^{(\mathscr{L}')}\delta_n\| \le 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}$ using Lemma 5.4. We can thus bound the first term by

$$\|R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}}\gamma^{p,(\mathscr{L}')}p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\| \le \|R_{\tilde{g}}^{(\mathscr{L}'),\mathsf{T}}\|\|\gamma^{p,(\mathscr{L}')}\|\|p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\|$$
$$\le 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\kappa\|\gamma^{p,(\mathscr{L}')}\|,$$

and similarly the third term, where we find

$$\frac{1}{2}\|(R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L})} - R_g^{(\mathscr{L}'),\mathsf{T}}p_w^{(\mathscr{L}')}R_g^{(\mathscr{L}')})\delta_n\| = \frac{1}{2}\|R_g^{(\mathscr{L}),\mathsf{T}}R_g^{(\mathscr{L}',\mathscr{L})}p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\|$$
$$\le \frac{1}{2}\|R_g^{(\mathscr{L}),\mathsf{T}}\|\|R_g^{(\mathscr{L}',\mathscr{L})}\|\|p_s^{(\mathscr{L}')}R_g^{(\mathscr{L}')}\delta_n\|$$
$$\le \frac{1}{2}2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\kappa^2.$$

By combining the three estimates we obtain Eq. (5.29). ∎

We finally prove our general approximation theorem.

Proof of Theorem 5.3. Choosing $\mathscr{L}' = \min\{\lfloor 2\log_2\frac{C}{\varepsilon}\rfloor, \mathscr{L}\}$, we see that

$$\varepsilon\mathscr{L}'\kappa + 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\max\{2\|\gamma^{p,(\mathscr{L}')}\|, 1\} \le \varepsilon\mathscr{L}'\kappa + 2^{-\frac{\mathscr{L}'-1}{2}}B^2 N^{\frac{3}{2}}\Omega$$
$$\le 2\varepsilon\kappa\log_2\frac{C}{\varepsilon} + \max\{C2^{-\frac{\mathscr{L}}{2}}, \varepsilon\}$$
$$\le 3\varepsilon\kappa\log_2\frac{C}{\varepsilon} + C2^{-\frac{\mathscr{L}}{2}},$$

where we have used that $\frac{C}{\varepsilon} \ge 2$. Now the result follows from Eq. (5.29) and Eq. (5.28) in Lemma 5.5, choosing $f = \delta_m$ or $f = \delta_m - \delta_n$ in the latter (and using that $\|\gamma^q\delta_m\| = \|\gamma^q\delta_0\|$). ∎

Finally, let us specialize to the case of the harmonic chain with mass M (as in the informal statement of our result in Section 2.2.1. The numerical values from Theorem 5.3 using the filter construction in Section 3.4 are illustrated in Fig. 5.3. The correlation functions are related to the covariance matrices as follows:

$$\langle p_i p_j\rangle = \gamma_{ij}^p,$$
$$\langle q_i q_j\rangle = \gamma_{ij}^q,$$

the latter assuming $m > 0$. If $m = 0$ then the latter has a divergence, so we instead define

$$\langle q_i q_j\rangle := \gamma_{ij}^q - \gamma_{ii}^q = \tilde{\gamma}_{ij}^q, \tag{5.33}$$
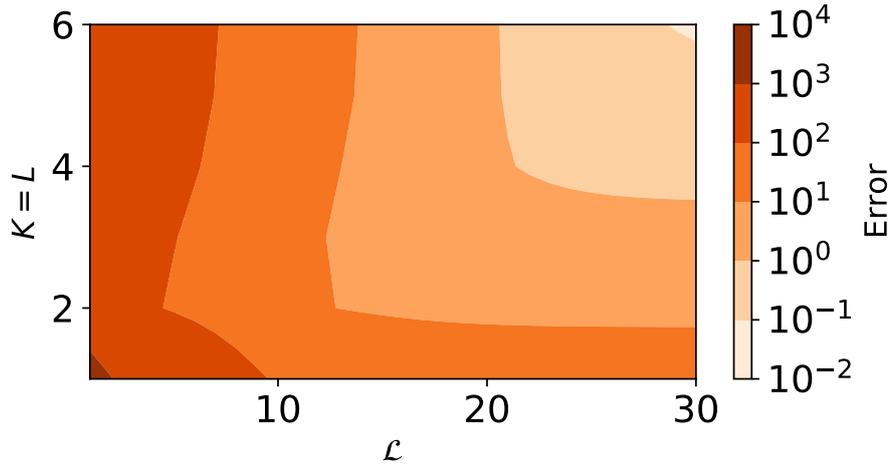
Figure 5.3: The error bound from Theorem 5.3 (Theorem 4.7) illustrated for a two-point function
$\langle p_i p_j \rangle$. It is obtained by evaluating the bound in Theorem 5.3 using Table 3.2 for an
aappropriate biorthogonal filter pair with parameters $K = L$. This is only an upper
bound to the error, see Fig. 5.2 for an example of the actual approximation error.

where $\tilde{\gamma}^q$ is the regulated covariance matrix defined in Eq. (5.23). Accordingly, we
would like to bound the quantities Δ_{ij}^p as well as Δ_{ij}^q (in the massive case) or $\tilde{\Delta}_{ij}^q$ (in
the massless case), which are defined in Eqs. (5.22) and (5.24). This is exactly achieved
by Theorem 5.3. We first normalize the dispersion relation of the harmonic chain $E(\theta)$
by a factor $\sqrt{M^2+1}$ to ω_{norm}, so that $\omega_{\text{norm}}(\pi) = 1$. There $\omega_{\text{norm}}^{(l)}(\theta) \leq 1$ and we may
apply Theorem 5.3 with $\Omega = 1$. We write γ for the original covariance matrix of the
harmonic chain and γ_{norm} for the covariance matrix where the dispersion relation has
been normalized, that is, $\gamma_{\text{norm}}^p = \frac{1}{\sqrt{M^2+1}}\gamma^p$ and $\gamma_{\text{norm}}^q = \sqrt{M^2+1}\gamma^q$. Then,

$$\frac{1}{M^2+1}\|\gamma_{\text{norm}}^q \delta_0\|^2 = \|\gamma^q \delta_0\|^2 = \int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{E(\theta)^2} = \int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{M^2+\sin^2\left(\frac{\theta}{2}\right)} \leq \frac{2\pi}{M^2},$$

so applying Theorem 5.3 using the covariance matrix γ_{norm} and restoring the fac-
tor $\sqrt{M^2+1}$ yields the results for Δ_{ij}^p and Δ_{ij}^q. In the massless case we can use Eq. (5.26)
and estimate

$$\|\gamma^q(\delta_i - \delta_j)\|^2 = \int_{-\pi}^{\pi}\frac{\sin^2\left(\frac{|i-j|\theta}{2}\right)}{\sin^2\left(\frac{\theta}{2}\right)}\mathrm{d}\theta$$

$$\leq 2\left(\int_0^{\frac{2}{|i-j|}}\frac{\pi^2|i-j|^2}{4}\mathrm{d}\theta + \int_{\frac{2}{|i-j|}}^{\pi}\frac{\pi^2}{\theta^2}\mathrm{d}\theta\right)$$

$$\leq 2\pi^2|i-j|,$$

since $|\sin(\frac{\theta}{2})| \geq \frac{|\theta|}{\pi}$ and $|\sin(\frac{n\theta}{2})| \leq \min\{\frac{n|\theta|}{2}, 1\}$ on the interval $(-\pi, \pi)$, yielding the esti-
mate for $\tilde{\Delta}_{ij}^q$.

Discussion and open questions

In this chapter we have explained how Gaussian entanglement renormalization circuits can be naturally contructed from the second quantizations of biorthogonal wavelet transforms. Moreover, as observed in Section 5.2.1 any Gaussian entanglement renormalization circuit of the form $A \oplus (A^\mathsf{T})^{-1}$ (so it does not mix the p and q modes) gives rise to a biorthogonal wavelet filter. There are a certain technical aspects that would be interesting to study in more detail. For example, if the system is not scale invariant then our notion of wavelet functions goes beyond the standard framework of wavelet theory, and one would have to identify suitable conditions on the filters that ensure that the wavelet and scaling functions as defined in Eq. (5.18) are well-behaved functions and that standard wavelet theory generalizes. Relatedly, one could extend the analysis of Chapter 4 to prove approximation bounds for correlation functions of smeared bosonic field.

The idea that wavelet theory should be a natural tool to discretize a field theory in order to perform renormalization has a long history [Bat99]. This approach differs from other works such as [BP13, BRSS15, SMMT20] which investigate the use of wavelets to discretize quantum field theories, in that we use biorthogonal wavelets, which moreover are specifically designed to target the Hamiltonian of the field theory. There is also a different approach to entanglement renormalization for quantum field theories, known as cMERA [HOVV13, FRV17]. This takes a different perspective by formulating a variational class of states directly in the continuum, rather than considering a discretization. In both cases, the correlation functions of the theory are accurately reproduced up to some cut-off. The precise relation between MERA and cMERA is not very well understood, for instance it is not clear that discretizing a cMERA state could yield a MERA. Intriguingly, cMERA is formally strongly reminiscent of the *continuous wavelet transform* (CWT). The continuous wavelet transform [Mal08] can be defined for a much broader class of wavelet functions ψ, and if ψ is a biorthogonal wavelet the CWT can be discretized to a discrete wavelet transform. Reformulating cMERA as the second quantization of a CWT would therefore give a clear relationship between MERA and cMERA for free bosonic systems. A starting point could be the cMERA in [ZGV19], which reproduces some scaling dimensions exactly. However, the CWT appears to break some of the symplectic properties of the discrete biorthogonal wavelet transform and it remains an open problem to make this connection more explicit.

Part II

Index theory for approximately local dynamics

Introduction to quantum cellular automata

While quantum dynamics of closed systems are always unitary, systems of interest often possess an additional property: information propagates at finite speeds. In quantum field theories or local quantum circuits, information is strictly constrained to spread within a region called the light cone, or causal cone. Systems with strict causal cones are called *quantum cellular automata* (QCA) [Mar86, SW04]; or sometimes *locality-preserving unitaries*. However, the effective theories governing real physical systems are often only constrained by an *approximate* causal cone, see Fig. 6.2. For instance, nontrivial time evolution by a fixed local lattice Hamiltonian never satisfies a strict causal cone, but it does exhibit an approximate causal cone, given by the Lieb-Robinson bounds [LR72]. We can ask general questions about this class of dynamics, e.g. when can the evolution be generated by some local Hamiltonian, or when can one evolution be continuously deformed into another? These fundamental questions also have application in the study of topological phases in many-body physics [PFM+16, HFH18].

In this part of the dissertation, we will generalize a well-known classification of QCAs on spin chains to approximately local dynamics. In the current chapter we will review some basic notions of the theory of QCAs, and provide motivation and context for our results. In Chapter 7 we introduce the operator algebra formalism and present a theorem on stability of subalgebras due to Christensen which is a key ingredient for our results. In Chapter 8 we review in detail the GNVW index and provide some additional properties and alternative definition. Chapter 9 contains the main contribution: a generalization of the GNVW index to approximately local dynamics. These chapters closely follow [RWW20].

6.1   Quantum cellular automata

We start by giving a (slightly informal) definition of a QCA (which we will make completely formal in Definition 8.1). The idea is that a QCA is a quantum operation which is both *unitary* and *local*. To formalize this we consider a lattice Γ and at each site $n \in \mathbb{Z}$ we have a local Hilbert space $\mathbb{C}^{d_n}$. We let $\mathscr{A}_n = \mathscr{M}_{d_n \times d_n}$ the algebra of d_n by d_n complex matrices. This is the algebra of observables at site n. If Γ is a finite lattice we let

$$\mathscr{A}_\Gamma := \bigotimes_{n \in \Gamma} \mathscr{A}_n$$

which is the full algebra of observables on the lattice Γ. If Γ is an infinite lattice it may not be immediately clear how to define an 'infinite tensor product'. The appropriate algebra of observables on all of Γ will be the quasi-local algebra. For now we will ignore this technical subtlety and for the purpose of this introduction we simply use a finite lattice. We will also write

$$\mathscr{A}_X := \bigotimes_{n \in X} \mathscr{A}_n$$

for any (finite) subset $X \subseteq \Gamma$, which is the algebra of all observables supported in X. The algebra $\mathscr{A}_X$ is a subalgebra of $\mathscr{A}_\Gamma$, simply by tensoring with identity operators.

We will define a QCA α in the Heisenberg picture, that is, we need to define how it acts on observables. Recall that in the Heisenberg picture, if x is some operator, and u is a unitary, the operator is mapped to $u^* x u$, where u^* is the Hermitian adjoint (and hence inverse) of u. If Γ is a finite lattice with some distance measure d, then α is a QCA if there exists a unitary $u \in \mathscr{A}_\Gamma$ such that

$$\alpha(x) = u^* x u$$

and if this unitary preserves locality in the sense that if x is an operator that is supported on a set of sites $X \subset \Gamma$, then $\alpha(x)$ is supported on the set of sites within distance R of X. Here R is a fixed value, and we say that α has radius R. This definition is illustrated in Fig. 6.1. Another way to say that $\alpha(x) = u^* x u$ is by demanding that α is a $*$-*automorphism* of the algebra of observables (see Chapter 7, this is equivalent for finite systems, but will be the correct formal definition for infinite systems).

Let us give some concrete examples. First of all, any *local quantum circuit* will give rise to a QCA. Here a local quantum circuit is a unitary of the form

$$u = u_n \circ \ldots \circ u_2 \circ u_1$$

where

$$u_k = \prod_{X \subset \Gamma} u_{k,X}$$

and where $u_{k,X}$ is supported on X, and the subsets X range over a collection of disjoint bounded subsets (which may be different for each k). An obvious example are one-dimensional 'brick-layer' circuits.

A second fairly obvious example consists of automorphisms which simply permute the sites. For this example, consider the case where all local dimensions d_n are equal. Then, any permutation of Γ which maps every site $n \in \Gamma$ to a site within distance R of n defines a QCA. Such QCAs are known as *shifts*. A one-dimensional example is the translation shift, which corresponds to a unitary u which shifts every site to the right by one. These two examples are illustrated in Fig. 6.1.

The composition of two QCAs with radius R_1 and radius R_2 is again a QCA with radius $R_1 + R_2$, so in particular any composition of a circuit and a shift is also a QCA. Can we find any other examples of QCAs? Perhaps surprisingly this is not very easy, and such examples do not exist when the lattice has dimension one or two.

There is an extensive literature on QCAs. They have found application in computational theory (as QCAs are capable of universal computation) and in many body physics, for which we will describe a few applications in this introduction See [Far20, Arr19] for recent reviews on the theory of QCAs.
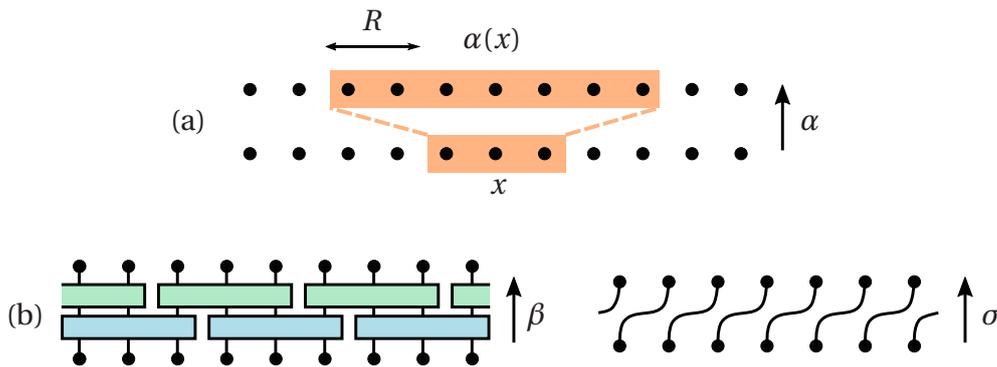
Figure 6.1: (a) Illustration of a QCA α with radius $R = 2$ mapping an operator x supported on three sites to an operator $\alpha(x)$ supported on seven sites. (b) A local circuit QCA β (left) and a translation QCA σ (right).

6.1.1 Index theory of one-dimensional QCAs

If the lattice has is one-dimensional (i.e. given by the integers $\mathbb{Z}$, or by a periodic chain), there is a beautiful classification of QCAs, which is based on an *index theory* provided by [GNVW12] (the GNVW index), based on ideas in [Kit06]. The intuition behind this index is that it measures the 'flow of information' in the QCA α. Roughly, it is defined by taking a cut along the chain and letting

$$\mathrm{ind}(\alpha) = \text{number of qubits moving to the right over the cut}$$
$$- \text{number of qubits moving to the left over the cut}$$

Of course this is a very informal definition (we will provide details in Chapter 8). If α is the translation QCA with local dimension d it is clear that from the above 'definition' that we should have $\mathrm{ind}(\alpha) = \log(d)$. On the other hand, it seems reasonable that if α is a quantum circuit it moves the same amount of information to the left as to the right, and we have $\mathrm{ind}(\alpha) = 0$. The index turns out to have various useful properties. For instance, it is additive under composition: $\mathrm{ind}(\alpha \circ \beta) = \mathrm{ind}(\alpha) + \mathrm{ind}(\beta)$. It takes discrete values in the set $\mathbb{Z}[\{\log(p_i)\}]$ where $\{p_i\}$ is the set of prime factors occurring in the local dimensions d_n. Finally, it can be shown that $\mathrm{ind}(\alpha) = 0$ *if and only if* α can be written as a circuit. This allowed [GNVW12] to classify one-dimensional QCAs: any such QCA can be written as a composition of a circuit and a tensor product of shifts.

6.1.2 QCAs in higher dimensions

In this dissertation we will focus on one-dimensional local dynamics. Higher dimensional lattices give rise to some very interesting open questions, and for this reason we give a brief overview of what is known about the structure of QCAs in higher dimensions.

   One way to understand the GNVW index is by the notion of a *boundary algebra*. To define this notion we consider the algebra $\mathscr{A}_X$ on some set of sites X, and a QCA α or radius R. The interior of X is the set of sites which are distance more than R away from the complement of X. Then if we look at $\alpha(\mathscr{A}_X)$ this certainly contains all operators which have support on the interior of X. The full algebra $\alpha(\mathscr{A}_X)$ is generated by the algebra of operators on the interior of X, and a boundary algebra. If we have a one-dimensional QCA, and we take X to be a sufficiently large interval, we find that the

boundary algebra consist of two commuting simple algebras, one supported on the left hand boundary of the interval and one supported on the right hand boundary of the interval. The GNVW index can now be computed as the logarithm of the ratio of the dimensions of these two algebras.

One can also study such boundary algebras in higher lattice dimensions. For a two-dimensional lattice it turns out that for a sufficiently large simply connected region its boundary algebra has a very specific structure. In this case, the boundary region is an annulus, and in [FH20] it was shown that the boundary algebra factorizes into a number of simple algebras $\mathscr{A}_i$ for $i = 1, \ldots, n$, where $\mathscr{A}_i$ has strictly local support, and the algebra $\mathscr{A}_i$ has overlapping support only with $\mathscr{A}_j$ for $j = i \pm 1 \mod n$. This structure allowed [FH20, Haa21] to prove that any two-dimensional QCA can be written, similarly to the one-dimensional case, as a composition of a circuit and a tensor product of shifts. An example of an application is given in [CX22], where the classification of two-dimensional QCAs is used to classify two-dimensional fermion-to-qubit mappings.

However, in [HFH18] it has been shown that (under assumption of a widely believed conjecture) the situation in three spatial dimensions is different. That is, there exist three-dimensional QCAs which can not be written as a composition of shifts and circuits. Similar examples where constructed in [FHH20, SCD$^+$22]. The idea in the construction of these examples is that the QCA disentangles a three-dimensional model with topological order. Then, one can show that the QCA is not a composition of a shift and a circuit, conditional on the conjecture that the three-dimensional model with topological order can not be described by a Hamiltonian which can be written as a sum of commuting local projectors. It is also known that the set of QCAs in a fixed dimension, modulo circuits and shifts, forms an abelian group under composition [FHH22]. As observed above, this group is trivial in one and two dimensions, but likely nontrivial in higher dimensions. For three spatial dimensions it has been conjectured that this group is given by the categorical Witt group of modular tensor categories [Haa21, SCD$^+$22]. Thus, there is a close connection between QCAs and certain aspects of topological order. Another approach to the classification problem is to study *Clifford quantum cellular automata*, which are QCAs which map Pauli operators to Pauli operators. For this class, modulo Clifford circuits and shifts, the group of Clifford QCAs is still trivial in one and two spatial dimensions, and nontrivial in three dimensions [Haa21].

6.1.3 Physical applications of QCAs

Locality is an important feature of quantum many-body physics, and in this light the notion of a QCA is very natural. Maybe the most fundamental appearance of locality is the fact that quantum field theories have strict causal cones (i.e. light cones). That is, if we consider a strictly local operator x on domain D in a spatial slice, then if we look at the time-evolved operator, it will be supported strictly in the light cone of D. This strict locality property is reminiscent of the definition of a QCA. A quantum field theory has both continuous space and time components. A natural thought is to use QCAs as a potential discretization (with space and time both discrete) which preserve the strict locality of the quantum field theory. For example [DP17, ABF20, BDPT18] and the review [Far20]. This perspective could be useful for simulation of quantum field theories on a quantum computer.

Another active line of research is the use of QCAs to study quantum hydrodynamics.

Here one starts from a QCA and iterates this QCA. Such models can be easier to solve than (integrable) spin chains, but still have interesting dynamics. For instance, one can study hydrodynamics and operator growth [ADM19, GHKV18, GNP21] in such models.

To mention a few other recent applications, QCAs have also been studied to construct symmetry-protected phases [SNBV$^+$19] and have been investigated from the perspective of tensor networks and matrix product operators [CPGSV17, ŞSBC18, GSSC20, PC20]. Finally, they have been used to classify topological phases of many-body localized dynamics [PFM$^+$16, ZL20], which we will briefly explain in Section 6.2.1. Also, as we saw above, three-dimensional QCAs are closely related to topological order.

6.2 Lieb-Robinson bounds and approximate locality

While in some contexts it is natural that quantum dynamics are *strictly* local (such as quantum field theory or in computational applications), for many realistic physical systems dynamics is only *approximately* local. The most iconic example is a lattice of spins where the dynamics is governed by a (quasi-)local Hamiltonian, in which case the so-called Lieb-Robinson bounds show that the dynamics induced by the Hamiltonian are approximately local. To formulate the Lieb-Robinson bounds, we again assume we have a lattice Γ (for instance $\mathbb{Z}^n$), and we assume that we have a local Hilbert space $\mathbb{C}^{d_n}$ at each site $n \in \Gamma$. Moreover, we assume that the systems has a Hamiltonian

$$H = \sum_{X \subseteq \Gamma} H_X$$

where each term H_X is supported on X. For instance, we could take a nearest neighbour Hamiltonian, where the subsets X are pairs of neighbouring sites, or we could more generally restrict X to local sets of some fixed bounded size k, so the interactions are k-local and spatially local. In this case it is well known that if we let $u(t) = e^{iHt}$ and if we let x_n be an operator supported at site n, and $x(t) = u(t)xu(t)^*$, then it holds that

$$\|[x_n(t), y_m]\| \le Ce^{\alpha(d(n,m)-vt)}\|x_n\|\|y_m\| \tag{6.1}$$

where y_m is any operator supported on site m, $d(n,m)$ is the distance between sites n and m and C, α and v are positive constants. These estimates are known as the Lieb-Robinson bounds, and the optimal possible v is the Lieb-Robinson velocity. The interpretation of Eq. (6.1) is that information spreads at most at linear speed (given by v) in the system. Indeed, at time t, $x_n(t)$ commutes up to an exponentially suppressed term with all operators supported at sites which are farther away than vt from n. Or in other words, $x_n(t)$ approximately has support in a ball of radius proportional to vt around site n. We introduce the Lieb-Robinson bounds more formally in Chapter 9. The key point is that in any discrete quantum many-body system one expects approximate locality rather than strict locality; hence for the application of theory of QCAs to many realistic physical systems it will be crucial to determine whether such results generalize to approximately local dynamics. In Chapter 9 we will generalize the notion of a QCA to an *approximately locality preserving unitary* (ALPU). Informally speaking, an ALPU is an automorphism α (i.e. conjugation by a unitary) which is such that it satisfies bounds of the type in Eq. (6.1), so for x supported on X it holds that

$$\|[\alpha(x), y]\| \le f(d(X, Y))\|x\|\|y\|$$
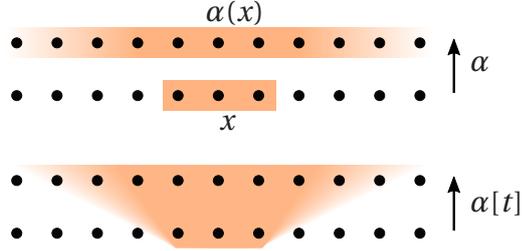
Figure 6.2: Illustration of an automorphism α with an approximate lightcone (an ALPU). Given such an α, does there exist a continuous dynamics $\alpha[t]$ with $\alpha[0] = I$ and $\alpha[1] = \alpha$ which remains approximately local at all times?

for any operator y supported on a set Y, where $d(X, Y)$ is the distance between these sets, and f is some fixed decaying function that determines how fast the tails decay (this could for instance be an exponentially decaying function or an inversely polynomially decaying function).

6.2.1   Floquet many-body localization and index theory

We will now provide an example of an ALPU which is *not* given as time evolution along a local Hamiltonian and where the GNVW index can be applied to understand quantum phases. This type of quantum systems are systems with *dynamical many-body localization* for Floquet systems in two dimensions [PFM$^+$16]. The intuitive idea is that on the two-dimensional lattice, under certain localization assumptions, time evolution of a subsystem with boundary defines an associated evolution on the one-dimensional boundary, and this evolution will be an ALPU in general, not necessarily given by a Hamiltonian evolution. The GNVW index of this boundary automorphism then captures whether the boundary has chiral transport, which is related to whether the two-dimensional system has vortex-like behavior.

Below we sketch the setup described by [PFM$^+$16]. See [HR17, DDP18, FPPV19, ZL20] for extensions. We consider a (time-dependent) local Hamiltonian H on a two-dimensional lattice, and we let U be the unitary obtained by time evolution for some fixed time T. The system exhibits many-body localization (MBL) if U can be written as a product of commuting unitaries which are all approximately local, i.e. when there exists a complete set of approximately local integrals of motion. More precisely, one says U is MBL in the sense of [PFM$^+$16] when it can be written

$$U = \prod_X u_X \tag{6.2}$$

where u_X is approximately supported in a set X of some bounded size and $[u_X, u_{X'}] = 0$ for all X, X'. What "approximately supported" means here depends on one's definition of many-body localization. A reasonable definition may be

$$\|[u_X, y]\| \leq C e^{-\gamma r} \|y\|$$

for any operator y which is supported on sites at least distance r away from X, and $C > 0$ and $\gamma > 0$ some constants. That is, the u_X satisfy Lieb-Robinson type bounds. To define

the invariant we let D denote the upper half plane (or in fact any simply connected infinite subset of the lattice) and we let U_D denote time evolution for time T using only the terms in the Hamiltonian strictly supported inside D. Also, we use (6.2) to define

$$U'_D = \prod_{X \subseteq D} U_X.$$

Then we let $V = U_D^{-1} U'_D$, so that the map $a \mapsto VaV^{-1}$ approximately preserves the algebra supported on a thick boundary strip ∂D of D. More precisely, the Lieb-Robinson bounds and (6.2) together show that for an operator a on a single site in ∂D, VaV^* is approximately supported within ∂D.

The key point is that while the original dynamics is generated by a two-dimensional lattice Hamiltonian, the induced dynamics on the boundary strip need not be generated by a Hamiltonian on the boundary strip. In [PFM$^+$16], one implicitly assumes that a small deformation of V actually defines a QCA on ∂D. Given this assumption, that MBL dynamics define a QCA on ∂D, one could then apply the index theory of one-dimensional QCAs to obtain a classification of MBL Floquet evolutions in 2D. The intuition is that if the Floquet dynamics has a 'vortex-like' behavior, the induced boundary dynamics will have a nonzero information flow, which is then captured by the GNVW index. To make this application of the GNVW index theory completely rigorous, it is required to study the robustness of the GNVW index, and see to what extent it generalizes to ALPUs.

6.2.2 Approximating approximately local dynamics by strictly local dynamics

It is well known that for Hamiltonian evolutions there is a close relationship between the approximately local time evolution and an approximation by strictly local quantum circuits. As explained in Section 1.2, given time evolution along a local (possibly time-dependent) Hamiltonian for time T, we can approximate the induced dynamics with a local quantum circuit. The most straightforward method is by Trotterization of the Hamiltonian. By increasing the depth of the circuit (i.e. decreasing the size of the time step in the Trotterization) one increases the accuracy of the approximation to arbitrary precision. A local quantum circuit is a QCA, and increased depth corresponds to a larger radius for the QCA. A key question, which we will address for one-dimensional systems, is whether one can generalize 'Trotterization' for more general approximately local dynamics. That is, given an ALPU, is there a sequence of QCAs of increasing radius which approximates the ALPU to arbitrary precision, and if so, what is the scaling of the radius with the precision of the approximation? An answer to this question will have to use methods which are fundamentally different from Trotterization, as this relies heavily on the fact that we have a Hamiltonian along which we evolve, and we can 'break up' this evolution in many small steps. In contrast, for an ALPU we are only given a single time step and there is no obvious way to approximate it by a QCA. There also exist methods for Hamiltonian simulation which are based more directly on the Lieb-Robinson bounds [HHKL18, TGS$^+$19], as mentioned in Section 1.2. Nevertheless, these are still based on the existence of a Hamiltonian and the possibility of restricting this Hamiltonian to a local region; therefore it is not clear how to apply such methods directly to an ALPU.
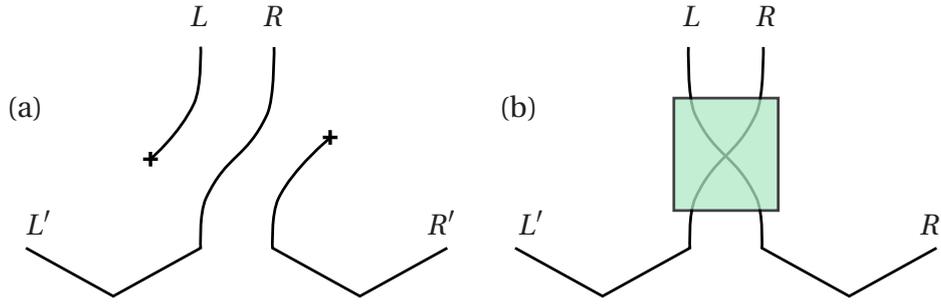
Figure 6.3: Illustration of (6.3). For the translation on a qudit with local dimension d in (a) we
have $I(L':R)_\phi = 2\log(d)$ and $I(L:R')_\phi = 0$, so the index equals d. For a circuit, one
can show that applying a local unitary as in (b) gives $I(L':R)_\phi = I(L:R')_\phi$, so the
index is zero.

6.3   Summary of contributions

6.3.1   Informal statement of results

A first result of this part of the dissertation is that we observe that one can re-formulate
the definition of the index as follows: one divides the chain into a left half L and right
half R, and one considers the Choi state $\phi_{LR,L'R'}$ of the automorphism α. Then the
mutual information difference

$$\mathrm{ind}(\alpha) = \frac{I(L':R)_\phi - I(L:R')_\phi}{2}, \tag{6.3}$$

is precisely the index of [GNVW12], but also well-defined for ALPUs with appropriately
decaying tails! This formula is illustrated in Fig. 6.3. In addition, the mutual information
enjoys much better continuity than the related expression for the index in Eq. (45)
of [GNVW12], which (in hindsight) can be understood as a difference of Rényi-2 en-
tropies. We show that the expression in Eq. (6.3) generalizes to the approximately local
setting.

   Our first main result consists of Theorems 9.16 and 9.18, summarized as

Approximation Theorem (informal)**.** *Suppose that α is an ALPU in one dimension. Then
there exists a sequence of QCAs α_j of increasing radius such that $\alpha_j(x)$ converges to $\alpha(x)$
for any local operator x, and that* $\mathrm{ind}(\alpha_j)$ *stabilizes for large j. We define*

$$\mathrm{ind}(\alpha) = \lim_{j\to\infty} \mathrm{ind}(\alpha_j).$$

If α has $\mathscr{O}(\frac{1}{r^{1+\delta}})$-tails for $\delta > 0$, the index defined in (6.3) *is finite and equal to* $\mathrm{ind}(\alpha)$. *The
exact index may also be computed locally through a rounding procedure.*

   To be precise, the error bounds are such that if α has tails like $\mathscr{O}(f(r))$, and α_j has
radius j, then for an operator x supported on an interval of n sites,

$$\|\alpha_j(x) - \alpha(x)\| = \mathscr{O}(f(j)(\frac{n}{j} + 1)\|x\|).$$

The key technical ingredient we use is a stability result for inclusions of possibly infinite algebras which we state as Theorem 7.6, an extension of results from [Chr77a, Chr80]. This result deals with the situation where $\mathscr{A}$ and $\mathscr{B}$ are algebras of observables and $\mathscr{A}$ is "nearly" included in $\mathscr{B}$, meaning that for each $x \in \mathscr{A}$ there is an element $y \in \mathscr{B}$ such that $\|x - y\| \le \varepsilon\|x\|$ for some small ε. Then (under some technical but very general assumptions on the algebras), there exists a unitary $u \in B(\mathscr{H})$ close to the identity in the sense that $\|u - I\| = \mathscr{O}(\varepsilon)$ with error independent of $\dim(\mathscr{H})$, such that $u\mathscr{A}u^*$ is strictly contained in $\mathscr{B}$. Loosely speaking, we construct the QCAs α_j by "localizing" the images $\alpha(\mathscr{A}_n)$ of the algebra $\mathscr{A}_n$ at each site n, by rotating $\alpha(\mathscr{A}_n)$ into an algebra supported within some radius of site n. The main technical effort in the construction is to ensure the rotations are compatible and the errors do not accumulate.

The index defines equivalence classes of ALPUs. These are characterized in our second main result, Theorem 9.25, sketched below:

Classification Theorem (informal). *Suppose α and β are ALPUs with $f(r)$-tails in one dimension. Then the following are equivalent conditions:*

(i) $\mathrm{ind}(\alpha) = \mathrm{ind}(\beta)$.

(ii) $\alpha = \beta\gamma$ *where* $\mathrm{ind}(\gamma) = 0$.

(iii) *There exists a "blended" ALPU which (up to small error) matches α on the left of the chain and matches β on the right.*

(iv) *There exists a strongly continuous path from α to β through the space of ALPUs with $g(r)$-tails for some $g(r) = o(1)$. If such a path exists, then it can be generated by evolving a time-dependent quasi-local Hamiltonian for unit time.*

In particular, (iv) provides a converse to the Lieb-Robinson bounds in one dimension: an automorphism can be generated by evolution along a time-dependent Hamiltonian with certain locality bounds if and only if it has index zero. Thus we see that the index theory of [GNVW12] completely generalizes to ALPUs and does not "collapse," with as only essential difference that the role of quantum circuits is replaced by time evolutions along time-dependent geometrically local Hamiltonians.

As an application, it follows immediately that the translation operator cannot be implemented by a finite time evolution of any (time-dependent) Hamiltonian satisfying Lieb-Robinson bounds. Moreover, there cannot exist a quasi-local "momentum density" that generates a lattice translation and also satisfies Lieb-Robinson bounds with $o(1)$-tails at all times. To show that it is necessary to impose some bound on the decay of the ALPU tails in our constructions, we give an example of a strongly continuous path of automorphisms generated by a Hamiltonian with $\frac{1}{r}$-decaying interactions that connects the identity map to a translation on a chain of qubits, showing that at this point the index theory does indeed collapse.

6.3.2 Prior work

To generalize the GNVW index to ALPUs, a natural concern is the sensitivity of the index to small perturbations of the QCA. However, the dependence of the sensitivity on the local Hilbert space dimension and the radius of the QCA is not immediately

clear from the considerations in [GNVW12], which yield a relatively weak continuity estimate. Without stronger bounds, it would have been possible that the homotopy classes established in [GNVW12] could collapse when considering ALPUs: two QCAs with different GNVW index might still be connected by a strongly continuous path through the space of ALPUs with some prescribed tails. Another concern was whether the generalized index would take values in the same discrete set as the GNVW index.

These questions are recognized in existing literature. In fact, one of the main open questions in the original work [GNVW12] was how to extend the index theory to some class of automorphisms with only approximate locality. Later work asked specifically whether ALPUs could be approximated by QCAs [Has13]. In the review [Far20] such questions were raised again, highlighting their relevance for the application of index theory to actual physical systems. Other recent work [FH20] also suggested the extension to ALPUs as an avenue for research, proposing that one approach might involve Ulam stability results for operator algebras. This is precisely the approach taken in this work. The stability results we use [Chr77a, Chr80] and augment were developed throughout the 1970s and 80s for studying how operator algebras behave under perturbation. Intriguingly, related questions about perturbations were tackled under a different guise in [CRSV17] (cf. their Theorem 3.6), in the context of quantum device certification.

Regarding the converse to the Lieb-Robinson bounds, see [WW20] for interesting work which develops a related converse with different assumptions. They show that if you already know α is generated by a k-local Hamiltonian satisfying a Lieb-Robinson-like condition, then the evolution can also be generated by a *geometrically* local k-local Hamiltonian. Their condition can be checked at infinitesimal times. In contrast, we do not assume the ALPU is generated by any k-local Hamiltonian.

Finally, our reformulation of the GNVW index in terms of the mutual information is similar to an expression which has been derived in [DDP18] in the context of two-dimensional Floquet phases. This expression has also been rederived, shortly after the appearance of our work [RWW20], in [GPC21].

Operator algebras and stability of subalgebras

For infinite dimensional quantum mechanical systems it is often more convenient to work with operator algebras (algebras of observables) rather than Hilbert spaces, and use the Heisenberg rather than Schrödinger picture of quantum mechanics. A standard reference for operator algebras and their relation to quantum physics is [BR12]; see [Naa13] for an accessible introduction. We review C^*-algebras and von Neumann algebras, focusing especially on facts used in subsequent proofs. Then we turn to methods for "perturbations" (e.g. small rotations) of operator algebras in Section 7.1.

7.0.1 C^*-algebras

The notion of an operator algebra is formalized by a C^*-*algebra*, which is a complex algebra $\mathscr{A}$ with a norm $\|\cdot\|$ and an anti-linear involution $x \mapsto x^*$, satisfying

- $\mathscr{A}$ is complete in $\|\cdot\|$,

- $\|xy\| \leq \|x\|\|y\|$,

- $\|x^*x\| = \|x\|^2$.

We will only use algebras with an identity element I. An important example is the C^*-algebra $B(\mathscr{H})$ of operators on some Hilbert space $\mathscr{H}$, where we take the operator norm as the norm, and the adjoint as the $*$-operation. In finite dimensions this reduces to the algebra $\mathscr{M}_{d \times d}$ of complex $d \times d$ matrices with the spectral norm and Hermitian conjugate. A C^*-algebra $\mathscr{A}$ is called *approximately finite-dimensional* (AF) if it contains a directed collection of finite-dimensional subalgebras whose union is dense in $\mathscr{A}$. If $\mathscr{A} \subseteq \mathscr{B}$ are C^*-algebras we define the *commutant* of $\mathscr{A}$ in $\mathscr{B}$ to be the set

$$\mathscr{A}' = \{x \in \mathscr{B} \text{ such that } [x, \mathscr{A}] = 0\},$$

which is again a C^*-algebra. We denote by $U(\mathscr{A})$ the set of elements $u \in \mathscr{A}$ that are *unitary*, meaning that $uu^* = u^*u = I$.

A $*$-*homomorphism* $\alpha \colon \mathscr{A} \to \mathscr{B}$ between C^*-algebras is an algebra homomorphism which also preserves the $*$-operation, $\alpha(x^*) = \alpha(x)^*$. Such a $*$-homomorphism is automatically continuous and indeed contractive, i.e., $\|\alpha(x)\| \leq \|x\|$. The latter can also

be written as $\|\alpha\| \leq 1$, where we define the notation $\|\beta\| = \sup_{\|x\| \leq 1} \|\beta(x)\|$ for any linear map β between C^*-algebras. A *(∗-)automorphism* is a bijective $*$-homomorphism. The inverse of an automorphism is again a $*$-homomorphism, and any automorphism is automatically unital and isometric. We write I for the identity automorphism. Finally, a *state* on a C^*-algebra $\mathscr{A}$ is given by a linear functional $\omega: \mathscr{A} \to \mathbb{C}$ which is positive (meaning that $\omega(x^*x) \geq 0$ for all $x \in \mathscr{A}$) and normalized (meaning that $\omega(I) = 1$).

It turns out that any C^*-algebra can be realized as a subalgebra of $B(\mathscr{H})$, the algebra of bounded operators on some Hilbert space $\mathscr{H}$. This is proven by the following result known as the *Gelfand-Naimark-Segal (GNS) construction* or *representation*:

Theorem 7.1 (Gelfand-Naimark-Segal). *Given a state ω on $\mathscr{A}$, there exists a Hilbert space $\mathscr{H}$, a $*$-homomorphism $\pi: \mathscr{A} \to B(\mathscr{H})$, and a cyclic vector ϕ (meaning $\pi(\mathscr{A})\phi$ is dense in $\mathscr{H}$) such that*

$$\omega(x) = \langle \phi, \pi(x)\phi \rangle$$

Moreover, if $(\mathscr{H}', \pi', \phi')$ is another triple as above then there exists a unique unitary $u: \mathscr{H} \to \mathscr{H}'$ such that $\phi' = u\phi$ and $\pi'(x) = u\pi(x)u^*$ for all $x \in \mathscr{A}$.

If ω is such that $\omega(x^*x) = 0$ implies $x = 0$, then the GNS representation is faithful (meaning that π_ω is injective). In that case, one way to construct the Hilbert space in Theorem 7.1 is by letting $\langle x, y \rangle = \omega(x^*y)$ define an inner product on $\mathscr{A}$ and letting $\mathscr{H}$ be the completion of $\mathscr{A}$ with respect to this inner product. Then $\mathscr{A}$ acts on $\mathscr{H}$ by left multiplication, which defines the $*$-homomorphism $\pi: \mathscr{A} \to B(\mathscr{H})$. The identity $I \in \mathscr{A}$ gives rise to a cyclic vector $\phi \in \mathscr{H}$.

7.0.2   Von Neumann algebras

A special class of C^*-algebras are von Neumann algebras. A C^*-algebra $\mathscr{A} \subseteq B(\mathscr{H})$ is a *von Neumann algebra* if it is equal to its double commutant in $B(\mathscr{H})$,

$$\mathscr{A} = \mathscr{A}''.$$

In fact, for any subset $S \subseteq B(\mathscr{H})$, the double commutant S'' is always a von Neumann algebra, called the von Neumann algebra generated by S. It is the smallest von Neumann algebra that contains S. There are various relevant topologies on $B(\mathscr{H})$. The *strong operator topology* is such that a net x_i converges to some operator x if and only if it holds that $x_i v \to xv$ for each vector $v \in \mathscr{H}$. The *weak operator topology* is such that a net x_i converges to some operator x if and only if $\langle w, x_i v \rangle \to \langle w, xv \rangle$ for each pair $v, w \in \mathscr{H}$. The weak operator topology is weaker than the strong operator topology, and both are weaker than the topology induced by the norm. Sometimes also the weak-$*$ topology is relevant, induced by interpreting $B(\mathscr{H})$ as the dual space of the trace class operators on $\mathscr{H}$. The weak operator topology is weaker than the weak-$*$ topology, but the two coincide on norm-bounded subsets of $B(\mathscr{H})$. On convex subsets the weak operator closure and strong operator closure coincide. The von *Neumann bicommutant theorem* states that for unital $*$-subalgebra $\mathscr{A} \subseteq B(\mathscr{H})$, $\mathscr{A}''$ is the weak operator closure of $\mathscr{A}$, so $\mathscr{A}$ is a von Neumann algebra if and only if $\mathscr{A}$ is weak operator closed. Any $*$-automorphism of a von Neumann algebra is continuous with respect to the weak-$*$ topology. Moreover, norm balls are compact in the weak operator topology (Theorem 5.1.3 in [KR97], a consequence of the Banach-Alaoglu theorem).

A useful fact in the study of von Neumann algebras is the *Kaplansky density theorem*, which states that for any self-adjoint subalgebra $\mathscr{A} \subseteq B(\mathscr{H})$ the unit ball of the strong operator closure of $\mathscr{A}$ equals the strong operator closure of the unit ball of $\mathscr{A}$. We refer to [BR12, §2.4] for more details. Another useful fact is that the norm is lower semi-continuous in the weak operator topology, i.e., if a net x_i converges to x in the weak operator topology then $\|x\| \le \liminf_i \|x_i\|$.

In infinite dimensions, working with a von Neumann algebra often confers advantages over more general C^*-algebras. For instance, the output of the Borel functional calculus (taking functions of operators) on a C^*-algebra $\mathscr{A} \subseteq B(\mathscr{H})$ produces operators that sometimes lie outside $\mathscr{A}$, but they always lie in the weak operator closure $\mathscr{A}''$. A von Neumann algebra therefore allows one to use spectral projections and other technical tools.

A von Neumann algebra $\mathscr{A} \subseteq B(\mathscr{H})$ is called a *factor* if it has trivial center, in other words $\mathscr{A}' \cap \mathscr{A} = \mathbb{C}I$. In particular, $\mathscr{A} = B(\mathscr{H})$ is a factor (a so-called type I factor). Any finite dimensional factor is of this form (for a finite dimensional Hilbert space), but there also exist infinite dimensional factors *not* of the form $B(\mathscr{H})$ (so-called type II and type III factors). A von Neumann algebra $\mathscr{A}$ is called *hyperfinite* (or approximately finite-dimensional) if it contains a directed collection of finite-dimensional subalgebras whose union is dense in the weak operator topology (equivalently, the weak-$*$ topology). Equivalently, $\mathscr{A}$ is hyperfinite when there exists an AF C^*-subalgebra $\mathscr{A}_0 \subseteq \mathscr{A}$ such that $\mathscr{A}_0'' = \mathscr{A}$.

Finally, if $\mathscr{M} \subseteq B(H)$ and $\mathscr{N} \subseteq B(\mathscr{K})$ are von Neumann algebras, we use $\mathscr{M} \otimes \mathscr{N} \subseteq B(\mathscr{H} \otimes \mathscr{K})$ to denote the von Neumann algebra tensor product, given by the weak operator closure of the algebraic tensor product of $\mathscr{M}$ and $\mathscr{N}$ in $B(\mathscr{H} \otimes \mathscr{K})$.

7.0.3   The quasi-local algebra

We will apply the formalism of operator algebras in the setting of lattice spin systems. If we have a system of a finite number of spins $\mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_n}$, the corresponding operator algebra is simply the full matrix algebra $\mathscr{M}_{d_1 \times d_1} \otimes \cdots \otimes \mathscr{M}_{d_n \times d_n}$. However, for infinitely many spins the tensor product structure becomes ambiguous. If the spins form a lattice the physically appropriate choice of C^*-algebra is the *quasi-local algebra*. Consider a lattice Γ, and associate a finite-dimensional matrix algebra $\mathscr{A}_n = \mathscr{M}_{d_n \times d_n}$ to each element n of the lattice. We assume that there is a uniform upper bound on the dimensions d_n. For any *finite* subset $X \subseteq \Gamma$ we can define the algebra $\mathscr{A}_X = \bigotimes_{n \in X} \mathscr{A}_n$. These algebras naturally form a local net, meaning that for any two subsets $X \subseteq X'$ we have a natural inclusion $\mathscr{A}_X \subseteq \mathscr{A}_{X'}$ (by tensoring with the identity on $X' \setminus X$), and for any two disjoint subsets $X \cap X' = \emptyset$ we have that $[\mathscr{A}_X, \mathscr{A}_{X'}] = 0$ (we embed the two algebras into any $\mathscr{A}_{X''}$ such that $X \cup X' \subseteq X''$). This allows us to define the algebra of all *strictly local* operators as

$$\mathscr{A}_\Gamma^{\text{strict}} = \bigcup_{X \subseteq \Gamma \text{ finite}} \mathscr{A}_X.$$

This is a $*$-algebra which inherits a norm from the $\mathscr{A}_X$, but it is not complete for this norm. We define the *quasi-local algebra* $\mathscr{A}_\Gamma$ to be the norm completion of $\mathscr{A}_\Gamma^{\text{strict}}$. Thus, $\mathscr{A}_\Gamma$ is a C^*-algebra. For infinite subsets $X \subseteq \Gamma$, we define $\mathscr{A}_X$ correspondingly as a norm-complete C^*-subalgebra of $\mathscr{A}_\Gamma$. Then we have inclusions $\mathscr{A}_X \subseteq \mathscr{A}_\Gamma$ for any subset $X \subseteq \Gamma$. If $x \in \mathscr{A}_X$, we say x is *supported* on X.

The quasi-local algebra has a natural state τ, called the *tracial state*, which can be thought of as the generalization of the maximally mixed state to an infinite lattice. It is defined on $x \in \mathscr{A}_X$ for finite $X \subseteq \Lambda$ by

$$\tau(x) = \frac{1}{d_X} \operatorname{tr}(x),$$

where $d_X = \prod_{n \in X} d_n$, and can be extended to the full algebra.

We consider the GNS representation $\pi \colon \mathscr{A}_\Gamma \to B(\mathscr{H})$ from Theorem 7.1 of the quasi-local algebra using the tracial state τ, and we let

$$\mathscr{A}_\Gamma^{\mathrm{vN}} = \pi(\mathscr{A}_\Gamma)'' \subseteq B(\mathscr{H}), \tag{7.1}$$

denote the von Neumann algebra generated by the GNS representation of the quasi-local algebra. The right-hand side is also the weak operator closure of the image $\pi(\mathscr{A}_\Gamma)$. It turns out $\mathscr{A}_\Gamma^{\mathrm{vN}}$ is a proper subalgebra of $B(\mathscr{H})$, which remains true even for Γ finite; in our case that Γ is infinite, $\mathscr{A}_\Gamma^{\mathrm{vN}}$ is the (unique up to unique isomorphism) *hyperfinite type II_1 factor*. This algebra is extensively studied, but for our purpose we will only need to observe that this factor is hyperfinite (as follows directly from its construction). If $X \subseteq \Gamma$ we denote $\mathscr{A}_X^{\mathrm{vN}} = \pi(\mathscr{A}_X)''$. Each $\mathscr{A}_X^{\mathrm{vN}}$ is hyperfinite and has the property that $(\mathscr{A}_X^{\mathrm{vN}})' \cap \mathscr{A}_\Gamma^{\mathrm{vN}} = \mathscr{A}_{\Gamma \setminus X}^{\mathrm{vN}}$. Since the tracial state is faithful, we may think of each $\mathscr{A}_X$ as a subalgebra of $\mathscr{A}_X^{\mathrm{vN}}$. Moreover, it holds that $\mathscr{A}_X^{\mathrm{vN}} \cap \mathscr{A}_{\mathbb{Z}} = \mathscr{A}_X$ for any $X \subseteq \mathbb{Z}$.[1]

The reason for introducing $\mathscr{A}_\Gamma^{\mathrm{vN}}$ is purely to be able to use technical tools, especially Theorem 7.6, from the study of von Neumann algebras. Our main results are all formulated in terms of the quasi-local algebra.

We observe that an automorphism α of $\mathscr{A}_\Gamma$ extends naturally to the associated von Neumann algebra in (7.1), as follows. If τ is the tracial state on the quasi-local algebra $\mathscr{A}_\Gamma$, then for any automorphism of $\mathscr{A}_\Gamma$ this state is left invariant, i.e., $\tau \circ \alpha = \tau$. (One way to see this is by using that τ is the unique state for which $\tau(xy) = \tau(yx)$ for all $x, y \in \mathscr{A}_\Gamma$.) By the uniqueness of the GNS construction this implies that α can be implemented by a unitary u on $\mathscr{H}$, in the sense that $\pi(\alpha(x)) = u\pi(x)u^*$. Therefore, α extends to an automorphism of the hyperfinite von Neumann algebra $\mathscr{A}_\Gamma^{\mathrm{vN}}$, which we denote by the same symbol α if there is no danger of confusion. Note that this extension is necessarily unique.

We will mostly consider the situation where $\Gamma = \mathbb{Z}$ is the discrete line. If

$$X = \{ m \in \mathbb{Z} \text{ such that } m \le n \}$$

---

[1] Since $\mathscr{A}_X^{\mathrm{vN}} = (\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$, it suffices to show that $(\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}} = \mathscr{A}_X$. We will argue that $(\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}} \subseteq \mathscr{A}_X$, since the other inclusion is immediate. To this end, let $x \in (\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}$. Since $x \in \mathscr{A}_{\mathbb{Z}}$, we can choose a sequence $x_i \in \mathscr{A}_{X_i}$ converging to x in norm and such that each X_i is a finite set. On the other hand, $x \in (\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})'$ implies that, for any $y \in \mathscr{A}_{\mathbb{Z} \setminus X}$ it holds that

$$\| [y, x_i] \| = \| [y, x_i - x] \| \le 2 \|y\| \|x - x_i\|. \tag{7.2}$$

Now let $\tilde{x}_i = \int_{U(\mathscr{A}_{X_i \cap (\mathbb{Z} \setminus X)})} u x_i u^* \, du$, similarly as in Eq. (7.4). Then $\tilde{x}_i \in \mathscr{A}_{X_i \cap X} \subseteq \mathscr{A}_X$, since it is an element of $\mathscr{A}_{X_i}$ that commutes with $\mathscr{A}_{X_i \cap (\mathbb{Z} \setminus X)} = \mathscr{A}_{X_i \setminus X}$. On the other hand, Eq. (7.2) implies (cf. Eqs. (7.5) and (7.6)) that

$$\| \tilde{x}_i - x_i \| \le \int_{U(\mathscr{A}_{X_i \cap (\mathbb{Z} \setminus X)})} \| u x_i u^* - x_i \| \, du \le \int_{U(\mathscr{A}_{X_i \cap (\mathbb{Z} \setminus X)})} \| [u, x_i] \| \, du \le 2 \|x - x_i\|.$$

Therefore $\| \tilde{x}_i - x \| \le 3 \|x - x_i\| \to 0$ and since each $\tilde{x}_i \in \mathscr{A}_X$, we conclude that $x \in \mathscr{A}_X$.

we will write $\mathscr{A}_{\leq n} := \mathscr{A}_X$ and similarly if $X = \{m \in \mathbb{Z} \text{ such that } m \geq n\}$ we will write $\mathscr{A}_{\geq n} := \mathscr{A}_X$. We use the same notation to describe subalgebras $\mathscr{A}_{\leq n}^{\mathrm{vN}}, \mathscr{A}_{\geq n}^{\mathrm{vN}}$ of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$.

7.1 Near inclusions and stability properties

We now define our notion of near inclusions of algebras and discuss related stability properties. The notion of a near inclusion follows e.g. [Chr80].

Definition 7.2 (Near inclusion)**.** For a C^*-algebra $\mathscr{B} \subseteq B(\mathscr{H})$ and an operator $a \in B(\mathscr{H})$, we write $a \overset{\varepsilon}{\in} \mathscr{B}$ when there exists $b \in \mathscr{B}$ such that $\|a - b\| \leq \varepsilon \|a\|$. Likewise for two C^*-algebras $\mathscr{A}, \mathscr{B} \subseteq B(\mathscr{H})$, we write $\mathscr{A} \overset{\varepsilon}{\subseteq} \mathscr{B}$ and say there is a *near inclusion* whenever it holds that $a \overset{\varepsilon}{\in} \mathscr{B}$ for all $a \in \mathscr{A}$.

We note that if $\mathscr{B}$ is a von Neumann algebra, we have $a \overset{\varepsilon}{\in} \mathscr{B}$ if and only if

$$\inf_{b \in \mathscr{B}} \|a - b\| \leq \varepsilon \|a\|.$$

That is, the infimum is attained by some $b \in \mathscr{B}$. Indeed, let b_i be a sequence in $\mathscr{B}$ such that $\lim_i \|a - b_i\| \leq \varepsilon \|a\|$. In particular, $\|b_i\|$ is bounded. Since (any multiple of) the closed unit ball in $B(\mathscr{H})$ is compact in the weak operator topology (e.g., Theorem 5.1.4 in [KR97]), the sequence b_i has a limit point $b \in \mathscr{B}$ in the weak operator topology. Because the norm is lower semi-continuous in the weak operator topology,

$$\|a - b\| \leq \lim_i \|a - b_i\| \leq \varepsilon \|a\|,$$

which concludes the argument.

When $\mathscr{B} \subseteq B(\mathscr{H})$ is a C^*-algebra and $x \in \mathscr{B}(\mathscr{H})$ is an operator that is nearly contained in its commutant, say $x \overset{\varepsilon}{\in} \mathscr{B}'$, then it is easy to see that, for any $b \in \mathscr{B}$,

$$\|[x, b]\| \leq 2\varepsilon \|x\| \|b\|. \tag{7.3}$$

Indeed $x \overset{\varepsilon}{\in} \mathscr{B}'$ means there exists $y \in \mathscr{B}'$ such that $\|x - y\| \leq \varepsilon \|x\|$. Then we have for any $b \in \mathscr{B}$ that

$$\|[x, b]\| = \|[x - y, b]\| \leq 2\|x - y\| \|b\| \leq 2\varepsilon \|x\| \|b\|.$$

We will be interested in the converse of this statement, which is rather less clear.

To gain some intuition, we consider the finite-dimensional setting. Suppose that we have a tensor product of finite-dimensional Hilbert spaces $\mathscr{H} = \mathscr{H}_A \otimes \mathscr{H}_B$, and let $\mathscr{B} = I \otimes B(\mathscr{H}_B) \subseteq B(\mathscr{H})$ be the algebra of operators supported on the second tensor factor. Then we can define a projection onto the commutant of $\mathscr{B}$ by twirling using Haar probability measure on the group $U(\mathscr{B})$ of unitaries on $\mathscr{B}$:

$$\mathbb{E}_{\mathscr{B}'}: B(\mathscr{H}) \to \mathscr{B}', \quad \mathbb{E}_{\mathscr{B}'}(x) = \int_{U(\mathscr{B})} uxu^* \, \mathrm{d}u. \tag{7.4}$$

In fact, the commutant is simply $\mathscr{A} = \mathscr{B}' = B(\mathscr{H}_A) \otimes I_B$, and the projection can equivalently be written in terms of the normalized partial trace, $\mathbb{E}_{\mathscr{B}'}(x) = \frac{1}{d_B} \mathrm{tr}_{\mathscr{B}}(x)$, where $d_B = \dim \mathscr{H}_B$. The projection exhibits the desirable property that if

$$\|[x, b]\| \leq \varepsilon \|x\| \|b\| \tag{7.5}$$

for all $b \in \mathscr{B}$, then

$$\|x - \mathbb{E}_{\mathscr{B}'}(x)\| \le \int_{U(\mathscr{B})} \|x - uxu^*\| \, \mathrm{d}u = \int_{U(\mathscr{B})} \|[x, u]\| \, \mathrm{d}u \le \varepsilon \|x\|. \qquad (7.6)$$

This shows that, in the finite-dimensional setting, the commutator bound (7.5) implies that $x \overset{\varepsilon}{\in} \mathscr{B}'$.

In infinite dimensions, where no Haar integral is available, we need a different way to define the projection. One way to do so is using the so-called "property P." If $\mathscr{B} \subseteq B(\mathscr{H})$ is a von Neumann algebra, it has *property P* if for any $x \in B(\mathscr{H})$, there exists some $y \in \mathscr{B}'$ such that y is also in the weak operator closure (equivalently, the weak-$*$ closure) of the convex hull of $\{uxu^* : u \in U(\mathscr{B})\}$. Note that in the finite-dimensional setting this is immediate from the definition in terms of the Haar integral. We can also generalize the notion of twirling by a (non-commutative) *conditional expectation* $\mathbb{E}_{\mathscr{A}}$ onto a von Neumann algebra $\mathscr{A} \subseteq B(\mathscr{H})$. This is defined to be contractive completely positive linear map $\mathbb{E}_{\mathscr{A}} : B(\mathscr{H}) \to \mathscr{A} \subseteq B(\mathscr{H})$ which is such that for $x \in B(\mathscr{H})$ and $a, a' \in \mathscr{A}$ we have $\mathbb{E}_{\mathscr{A}}(a) = a$ and $\mathbb{E}_{\mathscr{A}}(axa') = a\mathbb{E}_{\mathscr{A}}(x)a'$. A von Neumann algebra $\mathscr{A} \subseteq B(\mathscr{H})$ is called *injective* if there exists such a conditional expectation [Bla06, IV.2.1.4].

For von Neumann algebras acting on separable Hilbert spaces, these properties are equivalent to each other and to hyperfiniteness as defined earlier:

Theorem 7.3. *Let $\mathscr{B} \subseteq B(\mathscr{H})$ be a von Neumann algebra with $\mathscr{H}$ separable. Then the following are equivalent:*

  (i) *$\mathscr{B}$ is hyperfinite.*

 (ii) *$\mathscr{B}'$ is hyperfinite.*

(iii) *$\mathscr{B}$ has property P.*

(iv) *$\mathscr{B}'$ is injective.*

If $\mathscr{H}$ is not assumed to be separable, it is still true that (i) implies (iii) and (iii) implies (iv). Moreover, $\mathscr{B}$ has property P if and only if $\mathscr{B}'$ has property P, and the same is true for injectivity.

For a comprehensive account of the theory and classification of von Neumann algebras see [Tak03, Bla06]. Theorem 7.3 is proved in Proposition 4.1 of [NSW13] in the case that $\mathscr{B}$ is a factor. The general case follows similarly by combination of well-known results, as we sketch for convenience.

Proof. The implications (i)⇒(iii) and (iii)⇒(iv) are explained in [Bla06, IV.2.2.20], as is the fact that $\mathscr{B}$ has property P if and only if $\mathscr{B}'$ has property P. Moreover, [Bla06, IV.2.2.7] asserts that $\mathscr{B}$ is injective if and only if $\mathscr{B}'$ is injective. Now assume that $\mathscr{H}$ is separable or, equivalently, $\mathscr{B}$ has a separable predual. In this case, injectivity implies hyperfiniteness [Bla06, IV.2.6.1], so it follows that (iv)⇒(ii). Since $\mathscr{B}'' = \mathscr{B}$, (i)⇒(ii) also yields (ii)⇒(i), so that (i)–(iv) are all equivalent. ∎

When $\mathscr{B}$ is hyperfinite and $x \in B(\mathscr{H})$ is such that $\|[x, b]\| \le \varepsilon \|x\| \|b\|$ for all $b \in \mathscr{B}$, then $x \overset{\varepsilon}{\in} \mathscr{B}'$, providing a converse to the discussion above Eq. (7.3). Indeed, since $\mathscr{B}$ has property P by (i)⇒(iii), there exists some $y \in \mathscr{B}'$ in the weak operator closure of the

convex hull of $\{uxu^* : u \in U(\mathcal{B})\}$. Using lower semicontinuity of the norm with respect to the weak operator topology, we find

$$\|x - y\| \leq \sup_{u \in U(\mathcal{B})} \|x - uxu^*\| = \sup_{u \in U(\mathcal{B})} \|[x, u]\| \leq \varepsilon\|x\|,$$

which shows that $x \overset{\varepsilon}{\in} \mathcal{B}'$. Moreover, suppose that $x \in \mathcal{M}$, $\mathcal{B} \subseteq \mathcal{M}$ for a von Neumann algebra $\mathcal{M} \subseteq B(\mathcal{H})$, then we have that $x \overset{\varepsilon}{\in} \mathcal{B}' \cap \mathcal{M}$. Indeed, in this case $\{uxu^* : u \in U(\mathcal{B})\}$ is contained in $\mathcal{M}$, and the same is true for the weak operator closure of its convex hull. Since y was constructed as an element of the latter, it follows that $y \in \mathcal{B}' \cap \mathcal{M}$ and hence $x \overset{\varepsilon}{\in} \mathcal{B}' \cap \mathcal{M}$.

In turn, the above implies that any conditional expectation $\mathbb{E}_{\mathcal{B}'} : B(\mathcal{H}) \to \mathcal{B}'$ (and such conditional expectations exist due to (iii) $\Rightarrow$ (iv)) satisfies

$$\|\mathbb{E}_{\mathcal{B}'}(x) - x\| \leq \|\mathbb{E}_{\mathcal{B}'}(x) - \mathbb{E}_{\mathcal{B}'}(y)\| + \|y - x\| \leq 2\varepsilon\|x\|,$$

using that $\mathbb{E}_{\mathcal{B}'}(y) = y$ and that conditional expectations are contractions. When $\mathcal{B}$ is a factor, a different proof strategy shows that the constant 2 can be omitted; see Proposition 4.1 in [NSW13].

As an easy consequence we obtain:

Lemma 7.4 (Near inclusions and commutators [Chr77b, Theorem 2.3])**.** *Let $\mathcal{A}, \mathcal{B} \subseteq B(\mathcal{H})$ be two C^*-algebras. If $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}'$ is a near inclusion, then*

$$\|[a, b]\| \leq 2\varepsilon\|a\|\|b\|.$$

holds for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$.

Conversely, if $\mathcal{B}$ is a hyperfinite von Neumann algebra and

$$\|[a, b]\| \leq \varepsilon\|a\|\|b\|$$

holds for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$, then we have a near inclusion $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}'$. If moreover $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}$ for some von Neumann algebra $\mathcal{M} \subseteq B(\mathcal{H})$, then $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}' \cap \mathcal{M}$.

Proof. The first claim follows from Eq. (7.3), since $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}'$ means that $a \overset{\varepsilon}{\in} \mathcal{B}'$ for every $a \in \mathcal{A}$. For the converse claim, the discussion above the lemma shows that for every $a \in \mathcal{A}$ we have $a \overset{\varepsilon}{\in} \mathcal{B}'$, hence $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}'$; moreover, if $a \in \mathcal{M}$, $\mathcal{B} \subseteq \mathcal{M}$ then $a \overset{\varepsilon}{\in} \mathcal{B}' \cap \mathcal{M}$, and hence $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}' \cap \mathcal{M}$. ∎

As a straightforward consequence of Lemma 7.4, we in turn obtain the following:

Lemma 7.5 (Near inclusion of commutants)**.** *Let $\mathcal{C}, \mathcal{D} \subseteq B(\mathcal{H})$ be von Neumann algebras with $\mathcal{C}$ hyperfinite. If $\mathcal{C} \overset{\varepsilon}{\subseteq} \mathcal{D}$, then $\mathcal{D}' \overset{2\varepsilon}{\subseteq} \mathcal{C}'$. If moreover $\mathcal{C} \subseteq \mathcal{M}$ for a von Neumann algebra $\mathcal{M} \subseteq B(\mathcal{H})$, then $\mathcal{D}' \cap \mathcal{M} \overset{2\varepsilon}{\subseteq} \mathcal{C}' \cap \mathcal{M}$.*

Proof. It suffices to prove the second statement, since it reduces to the first if we choose $\mathcal{M} = B(\mathcal{H})$. Since $\mathcal{C} \overset{\varepsilon}{\subseteq} \mathcal{D} = (\mathcal{D}')'$, the first claim in Lemma 7.4 (with $\mathcal{A} = \mathcal{C}$ and $\mathcal{B} = \mathcal{D}'$) shows that

$$\|[a, b]\| \leq 2\varepsilon\|a\|\|b\| = 2\varepsilon\|a\|\|b\|$$

for all $a \in \mathcal{C}$ and $b \in \mathcal{D}'$. Since $\mathcal{C}$ is hyperfinite, we can now use the converse in Lemma 7.4 (with $\mathcal{A} = \mathcal{D}' \cap \mathcal{M}$ and $\mathcal{B} = \mathcal{C} \subseteq \mathcal{M}$) to conclude that $\mathcal{D}' \cap \mathcal{M} \overset{2\varepsilon}{\subseteq} \mathcal{C}' \cap \mathcal{M}$. ∎

We now come to a central and nontrivial result. For hyperfinite von Neumann algebras, if $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}$ for sufficiently small ε, there exists a unitary close to the identity that rotates $\mathcal{A}$ into $\mathcal{B}$.

Theorem 7.6 (Near inclusions of subalgebras)**.** *For hyperfinite von Neumann algebras* $\mathcal{A}, \mathcal{B} \subseteq B(\mathcal{H})$ *with* $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}$ *for some* $\varepsilon \leq \frac{1}{64}$*, there exists a unitary* $u \in (\mathcal{A} \cup \mathcal{B})''$ *such that* $u^* \mathcal{A} u \subseteq \mathcal{B}$ *and we have:*

(i) $\|I - u\| \leq 12\varepsilon$.

(ii) *If* $z \in B(\mathcal{H})$ *satisfies* $\|[z, c]\| \leq \delta \|z\| \|c\|$ *for all* $c \in \mathcal{A} \cup \mathcal{B}$*, then* $\|u^* z u - z\| \leq 10\delta \|z\|$*.*

Moreover, if $\mathcal{A}_0 \subseteq \mathcal{A}$ *is an AF C^*-algebra such that* $\mathcal{A}_0'' = \mathcal{A}$*, then u can be chosen such that also:*

(iii) *If* $z \in B(\mathcal{H})$ *satisfies* $z \overset{\delta}{\in} \mathcal{A}_0$ *and* $z \overset{\delta}{\in} \mathcal{B}$*, then* $\|u^* z u - z\| \leq 16\delta \|z\|$*.*

This theorem extends Theorem 4.1 of Christensen [Chr80]. The first item re-states his result, and we develop the remaining claims. A self-contained proof appears in Section 7.2. Similar stability theorems exist for various other classes of C^*-algebras [Chr80, CSS$^+$12]. The stability of subalgebra inclusions is closely related to what is often (especially in the context of groups) referred to as *Ulam stability* [Ula60, BOT13]. There, one is given a map that "almost" satisfies the homomorphism properties, and one asks whether the map can be slightly deformed into a true homomorphism. See for instance [Joh88, Par04] for Ulam stability results on C^*-algebras. The proof of Theorem 7.6 implicitly involves one such Ulam stability property: a completely positive map on a hyperfinite von Neumann algebra that is almost a homomorphism is then deformed to a true homomorphism; see e.g. [Joh88] more generally.

Using related methods, we also obtain the following useful lemma. Here, we control the global error between two homomorphisms using the sum of errors on their local restrictions.

Lemma 7.7. *Consider two injective weak-$*$ continuous unital $*$-homomorphisms*

$$\alpha_1, \alpha_2 \colon \mathcal{A} \to \mathcal{B}$$

between von Neumann algebras $\mathcal{A} \subseteq B(\mathcal{H})$ *and* $\mathcal{B} \subseteq B(\mathcal{K})$*, with hyperfinite von Neumann subalgebras* $\mathcal{A}_1, \ldots, \mathcal{A}_n \subseteq \mathcal{A}$ *that pairwise commute, i.e.,* $[\mathcal{A}_i, \mathcal{A}_j] = 0$ *for* $i \neq j$*, and generate* $\mathcal{A}$ *in the sense that* $(\cup_{i=1}^n \mathcal{A}_i)'' = \mathcal{A}$*. Define*

$$\varepsilon = \sum_{i=1}^n \|(\alpha_1 - \alpha_2)|_{\mathcal{A}_i}\|.$$

Then if $\varepsilon < 1$*,*

$$\|\alpha_1 - \alpha_2\| \leq 2\sqrt{2}\varepsilon \left(1 + (1 - \varepsilon^2)^{\frac{1}{2}}\right)^{-\frac{1}{2}} \leq 2\sqrt{2}\varepsilon,$$

where we note that the expression in the middle is $2\varepsilon + \mathcal{O}(\varepsilon^2)$*.*

The proof appears in Section 7.2. We find the difference between the homomorphisms α_1 and α_2 is controlled by the sum of their local differences. It appears possible that in general, a tighter bound $\|\alpha_1 - \alpha_2\| \le \varepsilon + \mathcal{O}(\varepsilon^2)$ may be correct. An easy example demonstrates the bound is optimal to within a constant factor. Let $\mathscr{A} = \mathscr{A}_1 \otimes \cdots \otimes A_n$ for matrix algebras A_i, and let $\alpha_1(x) = x$ and $\alpha_2(x) = u^* x u$ for $u = u_1 \otimes \cdots \otimes u_n$, choosing any $u_i \in U(\mathscr{A}_i)$ with spectrum $\{1, e^{i\frac{\varepsilon}{n}}\}$, so that $\|u_i - I\| = \frac{\varepsilon}{n} + \mathcal{O}(\varepsilon^2)$ and hence we find $\|u - I\| = \varepsilon + \mathcal{O}(\varepsilon^2)$. Note that by e.g. Theorem 26 of [JKP09], the map on operators $x \mapsto v x v^* - x$ for unitary v has norm given by the diameter of the smallest closed disk containing the spectrum of v. For u_i and u, that diameter is given by $\frac{\varepsilon}{n} + \mathcal{O}(\varepsilon^2)$ and $\varepsilon + \mathcal{O}(\varepsilon^2)$, respectively. Then $\|(\alpha_1 - \alpha_2)|_{\mathscr{A}_i}\| = \frac{\varepsilon}{n} + \mathcal{O}((\frac{\varepsilon}{n})^2)$, while $\|\alpha_1 - \alpha_2\| = \varepsilon + \mathcal{O}(\varepsilon^2)$.

7.2 Proof of Theorem 7.6

In this section, we prove Theorem 7.6 about near inclusions of von Neumann algebras. The result is an extension of Theorem 4.1 of Christensen [Chr80], but we give a self-contained proof. We follow closely the exposition in [Chr80, Chr77a]. Note that in [Chr80] it is assumed that injective von Neumann algebras have a property called D_1. However, whether this is true is unknown, see comments in [PTWW14]. We slightly adapt the arguments of [Chr80] to avoid this issue.

We start by two lemmas which bound commutators $[x, f(y)]$ in terms of commutators $[x, y]$, assuming that y is near the identity, which are relevant to our extension of the result in [Chr80].

In this appendix we bound commutators $[x, f(y)]$ in terms of commutators $[x, y]$, assuming that y is near the identity.

Lemma 7.8 (Commutators with powers)**.** *Let $\mathscr{A}$ be a C^*-algebra and let $y \in \mathscr{A}$ be a normal element with $\|I - y\| \le \varepsilon < 1$. Then, for any $s \in [-1, 1]$ and $x, y \in \mathscr{A}$, we have*

$$\|[x, y^s]\| \le \frac{|s|}{(1 - \varepsilon)^{1-s}} \|[x, y]\|.$$

For fractional powers, y^s is defined using the functional calculus, with branch cut on the negative imaginary axis (away from the spectrum because $\|y - I\| < 1$).

Proof. We assume that $s \notin \{0, 1\}$ since otherwise the claim holds trivially. Let $z = I - y$. The function $t \mapsto (1 - t)^s$ is holomorphic on the open unit disk, so we may expand

$$y^s = (I - z)^s = \sum_{n=0}^{\infty} c_n z^n.$$

The exact form of the coefficients c_n here is unimportant, but note $\operatorname{sgn}(c_n) = -\operatorname{sgn}(s)$ for $n \ge 1$ by our assumption that $s \notin \{0, 1\}$.

$$\|[x, y^s]\| = \|[x, (I - z)^s]\| \le \sum_{n=1}^{\infty} |c_n| \, \|[x, z^n]\| \le -\operatorname{sgn}(s) \sum_{n=1}^{\infty} c_n \, n \, \|z\|^{n-1} \|[x, z]\|$$

$$= -\operatorname{sgn}(s) \frac{d}{dw}(1 - w)^s \Big|_{w = \|z\|} \|[x, y]\| = \frac{|s|}{(1 - \|z\|)^{1-s}} \|[x, y]\| \le \frac{|s|}{(1 - \varepsilon)^{1-s}} \|[x, y]\|$$

as desired. ∎

Lemma 7.9 (Commutators with polar decompositions)**.** *Let $\mathscr{A}$ be a C^*-algebra and $y \in \mathscr{A}$ an element with $\|y - I\| \le \varepsilon \le \frac{1}{8}$. Let $y = u|y|$ be its polar decomposition, with $|y| = (y^* y)^{\frac{1}{2}}$. Then, for any $x \in \mathscr{A}$,*

$$\|[x, u]\| < 3\|[x, y]\| + 2\|[x, y^*]\|.$$

More generally for any $\varepsilon < \sqrt{2} - 1$, an estimate of the above form holds for some choice of constants on the right-hand side depending only on ε.

Proof. Note $\|y - I\| \le \varepsilon < 1$ implies that y is invertible, hence the unitary u in the polar decomposition is uniquely given by $u = y|y|^{-1} = y(y^* y)^{-\frac{1}{2}}$. Moreover, we must have $\|y\| \le 1 + \varepsilon$ and $\|y^* y - I\| \le (2 + \varepsilon)\varepsilon$, which also implies that

$$\|(y^* y)^{-\frac{1}{2}}\| \le (1 - (2 + \varepsilon)\varepsilon)^{-\frac{1}{2}},$$

since $\varepsilon < \sqrt{2} - 1$. We obtain

$$\|[x, u]\| = \|[x, y(y^* y)^{-\frac{1}{2}}]\| \le \|y\|\|[x, (y^* y)^{-\frac{1}{2}}\| + \|(y^* y)^{-\frac{1}{2}}\|\|[x, y]\|$$

$$\le (1 + \varepsilon)\|[x, (y^* y)^{-\frac{1}{2}}\| + \frac{1}{\left(1 - 2\varepsilon - \varepsilon^2\right)^{\frac{1}{2}}}\|[x, y]\|$$

$$\le \frac{1 + \varepsilon}{2(1 - 2\varepsilon - \varepsilon^2)^{\frac{3}{2}}}\|[x, y^* y]\| + \frac{1}{\left(1 - 2\varepsilon - \varepsilon^2\right)^{\frac{1}{2}}}\|[x, y]\|$$

$$\le \frac{(1 + \varepsilon)^2}{2(1 - 2\varepsilon - \varepsilon^2)^{\frac{3}{2}}}\left(\|[x, y^*]\| + \|[x, y]\|\right) + \frac{1}{\left(1 - 2\varepsilon - \varepsilon^2\right)^{\frac{1}{2}}}\|[x, y]\|$$

$$= \frac{(1 + \varepsilon)^2 + 2(1 - 2\varepsilon - \varepsilon^2)}{2(1 - 2\varepsilon - \varepsilon^2)^{\frac{3}{2}}}\|[x, y]\| + \frac{(1 + \varepsilon)^2}{2(1 - 2\varepsilon - \varepsilon^2)^{\frac{3}{2}}}\|[x, y^*]\|.$$

Here we use the above comments to bound the relevant norms, as well as Lemma 7.8 for $s = -\frac{1}{2}$. Using $\varepsilon \le \frac{1}{8}$ this implies the desired bounds. ∎

We continue with Proposition 7.10, which generalizes Proposition 4.2 of Christensen [Chr77a]. There Christensen considers two subalgebras $\mathscr{A}, \mathscr{B} \subseteq B(\mathscr{H})$ that are isomorphic via an isomorphism $\Phi \colon \mathscr{A} \to \mathscr{B}$. Note that Φ is defined only on $\mathscr{A}$, and not on $B(\mathscr{H})$. Roughly speaking, the theorem says that if the isomorphism nearly fixes $\mathscr{A}$, it is inner and implemented by a unitary near the identity. Our Proposition 7.10 below extends this result to the case of multiple commuting subalgebras $\mathscr{A}_i$. Our generalization will be useful for Lemma 7.7. We also extend Christensen's result with the following observation: for elements of $B(\mathscr{H})$ that nearly commute with $\mathscr{A}$ and $\mathscr{B}$, the distance these elements are moved by the inner automorphism is controlled by the size of their commutator with $\mathscr{A}$ and $\mathscr{B}$.

Proposition 7.10 (Making homomorphisms inner)**.** *Consider C^*-algebras $\mathscr{A}_i, \mathscr{B}_i \subseteq B(\mathscr{H})$ for $i = 1, \dots, n$, such that each $\mathscr{A}_i''$ is hyperfinite and $[\mathscr{A}_i, \mathscr{A}_j] = [\mathscr{B}_i, \mathscr{B}_j] = 0$ for $i \ne j$. Consider unital $*$-homomorphisms $\Phi_i \colon \mathscr{A}_i \to \mathscr{B}_i$, with $\|\Phi_i(a_i) - a_i\| \le \gamma_i \|a_i\|$ for all $a_i \in \mathscr{A}_i$ and $i = 1, \dots, n$. Denote $\mathscr{A} = (\cup_{i=1}^n \mathscr{A}_i)''$, $\mathscr{B} = (\cup_{i=1}^n \mathscr{B}_i)''$, and $\varepsilon = \sum_{i=1}^n \gamma_i$. If $\varepsilon < 1$, then there exists a unitary $u \in (\mathscr{A} \cup \mathscr{B})''$ such that $\Phi_i(a_i) = u^* a_i u$ for all i and $a_i \in \mathscr{A}_i$, with*

$$\|I - u\| \le \sqrt{2}\varepsilon(1 + (1 - \varepsilon^2)^{\frac{1}{2}})^{-\frac{1}{2}} \le \sqrt{2}\varepsilon,$$

where we note that the expression in the middle is in fact $\varepsilon + \mathcal{O}(\varepsilon^2)$.

Moreover, for $\varepsilon \leq \frac{1}{8}$, u can be chosen such that for any $z \in B(\mathcal{H})$, if $\|[z,c]\| \leq \delta\|z\|\|c\|$ for all $c \in \mathcal{A} \cup \mathcal{B}$, then $\|u^* z u - z\| \leq 10\delta\|z\|$.

Note $c \in \mathcal{A} \cup \mathcal{B}$ refers to the union of sets, i.e. $c \in \mathcal{A}$ or $c \in \mathcal{B}$. The proof extends the proof of Proposition 4.2 in [Chr77a].

Proof of Proposition 7.10. We will define an element $y \in (\mathcal{A} \cup \mathcal{B})''$ whose polar decomposition yields the desired unitary u. We construct the element y to satisfy the properties $\|I - y\| \leq \sum_{i=1}^{n} \gamma_i$ and $y\Phi_i(u_i) = u_i y$ for all $u_i \in U(\mathcal{A}_i)$ and $i = 1, \ldots, n$.[2]

By Proposition 7.11 further below, each homomorphism $\Phi_i \colon \mathcal{A}_i \to \mathcal{B}_i$ can be extended to a $*$-isomorphism $\Phi_i' \colon \mathcal{A}_i'' \to \Phi_i(\mathcal{A}_i)'' \subseteq \mathcal{B}_i''$. Moreover, $\|\Phi_i'(a_i) - a_i\| \leq \gamma_i\|a_i\|$. Without loss of generality, we may assume $\mathcal{A}_i$ is a hyperfinite von Neumann algebra and Φ_i is a weak-$*$ continuous unital homomorphism (this can always be achieved by replacing $\mathcal{A}_i$ by $\mathcal{A}_i''$, $\mathcal{B}_i$ by $\mathcal{B}_i''$, Φ_i by Φ_i'; the latter is weak-$*$ continuous because it is a $*$-isomorphism of von Neumann algebras).

Consider $B(\mathcal{H} \oplus \mathcal{H})$ with pairwise commuting subalgebras

$$\mathscr{C}_i = \left\{ \begin{pmatrix} a_i & 0 \\ 0 & \Phi_i(a_i) \end{pmatrix} : a_i \in \mathcal{A}_i \right\} \subseteq B(\mathcal{H} \oplus \mathcal{H}).$$

Since Φ_i is weak-$*$ continuous, the map $a_i \mapsto a_i \oplus \Phi_i(a_i)$ is a weak-$*$-continuous unital $*$-homomorphism. Therefore, $\mathscr{C}_i$, which is its image, is a von Neumann algebra is isomorphic to $\mathcal{A}_i$ and hence hyperfinite.

Therefore, by Theorem 7.3 $\mathscr{C}_i$ has property P and for

$$x_0 = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \in B(\mathcal{H} \oplus \mathcal{H})$$

there exists an element $x_1 \in \mathscr{C}_i'$ that is also in the weak operator closure of the convex hull of $\{c_1^* x_0 c_1 : c_1 \in U(\mathscr{C}_1)\}$. Note that unitaries $c_1 \in U(\mathscr{C}_1)$ are of the form

$$\begin{pmatrix} u_1 & 0 \\ 0 & \Phi_1(u_1) \end{pmatrix}$$

for $u_1 \in U(\mathcal{A}_1)$, so elements $c_1^* x_0 c_1$ are of the form

$$\begin{pmatrix} u_1^* & 0 \\ 0 & \Phi_1(u_1^*) \end{pmatrix} \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 & 0 \\ 0 & \Phi_1(u_1) \end{pmatrix} = \begin{pmatrix} 0 & u_1^* \Phi_1(u_1) \\ 0 & 0 \end{pmatrix}.$$

Hence x_1 is of the form

$$x_1 = \begin{pmatrix} 0 & y_1 \\ 0 & 0 \end{pmatrix} \tag{7.7}$$

---

[2]In finite dimension, one can define the element y using

$$y = \int_{U(\mathcal{A}_1)} \mathrm{d}u_1 \cdots \int_{U(\mathcal{A}_n)} \mathrm{d}u_n \, u_n^* \cdots u_1^* \Phi_1(u_1) \cdots \Phi_n(u_n),$$

using the Haar measure of the unitary groups $U(\mathcal{A}_i)$. This y is easily seen to satisfy the aforementioned properties.

for some $y_1 \in (\mathscr{A}_1 \cup \mathscr{B}_1)''$. By direct calculation, $x_1 \in \mathscr{C}_1'$ implies $y_1 \Phi(u_1) = u_1 y_1$ for any unitary $u_1 \in \mathscr{A}_1$, and hence

$$y_1 \Phi_1(a_1) = a_1 y_1$$

for any $a_1 \in \mathscr{A}_1$.

If $n = 1$, we take $y_1 = y$. Otherwise, we repeat the above construction but with x_1 taking the place of x_0, and applying property P of $\mathscr{C}_2$. We obtain $x_2 \in \mathscr{C}_2'$ and associated y_2, with $y_2 \Phi_2(a_2) = a_2 y_2$ for all $a_2 \in \mathscr{A}_2$. Also note $x_2 \in \mathscr{C}_1'$, so $y_2 \Phi_1(a_1) = a_1 y_2$ for all $a_1 \in U(\mathscr{A}_1)$. We continue in this way, until we obtain $y := y_n$, with the property

$$y \Phi_i(a_i) = a_i y \tag{7.8}$$

for all $a_i \in \mathscr{A}_i$ and $i = 1, \dots, n$.

By construction, y_1 is in the weak operator closure of the convex hull of the set

$$\{u_1^* \Phi_1(u_1) : u_1 \in U(\mathscr{A}_1)\},$$

and likewise y_2 is in the weak operator closure of the convex hull of

$$\{u_2^* y_1 \Phi_2(u_2) : u_2 \in U(\mathscr{A}_2)\},$$

and so on. Then y is in the weak operator closure of the convex hull of

$$S := \{u_n^* \dots u_1^* \Phi_1(u_1) \dots \Phi_n(u_n) : u_1 \in U(\mathscr{A}_1), \dots, u_n \in U(\mathscr{A}_n)\}. \tag{7.9}$$

Elements of this form are near the identity,

$$\|I - u_n^* \dots u_1^* \Phi_1(u_1) \dots \Phi_n(u_n)\| \leq \|I - u_n^* \dots u_2^* \Phi_2(u_2) \dots \Phi_n(u_n)\| + \|\Phi_1(u_1) - u_1\|$$

$$\leq \sum_{i=1}^{n} \|\Phi_i(u_i) - u_i\|,$$

and thus, by convexity and lower semicontinuity of the norm in the weak operator topology,

$$\|I - y\| \leq \sum_{i=1}^{n} \gamma_i = \varepsilon.$$

Define $u = y|y|^{-1}$ as the unitary in the polar decomposition of y. By the above estimate, it generally follows (Lemma 2.7 of [Chr75]) that

$$\|u - I\| \leq \sqrt{2}\varepsilon (1 + (1 - \varepsilon^2)^{\frac{1}{2}})^{-\frac{1}{2}} \leq \sqrt{2}\varepsilon.$$

We now show

$$u^* a_i u = \Phi_i(a_i)$$

for all $a_i \in \mathscr{A}_i$ and $i = 1, \dots, n$. To see this, first note that (7.8) implies

$$y^* y = \Phi_i(u_i)^* y^* y \Phi(u_i)$$

for any $u_i \in U(\mathscr{A}_i)$, so that $[\Phi_i(u_i), y^* y] = 0$. Then, since any $a_i \in \mathscr{A}_i$ can be written as a linear combination of unitary elements, $[\Phi_i(a_i), y^* y] = 0$, hence $[\Phi_i(a_i), |y|^{-1}] = 0$ and

$$u^* a_i u = |y|^{-1} y^* a_i y |y|^{-1} = |y|^{-1} y^* y \Phi_i(a_i) |y|^{-1} = |y|^{-1} y^* y |y|^{-1} \Phi_i(a_i) = \Phi(a_i)$$

where we first used (7.8) and then that $[\Phi_i(a_i), |y|^{-1}] = 0$.

Finally, we show the last claim of the theorem. Consider any $z \in B(\mathscr{H})$ with the property that $\|[z, c]\| \le \delta \|z\| \|c\|$ for all $c \in \mathscr{A} \cup \mathscr{B}$. Then, $\|[z, s]\| \le 2\delta \|z\|$ for any $s \in S$, since any element of S is a product of a unitary in $U(\mathscr{A})$ and a unitary in $U(\mathscr{B})$, and likewise $\|[z, s^*]\| \le 2\delta \|z\|$. We find that

$$\|[z, y]\| \le 2\delta \|z\|,$$

using that y is in the weak operator closure of the convex hull of S as defined in (7.9). To see this, let y_i be a net of elements in the convex hull of S that converges to y in the weak operator topology. Since the elements in S have norm at most one, by convexity it holds that $\|y_i\| \le 1$ as well and hence $\|[z, y_i]\| \le 2\delta \|z\|$. The norm is lower semicontinuous in the weak operator topology which implies that $\|[z, y]\| \le \liminf_i \|[z, y_i]\| \le 2\delta \|z\|$. The above reasoning holds for y^* as well. Then we can apply Lemma 7.9, where we use that $\|I - y\| \le \varepsilon \le \frac{1}{8}$. We find

$$\|u^* z u - z\| = \|[z, u]\| \le 3\|[z, y]\| + 2\|[z, y^*]\| \le 10\delta \|z\|$$

as desired. $\blacksquare$

The above proof is completed by the technical proposition below. The proof follows from the proof of Theorem 5.4 in [Chr77a].

Proposition 7.11. *Let $\mathscr{A} \subseteq B(\mathscr{H})$ be C^*-algebra and let $\Phi: \mathscr{A} \to B(\mathscr{H})$ be a unital $*$-homomorphism with $\|\Phi(a) - a\| \le \varepsilon \|a\|$ for all $a \in \mathscr{A}$ and some $\varepsilon < 1$, then Φ can be extended to a $*$-isomorphism $\Phi': \mathscr{A}'' \to \Phi(\mathscr{A})''$ with $\|\Phi(a) - a\| \le \varepsilon \|a\|$ for all $a \in \mathscr{A}''$.*

Proof. To extend Φ, consider $B(\mathscr{H} \oplus \mathscr{H})$ with subalgebra

$$\mathscr{C} = \left\{ \begin{pmatrix} a & 0 \\ 0 & \Phi(a) \end{pmatrix} : a \in \mathscr{A} \right\} \subseteq B(\mathscr{H} \oplus \mathscr{H}).$$

We first show that for any $a \in \mathscr{A}''$, there exists unique $b \in \Phi(A)''$ such that

$$c = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathscr{C}''.$$

For $a \in \mathscr{A}''$, by Kaplansky's density theorem, there exists a net $\{a_i\}$ in $\mathscr{A}$ converging in the strong and hence in the weak operator topology to a, with $\|a_i\| \le \|a\|$. Then we have $\|\Phi(a_i) - a_i\| \le \varepsilon \|a\|$, and $\|\Phi(a_i)\| \le (1 + \varepsilon)\|a\|$, so we can define a net

$$c_i = \begin{pmatrix} a_i & 0 \\ 0 & \Phi(a_i) \end{pmatrix}$$

within a ball of finite radius in $B(\mathscr{H})$. Since such balls are compact in the weak operator topology, c_i must have a convergent subnet, which then converges to some

$$c = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathscr{C}'',$$

as claimed. To see the uniqueness of b given a, suppose otherwise that there exist corresponding $b_1, b_2 \in \Phi(A)''$ with $(a, b_1), (a, b_2) \in \mathscr{C}''$, so that $z = b_1 - b_2 \in \Phi(\mathscr{A})''$ with

$$c = \begin{pmatrix} 0 & 0 \\ 0 & z \end{pmatrix} \in \mathscr{C}''.$$

By Kaplansky's density theorem, there exists a net $\{c_i\}$ in $\mathscr{C}$ converging strongly to c with $\|c_i\| \leq \|c\| = \|z\|$. Write $c_i = (a_i, \Phi(a_i))$ for $a_i \in \mathscr{A}$. Then $\|a_i\| \leq \|z\|$, $\{a_i\}$ converges strongly to zero, and $\{\Phi(a_i)\}$ converges strongly to z, so $\Phi(a_i) - a_i$ converges strongly to z and hence also weakly. By the lower semicontinuity of the norm for the weak operator topology, $\|z\| \leq \liminf_i \|\Phi(a_i) - a_i\| \leq \varepsilon \|z\|$, so that $\|z\| = 0$ and $b_1 = b_2$, demonstrating uniqueness.

A similar argument shows that for any $b \in \Phi(\mathscr{A})''$, there exists unique $a \in \mathscr{A}''$ such that

$$c = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathscr{C}''.$$

The above maps $a \mapsto b$ and $b \mapsto a$ define a bijection $\Phi' : \mathscr{A}'' \to \Phi(\mathscr{A})''$. The linearity, multiplicativity, and $*$-property of Φ' follow from the above uniqueness. Thus Φ' is a $*$-isomorphism. Finally, we show $\|\Phi'(a) - a\| \leq \|a\|\varepsilon$ for all $a \in \mathscr{A}''$. By Kaplansky's density theorem, there exists a net $\{a_i\}$ strongly converging to a for $a_i \in \mathscr{A}$ with $\|a_i\| \leq \|a\|$. By the above constructions, there exists a subnet such that $\Phi(a_i)$ converges in the weak operator topology to $\Phi'(a)$. Then, again by the lower semicontinuity of the norm, we find that $\|\Phi'(a) - a\| \leq \liminf_i \|\Phi(a_i) - a_i\| \leq \varepsilon \|a\|$, as desired. ∎

Now we turn to Theorem 7.6. In Theorem 4.1 of [Chr80], Christensen proves that if a subalgebra $\mathscr{A}$ is approximately contained in another subalgebra $\mathscr{B}$ then there exists a unitary near the identity that rotates $\mathscr{A}$ into $\mathscr{B}$. Our Theorem 7.6 extends his result with the following observations. First, elements of $B(\mathscr{H})$ already close to both $\mathscr{A}$ and $\mathscr{B}$ are not moved much by the automorphism. Second, elements that nearly commute with both $\mathscr{A}$ and $\mathscr{B}$ are are not moved much either. Thus the automorphism "does no more than it needs."

For convenience, we recall the notion of near inclusions in Definition 7.2. We write $a \overset{\varepsilon}{\in} \mathscr{B}$ when there exists $b \in \mathscr{B}$ such that

$$\|a - b\| \leq \varepsilon \|a\|,$$

and we write $\mathscr{A} \overset{\varepsilon}{\subseteq} \mathscr{B}$ when $a \overset{\varepsilon}{\in} \mathscr{B}$ for all $a \in \mathscr{A}$. Also recall the notion of hyperfinite von Neumann algebras, reviewed in Section 7.0.2. Then we are equipped to state Theorem 7.6, repeated below.

Theorem 7.6 (Near inclusions of subalgebras)**.** *For hyperfinite von Neumann algebras $\mathscr{A}, \mathscr{B} \subseteq B(\mathscr{H})$ with $\mathscr{A} \overset{\varepsilon}{\subseteq} \mathscr{B}$ for some $\varepsilon \leq \frac{1}{64}$, there exists a unitary $u \in (\mathscr{A} \cup \mathscr{B})''$ such that $u^* \mathscr{A} u \subseteq \mathscr{B}$ and we have:*

(i) $\|I - u\| \leq 12\varepsilon$.

(ii) If $z \in B(\mathscr{H})$ satisfies $\|[z, c]\| \leq \delta \|z\| \|c\|$ for all $c \in \mathscr{A} \cup \mathscr{B}$, then $\|u^* z u - z\| \leq 10\delta \|z\|$.

Moreover, if $\mathcal{A}_0 \subseteq \mathcal{A}$ is an AF C^*-algebra such that $\mathcal{A}_0'' = \mathcal{A}$, then u can be chosen such that also:

(iii) If $z \in B(\mathcal{H})$ satisfies $z \overset{\delta}{\in} \mathcal{A}_0$ and $z \overset{\delta}{\in} \mathcal{B}$, then $\|u^* z u - z\| \le 16\delta\|z\|$.

The first item re-states Theorem 4.1 of Christensen [Chr80], or specifically part (b) of his Corollary 4.2 (noting that hyperfinite algebras are injective). The remaining items constitute our extension.

Now we proceed with the proof of Theorem 7.6, closely following and elaborating on some technical details and then extending the proof of Theorem 4.1 in [Chr80].

Proof of Theorem 7.6. By Theorem 7.3, since $\mathcal{B}$ is hyperfinite, it is injective and hence there exists a conditional expectation

$$\mathbb{E}_\mathcal{B} : B(\mathcal{H}) \to \mathcal{B} \subseteq B(\mathcal{H}).$$

This map is completely positive and unital, and thus it has a Stinespring dilation [Sti55]. That is, there exists a Hilbert space $\mathcal{K}$, a unital $*$-homomorphism $\pi\colon B(\mathcal{H}) \to B(\mathcal{K})$, and an isometry $v\colon \mathcal{H} \to \mathcal{K}$ such that

$$\mathbb{E}_\mathcal{B}(x) = v^* \pi(x) v \quad \forall x \in B(\mathcal{H}). \tag{7.10}$$

Let $p = vv^* \in B(\mathcal{K})$ be the projection onto the image of v. Then $p \in \pi(\mathcal{B})'$, since $\mathbb{E}_\mathcal{B}$ restricted to $\mathcal{B}$ is an isomorphism.[34]

Next we show that p nearly commutes with $\pi(\mathcal{A})$ as well. For any $a \in \mathcal{A}$, we choose $b \in \mathcal{B}$ with $\|a - b\| \le \varepsilon\|a\|$, using $\mathcal{A} \overset{\varepsilon}{\subseteq} \mathcal{B}$. Then,

$$\|[\pi(a), p]\| = \frac{1}{2}\|[\pi(a-b), 2p-I]\| \le \|\pi(a-b)\|\,\|2p-I\| \le \varepsilon\|a\|, \tag{7.11}$$

noting that for any projection, $\|2p - I\| = 1$.

Note that although $\mathcal{A}$ itself is hyperfinite, $\pi(\mathcal{A})''$ is not immediately guaranteed hyperfinite, because π is not guaranteed weak-$*$ continuous. On the other hand, since $\mathcal{A}$ is hyperfinite, there exists an AF C^*-algebra $\mathcal{A}_0 \subseteq \mathcal{A}$ such that $\mathcal{A}_0'' = \mathcal{A}$, as in the theorem statement. Then $\pi(\mathcal{A}_0)$ is also AF, and $\pi(\mathcal{A}_0)''$ is hyperfinite.

Because $\pi(\mathcal{A}_0)''$ hyperfinite, it satisfies property P, so there exists $\tilde{p}$ in the weak operator closure of the convex hull of $\{upu^* : u \in U(\pi(\mathcal{A}_0)'')\}$ such that $\tilde{p} \in \pi(\mathcal{A}_0)'$. By

---

[3]In more detail, to see $p \in \pi(\mathcal{B})'$, first note $\pi(\mathcal{B}) \to B(\mathcal{K}), \pi(b) \mapsto p\pi(b)p$ is a $*$-homomorphism. Then note the following general fact: for any algebra $\mathcal{A} \subset B(\mathcal{H})$ and projection $p \in B(\mathcal{H})$, if the map $f(a) = pap$ is a $*$-homomorphism, then $p \in \mathcal{A}'$. To see this, note for any $a \in \mathcal{A}$, $f(a^* a) = f(a^*)f(a) = pa^* pap$ and $f(a^* a) = pa^* ap = pa^*(p + p^\perp)ap = pa^* pap + pa^* p^\perp ap$, so the difference yields $0 = pa^* p^\perp ap = (p^\perp ap)^*(p^\perp ap)$, so $p^\perp ap = 0$. The same is true for a^*, so $pap^\perp = 0$ also. Then $[p, a] = (p + p^\perp)[p, a](p + p^\perp) = pap^\perp - p^\perp ap = 0$.

[4]It may be helpful to understand the Stinespring dilation explicitly in finite dimensions where $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{B} = I_A \otimes B(\mathcal{H}_B)$, with commutant $\mathcal{B}' = \mathcal{A} = B(\mathcal{H}_A) \otimes I_B$. Then the conditional expectation is the normalized partial trace $\mathbb{E}_\mathcal{B}(x) = \frac{1}{d_A}\mathrm{tr}_\mathcal{A}(x)$. For a minimal Stinespring dilation we can take the Hilbert space $\mathcal{K} = \mathcal{H}_A^1 \otimes H_A^2 \otimes \mathcal{H}_A^3 \otimes \mathcal{H}_B$, where the $\mathcal{H}_A^i$ are three copies of the Hilbert space $\mathcal{H}_A$. We define $\pi\colon B(\mathcal{H}) \to B(\mathcal{K})$ by identifying operators on $\mathcal{H}$ with operators on $\mathcal{H}_A^1 \otimes \mathcal{H}_B$. Finally, we take the isometry v as adding a maximally entangled state on $\mathcal{H}_A^1 \otimes \mathcal{H}_A^2$. Note that the projection p onto the image of v commutes with $\pi(\mathcal{B}) = I_{A^1 A^2 A^3} \otimes B(\mathcal{H}_B)$.

convexity of the norm and lower semicontinuity of the norm with respect to the weak operator topology,

$$\|\tilde{p} - p\| \le \sup_{u \in U(\pi(\mathscr{A}_0)'')} \|\tilde{p} - upu^*\| = \sup_{u \in U(\pi(\mathscr{A}_0)'')} \|[u, p]\| \le \varepsilon.$$

The final inequality follows from Eq. (7.11) in the following way. First we observe that $\|[u, p]\| \le \varepsilon$ for any $u \in \pi(U(\mathscr{A}))$. By the lower semicontinuity of the norm in the weak operator topology, this estimate extends directly to the weak operator closure of $\pi(U(\mathscr{A}))$. Accordingly, it suffices to show that any $u \in U(\pi(\mathscr{A}_0)'')$ is contained in the weak operator closure of $\pi(U(\mathscr{A}))$. This can be seen as follows. By Theorem 5.2.5 in [KR97] there exists self-adjoint $y \in \pi(\mathscr{A}_0)''$ such that $u = e^{iy}$. By a version of the Kaplansky density theorem, there exists a net $\{y_n\}$ of self-adjoint elements $y_n \in \pi(\mathscr{A}_0)$ converging strongly to y, with $\|y_n\| \le \|y\|$. Then $\{e^{iy_n}\}$ is a net of elements in $\pi(U(\mathscr{A}_0))$, since we can always write $y_n = \pi(x_n)$ with self-adjoint $x_n \in \mathscr{A}_0$, hence $e^{iy_n} = e^{i\pi(x_n)} = \pi(e^{ix_n})$ and e^{ix_n} is unitary. On the other hand, by Proposition 5.3.2 in [KR97], $\{e^{iy_n}\}$ converges strongly to $e^{iy} = u$. We conclude that u is in the strong (and hence in the weak) operator closure of $\pi(U(\mathscr{A}_0))$, hence in particular of $\pi(U(\mathscr{A}))$. Note that, by construction, $\|\tilde{p}\| \le \|p\| = 1$.

Next we would like to project $\tilde{p}$ onto $(\pi(B(\mathscr{H})) \cup \{p\})''$, the von Neumann algebra generated by $\pi(B(\mathscr{H}))$ and the projection p inside $B(\mathscr{K})$.[5] By Corollary 1.3.2 in [Arv69], $(\pi(B(\mathscr{H})) \cup \{p\})'$ is isomorphic to $\mathscr{B}'$. Because the commutant of an injective von Neumann algebra is injective, and $\mathscr{B}$ is injective, $\mathscr{B}'$ is also injective, hence also $(\pi(B(\mathscr{H})) \cup \{p\})'$ and $(\pi(B(\mathscr{H})) \cup \{p\})''$. Thus we can use a conditional expectation to define

$$x = \mathbb{E}_{(\pi(B(\mathscr{H})) \cup \{p\})''}(\tilde{p}) \in (\pi(B(\mathscr{H})) \cup \{p\})'', \quad \|x - p\| \le \varepsilon,$$

where the norm bound follows from $\|\tilde{p} - p\| \le \varepsilon$, because the conditional expectation is a contraction. Moreover, it holds that $x \in \pi(\mathscr{A}_0)'$. To see this, compute $[x, z] = 0$ for $z \in \pi(\mathscr{A}_0)$, using that $\tilde{p} \in \pi(\mathscr{A}_0)'$ and $\pi(\mathscr{A}_0) \subseteq (\pi(B(\mathscr{H})) \cup \{p\})''$, and the general property of conditional expectations that $\mathbb{E}_{\mathscr{Z}}(z_1 y z_2) = z_1 \mathbb{E}_{\mathscr{Z}}(y) z_2$ for $z_1, z_2 \in \mathscr{Z}$.

The next steps follow Lemma 3.3 of [Chr77a]. Note that because x is self-adjoint, we have $\|x - p\| \le \varepsilon$, and $\|x\| \le 1$ (as conditional expectations are contractions), its spectrum is in $[-\varepsilon, \varepsilon] \cup [1 - \varepsilon, 1]$. Define the projection $q \in \pi(\mathscr{A}_0)'$ as the spectral projection of x corresponding to the part of the spectrum in $[1 - \varepsilon, 1]$. Then, $\|q - x\| \le \varepsilon$ and $\|q - p\| \le 2\varepsilon$.

Using the projection $p \in \pi(\mathscr{B})'$ and the nearby projection $q \in \pi(\mathscr{A}_0)'$, define

$$y = qp + q^\perp p^\perp,$$

where $p^\perp = (I - p)$ denotes the projection onto the orthogonal complement. Then

$$\|y - I\| = \|(2q - I)(p - q)\| \le \|p - q\| \le 2\varepsilon. \tag{7.12}$$

In particular, y is invertible. Now consider the unitary $w = y|y|^{-1}$ from the polar decomposition $y = w|y|$. Because y is near the identity, w must be as well. Namely, by Lemma 2.7 of [Chr75], we find

$$\|w - I\| \le 2\sqrt{2}\varepsilon. \tag{7.13}$$

---

[5]In the finite-dimensional setting of Footnote 4, where again $\mathscr{K} = \mathscr{H}_A^1 \otimes \mathscr{H}_A^2 \otimes \mathscr{H}_A^3 \otimes \mathscr{H}_B$, we have $(\pi(B(\mathscr{H})) \cup \{p\})' = B(\mathscr{H}_A^3)$ and hence $(\pi(B(\mathscr{H})) \cup \{p\})'' = B(\mathscr{H}_A^1 \otimes \mathscr{H}_A^2 \otimes \mathscr{H}_B)$.

Since $y^*y = pqp + p^\perp q^\perp p^\perp$, we have $[p, y^*y] = 0$ and hence $[p, |y|^{-1}] = 0$. Moreover, $yp = qy$, so

$$wpw^* = y|y|^{-1}p|y|^{-1}y^* = yp|y|^{-1}|y|^{-1}y^* = qy|y|^{-1}|y|^{-1}y^* = q. \qquad (7.14)$$

With the unitary $w \in (p \cup \pi(B(\mathcal{H})))''$, we can finally define the homomorphism to which we soon apply Proposition 7.10. Let

$$\Phi \colon \mathscr{A}_0 \to \mathscr{B} \subseteq B(\mathcal{H}), \quad \Phi(a) = v^*w^*\pi(a)wv.$$

This is a unital $*$-homomorphism, since it clearly preserves the $*$-operation and we have for all $a_1, a_2 \in \mathscr{A}_0$ that

$$\Phi(a_1)\Phi(a_2) = v^*w^*\pi(a_1)wpw^*\pi(a_2)wv = v^*w^*\pi(a_1)q\pi(a_2)wv$$
$$= v^*w^*\pi(a_1a_2)qwv = v^*w^*\pi(a_1a_2)wpv = \Phi(a_1a_2),$$

using (7.14), $q \in \pi(\mathscr{A}_0)'$, $p = vv^*$, and that v is an isometry. To see that its image lies in $\mathscr{B}$, note that $w^*\pi(a)w \in (p \cup \pi(B(\mathcal{H})))''$ and recall the original construction of the Stinespring dilation in (7.10). Moreover, for any $a \in \mathscr{A}_0$, by assumption there exists $b \in \mathscr{B}$ with $\|b - a\| \le \varepsilon\|a\|$, so that

$$\begin{aligned}
\|\Phi(a) - a\| &\le \|\Phi(a) - b\| + \|b - a\| \\
&= \|v^*\big(w^*\pi(a)w - \pi(b)\big)v\| + \|b - a\| \\
&\le \|w^*\pi(a)w - \pi(b)\| + \|b - a\| \\
&\le \|w^*\pi(a)w - \pi(a)\| + 2\|b - a\| \\
&= \|[\pi(a), w]\| + 2\|b - a\| \qquad\qquad\qquad (7.15) \\
&= \|[\pi(a), w - I]\| + 2\|b - a\| \\
&\le 2\|w - I\|\|a\| + 2\|b - a\| \le 8\varepsilon\|a\|.
\end{aligned}$$

using $b = \mathbb{E}_{\mathscr{B}}(b) = v^*\pi(b)v$ in the second step and (7.13) in the last step.

We can thus apply Proposition 7.10 (for $n = 1$) to obtain a unitary $u \in (\mathscr{A} \cup \mathscr{B})''$ which is such that $\|u - I\| \le \sqrt{2} \cdot 8\varepsilon \le 12\varepsilon$ such that $u^*au = \Phi(a) \in \mathscr{B}$ for all $a \in \mathscr{A}_0$, and we extend the map $\Phi \colon \mathscr{A} \to \mathscr{B}$ by $\Phi(a) = u^*au$ for $a \in \mathscr{A} = \mathscr{A}_0''$. Moreover, by Proposition 7.10, we are already ensured the desired property of Theorem 7.6 that if $z \in B(\mathcal{H})$ satisfies $\|[z, c]\| \le \delta\|z\|\|c\|$ for all $c \in \mathscr{A} \cup \mathscr{B}$, then $\|uzu^* - z\| \le 10\delta\|z\|$.

Finally, we need to show the additional property that for any $z \in B(\mathcal{H})$ with $z \overset{\delta}{\in} \mathscr{A}_0$ and $z \overset{\delta}{\in} \mathscr{B}$, we have $\|u^*zu - z\| \le 16\delta\|z\|$. First take $a \in \mathscr{A}_0$ with $\|z - a\| \le \delta\|z\|$. Then,

$$\|u^*zu - z\| = \|u^*(z - a)u - (z - a) + u^*au - a\| \le 2\delta\|z\| + \|\Phi(a) - a\|.$$

Now take $b \in \mathscr{B}$ with $\|z - b\| \le \delta\|z\|$, hence also $\|a - b\| \le 2\delta\|z\|$. Then we can bound just like in (7.15) to obtain (note that $a \in \mathscr{A}_0$)

$$\begin{aligned}
\|\Phi(a) - a\| &\le \|[\pi(a), w]\| + 2\|a - b\| \\
&\le \|[\pi(a), w]\| + 4\delta\|z\| \\
&\le 3\|[\pi(a), y]\| + 2\|[\pi(a), y^*]\| + 4\delta\|z\|,
\end{aligned}$$

where we used Lemma 7.9 in the last line, noting that $\|y - I\| \le 2\varepsilon \le \frac{1}{8}$ by Eq. (7.12) and our assumption on ε. To bound the commutators' norms, recall that $p \in \pi(\mathcal{B})'$ and $q \in \pi(\mathcal{A}_0)'$. Hence,

$$\|[\pi(a), y]\| = \|[\pi(a), qp + q^\perp p^\perp]\| = \|(2q - I)[\pi(a), p]\|$$

$$\le \|[\pi(a), p]\| = \|[\pi(a - b), p]\| = \frac{1}{2}\|[\pi(a - b), 2p - I]\| \le \|a - b\| \le 2\delta\|z\|,$$

and likewise for $[\pi(b), y^*]$. Therefore, $\|\Phi(a) - a\| \le 14\delta\|z\|$, and hence

$$\|u^* z u - z\| \le 16\delta\|z\|$$

as desired.   ∎

As another application of Proposition 7.10, Lemma 7.7 controls the distance between homomorphisms using the distance between their local restrictions. We repeat the statement for convenience.

Lemma 7.7. *Consider two injective weak-$*$ continuous unital $*$-homomorphisms*

$$\alpha_1, \alpha_2 : \mathcal{A} \to \mathcal{B}$$

between von Neumann algebras $\mathcal{A} \subseteq B(\mathcal{H})$ and $\mathcal{B} \subseteq B(\mathcal{K})$, with hyperfinite von Neumann subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_n \subseteq \mathcal{A}$ that pairwise commute, i.e., $[\mathcal{A}_i, \mathcal{A}_j] = 0$ for $i \ne j$, and generate $\mathcal{A}$ in the sense that $(\cup_{i=1}^n \mathcal{A}_i)'' = \mathcal{A}$. Define

$$\varepsilon = \sum_{i=1}^n \|(\alpha_1 - \alpha_2)|_{\mathcal{A}_i}\|.$$

Then if $\varepsilon < 1$,

$$\|\alpha_1 - \alpha_2\| \le 2\sqrt{2}\varepsilon \left(1 + (1 - \varepsilon^2)^{\frac{1}{2}}\right)^{-\frac{1}{2}} \le 2\sqrt{2}\varepsilon,$$

where we note that the expression in the middle is $2\varepsilon + \mathcal{O}(\varepsilon^2)$.

Proof. Since we assume the α_i to be weak-$*$ continuous, $\alpha_1(\mathcal{A}_i)$ and $\alpha_2(\mathcal{A}_i)$ are von Neumann algebras which are isomorphic to $\mathcal{A}_i$ (and in particular are hyperfinite). Define $*$-isomorphisms Φ_i between $\alpha_1(\mathcal{A}_i)$ and $\alpha_2(\mathcal{A}_i)$, given by $\alpha_1(a_i) \mapsto \alpha_2(a_i)$ for $a_i \in \mathcal{A}_i$. Then we apply Proposition 7.10 (with $\alpha_1(\mathcal{A}_i)$ as $\mathcal{A}_i$, $\alpha_2(\mathcal{A}_i)$ as $\mathcal{B}_i$, and where we let $\gamma_i = \|(\alpha_1 - \alpha_2)|_{\mathcal{A}_i}\|$) to find a unitary $u \in \mathcal{B}$ such that $\alpha_2(a_i) = u^* \alpha_1(a_i) u$ for all $a_i \in \mathcal{A}_i$ for $i = 1, \ldots n$ with

$$\|I - u\| \le \sqrt{2}\varepsilon \left(1 + (1 - \varepsilon^2)^{\frac{1}{2}}\right)^{-\frac{1}{2}}.$$

This implies that $\alpha_2(a) = u^* \alpha_1(a) u$ for all $a \in \mathcal{A}$, and hence

$$\|\alpha_1 - \alpha_2\| \le 2\sqrt{2}\varepsilon \left(1 + (1 - \varepsilon^2)^{\frac{1}{2}}\right)^{-\frac{1}{2}}.$$

                                                                                     ∎

Finally, we mention another result about simultaneous near inclusions. If several mutually commuting subalgebras $\mathscr{A}_i$ each nearly include into $\mathscr{B}$, then so does the algebra they generate. We use this lemma to prove Lemma 9.3 which shows that Lieb-Robinson type bounds for single site operators imply Lieb-Robinson bounds for operators supported on arbitrary sets.

Lemma 7.12 (Simultaneous near inclusions)**.** *Let $\mathscr{A}_i \subseteq B(\mathscr{H})$ for $i = 1, ..., n$ and let $\mathscr{B} \subseteq B(\mathscr{H})$ be von Neumann algebras, where the $\mathscr{A}_i$ are hyperfinite and $[\mathscr{A}_i, \mathscr{A}_j] = 0$ for $i \neq j$. If $\mathscr{A}_i \overset{\varepsilon_i}{\subseteq} \mathscr{B}$ for each i, then for $\varepsilon := \sum_{i=1}^{n} \varepsilon_i$ we have*

$$\mathscr{B}' \overset{2\varepsilon}{\subseteq} (\cup_i \mathscr{A}_i)'. \tag{7.16}$$

If additionally $\mathscr{A}_i \subseteq \mathscr{M}$ for some von Neumann algebra $\mathscr{M} \subseteq B(\mathscr{H})$ for $i = 1, \ldots, n$, then

$$\mathscr{B}' \cap \mathscr{M} \overset{2\varepsilon}{\subseteq} (\cup_i \mathscr{A}_i)' \cap \mathscr{M}. \tag{7.17}$$

Finally, if $\mathscr{B}'$ is hyperfinite then

$$(\cup_i \mathscr{A}_i)'' \overset{4\varepsilon}{\subseteq} \mathscr{B}.$$

Proof. First we show $\mathscr{B}'$ nearly includes into $(\cup_i \mathscr{A}_i)'$. By hyperfiniteness (and therefore property P) of $\mathscr{A}_1$, for each $b_0' \in \mathscr{B}'$ there exists $b_1' \in \mathscr{A}_1'$ in the weak operator closure of the convex hull of $\{u_1^* b_0' u_1 \,:\, u_1 \in U(\mathscr{A}_1)\}$. Then by property P of $\mathscr{A}_2$, there exists $b_2' \in \mathscr{A}_2'$ in the weak operator closure of the convex hull of $\{u_2^* b_1' u_2 \,:\, u \in U(\mathscr{A}_2)\}$. Note that $b_2' \in \mathscr{A}_1'$ still, using $[\mathscr{A}_1, \mathscr{A}_2] = 0$. We continue in this way until we find b_n' in the weak operator closure of the convex hull of[6]

$$\{u_n^* \cdots u_1^* b_0' u_1 \cdots u_n \,:\, u_1 \in U(\mathscr{A}_1), \ldots, u_n \in U(\mathscr{A}_n)\}.$$

Note $\|[u_i, b_0']\| \leq 2\varepsilon_i \|b_0'\|$, by $\mathscr{A}_i \overset{\varepsilon_i}{\subseteq} \mathscr{B}$ and Lemma 7.4. Thus, elements in the above set are near b_0', since by a telescoping sum

$$\|b_0' - u_n^* \cdots u_1^* b_0' u_1 \cdots u_n\| \leq \sum_{i=1}^{n} \|u_n^* \cdots u_{i+1}^* b_0' u_{i+1} \cdots u_n - u_n^* \cdots u_i^* b_0' u_i \cdots u_n\|$$

$$= \sum_{i=1}^{n} \|b_0' - u_i^* b_0' u_i\| = \sum_{i=1}^{n} \|[u_i, b_0']\| \leq 2\varepsilon \|b_0'\|$$

and hence, using the convexity of the norm and its lower semicontinuity with respect to the weak operator topology,

$$\|b_0' - b_n'\| \leq 2\varepsilon \|b_0'\|.$$

By construction, $b_n' \in \mathscr{A}_i'$ for each i, so $b_n' \in (\cup_i \mathscr{A}_i)'$. The above construction held for any $b_0' \in \mathscr{B}'$, so Eq. (7.16) follows. Note that if we assume that each $\mathscr{A}_i \subseteq \mathscr{M}$ and we take $b_0' \in \mathscr{B}' \cap \mathscr{M}$, then also $b_n' \in \mathscr{M}$, which shows Eq. (7.17). Finally, by Lemma 7.5 and the assumption that $\mathscr{B}'$ is hyperfinite we conclude that $(\cup_i \mathscr{A}_i)'' \overset{4\varepsilon}{\subseteq} \mathscr{B}$.   ∎

---

[6]In the finite-dimensional case, we could immediately define

$$b_n' = \int_{U(\mathscr{A}_1)} du_1 \cdots \int_{U(\mathscr{A}_n)} du_n \ u_n^* \cdots u_1^* b_0' u_1 \cdots u_n$$

using the Haar integral, rather than make use of property P.

Index theory for quantum cellular automata

In this section we first give a careful definition of QCAs. Next, we will discuss the index theory of one-dimensional QCAs, following [GNVW12]. We prove some useful properties on the robustness of this index, and we provide an alternative expression for the index in terms of the mutual information. We start by giving a formal definition of the notion of a QCA in Definition 8.1. In Section 8.1 we review the definition and some of the most important properties of the GNVW index as proven in [GNVW12]. In Section 8.2 we prove new results about the GNVW index: in Section 8.2.1 we provide an alternative formula for the index in terms of a difference of mutual informations. In Section 8.2.2 we prove some results about QCAs in one dimension which are locally close to each other. These results are interesting in their own right, but will also be crucial when extending the index to ALPUs.

We define a QCA as an automorphism of the quasi-local algebra, with a notion of strict locality. To this end, consider a spin system on a lattice Γ with some metric d and the associated quasi-local algebra $\mathscr{A}_\Gamma$. If $X \subseteq \Gamma$ we will denote by

$$B(X, r) = \{n \in \Gamma \text{ such that } d(n, X) \leq r\}$$

the set of sites within distance r of X.

Definition 8.1 (QCA)**.** A *quantum cellular automaton* (QCA) *with radius R* is an automorphism $\alpha \colon \mathscr{A}_\Gamma \to \mathscr{A}_\Gamma$ such that if x is an operator supported on a finite subset $X \subseteq \Gamma$, then $\alpha(x)$ is supported on $B(X, R)$. We call R the *radius* of the QCA.

In this dissertation we will only be concerned with QCAs on a one-dimensional lattice, which we simply take to be $\Gamma = \mathbb{Z}$. In Chapter 9 we will generalize this notion to an approximately local version (an approximately locality preserving unitary, or ALPU).

8.1 One-dimensional QCAs and the GNVW index

One-dimensional QCAs have a beautifully simple structure theory, which we will now review. The material in this section is based on [GNVW12], which we recommend for a more extensive discussion. The same material is also covered in the review [Far20].

Suppose that α is a nearest-neighbour QCA, which we may assume without loss of generality after blocking sites. Let

$$\begin{aligned}
\mathcal{B}_n &= \mathcal{A}_{\{2n,2n+1\}} \\
\mathcal{C}_n &= \mathcal{A}_{\{2n-1,2n\}}
\end{aligned} \tag{8.1}$$

be algebras on pairs of adjacent sites; with $\mathcal{B}_n$ and $\mathcal{C}_n$ corresponding to pairs staggered by one. In particular, $\alpha(\mathcal{B}_n) \subseteq \mathcal{C}_n \otimes \mathcal{C}_{n+1}$. Define

$$\begin{aligned}
\mathcal{L}_n &= \alpha(\mathcal{B}_n) \cap \mathcal{C}_n \\
\mathcal{R}_n &= \alpha(\mathcal{B}_n) \cap \mathcal{C}_{n+1}.
\end{aligned} \tag{8.2}$$

See Fig. 8.1 as a mnemonic. These are manifestly algebras, but naively they might be trivial. Instead, it turns out that they provide factorizations of $\mathcal{C}_n$ and $\mathcal{B}_n$. Using the notation $\mathcal{M} \otimes \mathcal{N} := (\mathcal{M} \cup \mathcal{N})''$ for finite-dimensional mutually commuting subalgebras $\mathcal{M}, \mathcal{N} \subset \mathcal{A}_{\mathbb{Z}}$, one has the following result.

Theorem 8.2 (Factorization [GNVW12])**.**

$$\mathcal{C}_n := \mathcal{A}_{\{2n-1,2n\}} = \mathcal{L}_n \otimes \mathcal{R}_{n-1} \tag{8.3}$$

$$\mathcal{B}_n := \mathcal{A}_{\{2n,2n+1\}} = \alpha^{-1}(\mathcal{L}_n) \otimes \alpha^{-1}(\mathcal{R}_n). \tag{8.4}$$

Thus $\alpha^{-1}(\mathcal{L}_n)$ is the part of $\mathcal{B}_n$ that α sends to the left, and $\alpha^{-1}(\mathcal{R}_n)$ is the part of $\mathcal{B}_n$ that α sends to the right. Likewise, $\mathcal{C}_n$ is composed of a part $\mathcal{L}_n$ that was sent leftward from $\mathcal{B}_n$, and a part $\mathcal{R}_{n-1}$ that was sent rightward from $\mathcal{B}_{n-1}$.

Proof. Recall from Eq. (7.4) that in general, for a finite-dimensional subalgebra $\mathcal{M} \subset \mathcal{A}_{\mathbb{Z}}$, we have the conditional expectation $\mathbb{E}_{\mathcal{M}'}(x) = \int_{U(\mathcal{M})} u x u^* \, \mathrm{d}u$. We first show

$$\mathcal{L}_n := \alpha(\mathcal{B}_n) \cap \mathcal{C}_n = \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(\mathcal{B}_n)) \tag{8.5}$$

$$\mathcal{R}_{n-1} := \alpha(\mathcal{B}_{n-1}) \cap \mathcal{C}_n = \mathbb{E}_{\mathcal{C}'_{n-1}}(\alpha(\mathcal{B}_{n-1})). \tag{8.6}$$

Clearly, $\mathcal{L}_n \subseteq \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(\mathcal{B}_n))$. To show the reverse inclusion, let $y = \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(x))$ for some $x \in \mathcal{B}_n$, i.e.

$$y = \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(x)) = \int_{U(\mathcal{C}_{n+1})} u \alpha(x) u^* \, \mathrm{d}u.$$

From this expression, we see $[y, \alpha(\mathcal{B}_{n-1})] = 0$ because we have $[\alpha(x), \alpha(\mathcal{B}_{n-1})] = 0$ and $[\mathcal{C}_{n+1}, \alpha(\mathcal{B}_{n-1})] = 0$ (the latter because $\alpha(\mathcal{B}_{n-1}) \subseteq \mathcal{C}_{n-1} \otimes \mathcal{C}_n$). On the other hand, it follows from $\alpha(\mathcal{B}_n) \subseteq \mathcal{C}_n \otimes \mathcal{C}_{n+1}$ that $y \in \mathcal{C}_n$. Moreover, α^{-1} is again a nearest neighbour QCA, so we have $\alpha^{-1}(\mathcal{C}_n) \subset \mathcal{B}_{n-1} \otimes \mathcal{B}_n$, so we find that $y \in \alpha(\mathcal{B}_{n-1} \otimes \mathcal{B}_n)$. Then $[y, \alpha(\mathcal{B}_{n-1})] = 0$ implies $y \in \alpha(\mathcal{B}_n)$. We conclude Eq. (8.5) holds; a similar argument shows Eq. (8.6).

Finally we demonstrate $\mathcal{C}_n \subseteq \mathcal{L}_n \otimes \mathcal{R}_{n-1}$, which then becomes an equality. For any $c \in \mathcal{C}_n$, we can express $\alpha^{-1}(c) \in \mathcal{B}_{n-1} \otimes \mathcal{B}_n$ as $\alpha^{-1}(c) = \sum_i a_i b_i$ for some elements $a_i \in \mathcal{B}_{n-1}, b_i \in \mathcal{B}_n$. Then

$$\begin{aligned}
c = \mathbb{E}_{\mathcal{C}'_{n-1}} \mathbb{E}_{\mathcal{C}'_{n+1}}(c) &= \sum_i \mathbb{E}_{\mathcal{C}'_{n-1}} \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(a_i)\alpha(b_i)) \\
&= \sum_i \mathbb{E}_{\mathcal{C}'_{n-1}}(\alpha(a_i)) \mathbb{E}_{\mathcal{C}'_{n+1}}(\alpha(b_i)) \in \mathcal{L}_n \otimes \mathcal{R}_{n-1},
\end{aligned}$$
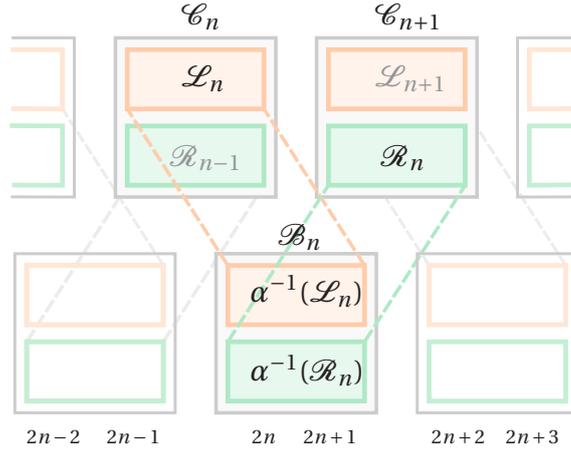
Figure 8.1: The factorization Theorem 8.2 decomposes every two-site algebra into a left-moving and right-moving part.

as desired. The final equality follows from $\alpha(a_i) \in \mathscr{C}_{n-1} \otimes \mathscr{C}_n$ and $\alpha(b_i) \in \mathscr{C}_n \otimes \mathscr{C}_{n+1}$, and the final inclusion is manifest from Eqs. (8.5) and (8.6). Thus we have proved Eq. (8.3).

Noting again that α^{-1} is a nearest neighbour QCA, similar logic applied to α^{-1} yields Eq. (8.4). Specifically, Eqs. (8.5) and (8.6) are replaced by

$$\alpha^{-1}(\mathscr{L}_n) = \mathscr{B}_n \cap \alpha^{-1}(\mathscr{C}_n) = \mathbb{E}_{\mathscr{B}'_{n-1}}(\alpha^{-1}(\mathscr{C}_n))),$$
$$\alpha^{-1}(\mathscr{R}_n) = \mathscr{B}_n \cap \alpha^{-1}(\mathscr{C}_{n+1}) = \mathbb{E}_{\mathscr{B}'_{n+1}}(\alpha^{-1}(\mathscr{C}_{n+1}))),$$

which follow using the inclusions $\alpha^{-1}(\mathscr{C}_n) \subseteq \mathscr{B}_{n-1} \otimes \mathscr{B}_n$, $\alpha^{-1}(\mathscr{C}_{n+1}) \subseteq \mathscr{B}_n \otimes \mathscr{B}_{n+1}$, as well as $\alpha(\mathscr{B}_n) \subseteq \mathscr{C}_n \otimes \mathscr{C}_{n+1}$, and one then uses this to prove the nontrivial inclusion

$$\mathscr{B}_n \subseteq \alpha^{-1}(\mathscr{L}_n)) \otimes \alpha^{-1}(\mathscr{R}_n).$$

∎

For later use in Section 9.2, below we note Theorem 8.2 also holds for weaker assumptions, by an identical argument.

Remark 8.3. Although in Theorem 8.2 we assumed the automorphism α was a QCA, the only locality properties of α required to achieve $\mathscr{C}_n = \mathscr{L}_n \otimes \mathscr{R}_{n-1}$ were

$$\alpha(\mathscr{B}_{n-1}) \subseteq \mathscr{C}_{n-1} \otimes \mathscr{C}_n, \qquad \alpha(\mathscr{B}_n) \subseteq \mathscr{C}_n \otimes \mathscr{C}_{n+1}, \qquad \alpha^{-1}(\mathscr{C}_n) \subseteq \mathscr{B}_{n-1} \otimes \mathscr{B}_n.$$

Similarly, to achieve $\mathscr{B}_n = \alpha^{-1}(\mathscr{L}_n) \otimes \alpha^{-1}(\mathscr{R}_n)$ we need only

$$\alpha^{-1}(\mathscr{C}_n) \subseteq \mathscr{B}_{n-1} \otimes \mathscr{B}_n, \qquad \alpha^{-1}(\mathscr{C}_{n+1}) \subseteq \mathscr{B}_n \otimes \mathscr{B}_{n+1}, \qquad \alpha(\mathscr{B}_n) \subseteq \mathscr{C}_n \otimes \mathscr{C}_{n+1}.$$

Based on Theorem 8.2 we can show that the ratio of $\dim(\mathscr{L}_n)$ and $\dim(\mathscr{A}_{2n})$ is independent of n, motivating the following definition:

Definition 8.4 (Index of QCA)**.** Suppose α is a one-dimensional nearest neighbour QCA. Let $\mathscr{L}_n$ and $\mathscr{R}_n$ be defined as in (8.2), then the *index* of α is given by

$$\text{ind}(\alpha) := \frac{1}{2}\big(\log(\dim(\mathscr{L}_n)) - \log(\dim(\mathscr{A}_{2n}))\big) \tag{8.7}$$
$$= \frac{1}{2}\big(\log(\dim(\mathscr{A}_{2n+1})) - \log(\dim(\mathscr{R}_n))\big).$$

The value of $\text{ind}(\alpha)$ is independent of the choice of n [GNVW12]. We choose to take the logarithm of the original definition. The index of a QCA with radius $R > 1$ may be defined by blocking sites such that the resulting QCA is nearest neighbour, and one can show that the index is independent of the choice of blocking. This index can be thought of as a 'flux', measuring the difference between how much quantum information is flowing to the right vs. left. From the definition it is clear it cannot take arbitrary values, but is restricted to integer linear combinations $\mathbb{Z}[\{\log(p_i)\}]$ where the p_i are all prime factors of local Hilbert space dimensions d_n.

The index can be used to characterize all one-dimensional QCAs. In order to do so, we introduce two types of QCAs: circuits and shifts. We will say a QCA α is a *block partitioned unitary* if it can be written as

$$\alpha(x) = \left(\prod_j u_j^*\right) x \left(\prod_j u_j\right)$$

where the u_j are a family of local unitaries, the u_j having disjoint and finite support. We will say α is a *circuit* (in [GNVW12] a similar notion is called *locally implementable*) if it can be written as a composition of block partitioned unitaries where each local unitary is supported on a uniformly bounded finite set. In one dimension any circuit QCA of radius R can be written as a composition of at most two block partitioned unitaries where each local unitary is supported on at most $2R$ contiguous sites. We denote by σ_d^k the *translation* QCA which has local Hilbert space dimension d and which translates any operator by k sites, mapping $\sigma_d^k(\mathscr{A}_n) = \mathscr{A}_{n-k}$. Here k can be negative. We will say a QCA is a *shift* if it is a tensor product of translations of the form σ_d^k.

Theorem 8.5 (Properties of GNVW index [GNVW12])**.** *Let α, β be one-dimensional QCAs. Then:*

(i) $\text{ind}(\alpha \otimes \beta) = \text{ind}(\alpha) + \text{ind}(\beta)$

(ii) *If α and β are defined on the same quasi-local algebra (i.e., with the same local dimensions),* $\text{ind}(\alpha\beta) = \text{ind}(\alpha) + \text{ind}(\beta)$.

(iii) *α is a circuit if and only if* $\text{ind}(\alpha) = 0$.

(iv) $\text{ind}(\sigma_d^k) = k\log(d)$.

(v) *Every one-dimensional QCA is a composition of a shift and a circuit.*[1]

(vi) *If α and β are defined on the same quasi-local algebra the following are equivalent:*

    (a) $\text{ind}(\alpha) = \text{ind}(\beta)$.

    (b) *There exists a circuit γ such that $\alpha = \beta\gamma$.*

    (c) *There exists a strongly continuous path from α to β through the space of QCAs with a uniform bound on the radius.*

---

[1]Strictly speaking this only makes sense if all the local dimensions are equal. We can always achieve this by taking a tensor product with the identity automorphism on a quasi-local algebra with appropriate local dimensions.

(d) *There exists a* blending *of α and β, meaning a QCA γ which is identical to α on a region extending to left infinity and equal to β on a region extending to right infinity.*

The "classification" of one-dimensional QCAs refers to the set of QCAs modulo an equivalence relation, given either by circuits (b), continuous deformations (c), or blending (d). These equivalence classes are identical and characterized by the index, as expressed in (vi). If α and β are not defined on the same quasi-local algebra (i.e. have different local dimensions), analogous statements to (b), (c), (d) hold after separately tensoring α and β with appropriate identity automorphisms, i.e. adding extra tensor factors on which they act trivially, such that α and β then have the same local dimensions. The notion of equivalence between QCAs that further allows extra tensor factors is called 'stable equivalence,' discussed in [FH20]. We will prove generalizations of all these properties for ALPUs.

As observed in [GNVW12], the tensor product property together with the normalization on shifts and circuits completely determines the index.

Lemma 8.6. *Suppose I assigns a real number $I(\alpha)$ to any one-dimensional QCA α such that*

(i) *$I(\alpha \otimes \beta) = I(\alpha) + I(\beta)$ for all one-dimensional QCAs α and β.*

(ii) *I takes the same values as* ind *on circuits and on σ_d^k.*

Then $I(\alpha) = \text{ind}(\alpha)$ for any one-dimensional QCA α.

Proof. Let α be any one-dimensional QCA and let β be a shift with $I(\beta) = \text{ind}(\beta) = -\text{ind}(\alpha)$, using (ii). Then $I(\alpha \otimes \beta) = I(\alpha) + I(\beta) = I(\alpha) - \text{ind}(\alpha)$ by (i). On the other hand, $\text{ind}(\alpha \otimes \beta) = 0$ so it is a circuit. Again by property (ii) this implies that $I(\alpha \otimes \beta) = 0$, showing that $I(\alpha) = \text{ind}(\alpha)$. ■

8.2 The GNVW index revisited

We will now discuss some new results regarding the GNVW index for QCAs. First, in Section 8.2.1 we provide an alternative formula for the index in terms of a difference of mutual informations. In Section 8.2.2 we prove some results about QCAs which are locally close to each other. These results are interesting in their own right, but will also be crucial when extending the index to ALPUs.

8.2.1 An entropic definition of the GNVW index

Here we provide a new formula for the index in terms of the mutual information, which can also be defined for infinite C^*-algebras. This reformulation is not strictly necessary to develop an index theory for ALPUs, but it does allow us to give a clean expression for the index of an ALPU.

We consider two copies of the quasi-local algebra $\mathscr{A}_{\mathbb{Z}}$. Then the tensor product algebra $\mathscr{A}_{\mathbb{Z}} \otimes \mathscr{A}_{\mathbb{Z}}$ is uniquely defined as a C^*-algebra since $\mathscr{A}_{\mathbb{Z}}$ is nuclear (so there is no ambiguity in the norm completion of the tensor product). We choose a transposition on
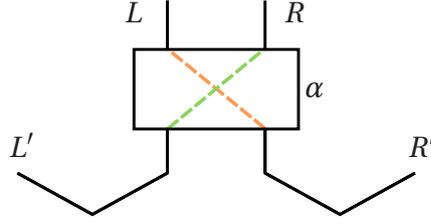
Figure 8.2: Illustration of (8.11). The index measures the difference in information flows, left to right minus right to left, as $\mathrm{ind}(\alpha) = \frac{1}{2}\left(I(L':R)_\phi - I(L:R')_\phi\right)$.

each local algebra, which gives rise to a transposition $x \mapsto x^T$ on $\mathscr{A}_\mathbb{Z}$. Let τ be the tracial state on $\mathscr{A}_\mathbb{Z}$. Then we define the *maximally entangled state ω* by

$$\omega(x \otimes y) = \tau(xy^T) \tag{8.8}$$

for $x \otimes y \in \mathscr{A}_\mathbb{Z} \otimes \mathscr{A}_\mathbb{Z}$. It is not hard to see that if we restrict to a finite number of sites, ω indeed restricts to the usual maximally entangled state. Then we define

$$\phi = (\alpha^\dagger \otimes I)(\omega).$$

where I is the identity channel, and $\alpha^\dagger$ is the adjoint channel. In other words, ϕ is the *Choi state* of α.

Split the algebra $\mathscr{A}_\mathbb{Z}$ at any point n in the chain, letting

$$\begin{aligned}\mathscr{A}_L &:= \mathscr{A}_{\leq n}\\ \mathscr{A}_R &:= \mathscr{A}_{>n}.\end{aligned} \tag{8.9}$$

and similarly split the copy as $\mathscr{A}_{L'}$ and $\mathscr{A}_{R'}$. For a QCA with radius r, we will also consider

$$\begin{aligned}\mathscr{A}_{L_1} &= \mathscr{A}_{n-r+1,\dots,n}, &\mathscr{A}_{L_2} &= \mathscr{A}_{\leq n-r},\\ \mathscr{A}_{R_1} &= \mathscr{A}_{n+1,\dots,n+r}, &\mathscr{A}_{R_2} &= \mathscr{A}_{\geq n+r+1}.\end{aligned} \tag{8.10}$$

We will define the index in terms of a difference of mutual informations of the Choi state. If ϕ, ψ are states on a C^*-algebra we may define the relative entropy $D(\phi, \psi)$ [OP04]. The mutual information of a state ϕ on $\mathscr{A}_A \otimes \mathscr{A}_B$ can then defined using the relative entropy as $I(A:B)_\phi = D(\phi, \phi|_{\mathscr{A}_A} \otimes \phi|_{\mathscr{A}_B})$. On finite dimensional subsystems this definition coincides with the usual one. The only property we need is that relative entropies, and hence mutual informations, on the full algebra can be computed as limits:

Proposition 8.7 (Proposition 5.23 in [OP04])**.** *Let $\mathscr{A}$ be a C^*-algebra and let $\{\mathscr{A}_i\}_i$ be an increasing net of C^*-subalgebras so that $\cup_i \mathscr{A}_i$ is dense in $\mathscr{A}$. Then for any two states ϕ, ψ on $\mathscr{A}$ the net $D(\phi_i, \psi_i)$ converges to $D(\phi, \psi)$ where $\phi_i = \phi|_{\mathscr{A}_i}, \psi_i = \psi|_{\mathscr{A}_i}$.*

Proposition 8.8. *For any choice of n in Eq. (8.9) the index of a one-dimensional QCA α is given by*

$$\mathrm{ind}(\alpha) = \frac{1}{2}\left(I(L':R)_\phi - I(L:R')_\phi\right). \tag{8.11}$$

For a QCA with radius r, this can also be computed locally as

$$\mathrm{ind}(\alpha) = \frac{1}{2}\left(I(L_1':R_1)_\phi - I(L_1:R_1')_\phi\right). \tag{8.12}$$

Here, the mutual information terms are computed with respect to the corresponding subalgebras of $\mathscr{A}_{\mathbb{Z}} \otimes \mathscr{A}_{\mathbb{Z}}$ (with primed systems corresponding to subalgebras of the second factor).

Proof. Denote by $I(\alpha)$ the expression in (8.11). First we will argue that

$$I(L' : R)_\phi = I(L'_1 : R_1)$$
$$I(L : R')_\phi = I(L_1 : R'_1).$$

One sees this by verifying that

$$\phi_{L'R} = \phi_{L'_1 R_1} \otimes \tau_{L'_2 R_2}$$
$$\phi_{LR'} = \phi_{L_1 R_1} \otimes \tau_{L_2 R'_2}$$

where the τ denote tracial (i.e maximally mixed) states. Next, to see that $I(\alpha) = \text{ind}(\alpha)$ we will apply Lemma 8.6. From the definition it is clear that $I(\alpha \otimes \beta) = I(\alpha) + I(\beta)$, so it suffices to compute $I(\alpha)$ for a circuit and a shift. For a shift $\alpha = \sigma_d^k$ it is clear from the definition that for positive k

$$I(L' : R)_\phi = 2k \log(d)$$
$$I(L : R')_\phi = 0$$

and for negative k

$$I(L' : R)_\phi = 0$$
$$I(L : R')_\phi = 2k \log(d).$$

Finally, for a circuit α, notice that we can ignore any unitaries that act only on L or R as they keep the mutual information invariant. In this way, we may also reduce to the finite subsystem $L_1 R_1 L'_1 R'_1$. In order to see that $I(\alpha) = 0$ we thus only need to check that

$$I(L'_1 : R_1)_\phi = I(L_1 : R'_1)_\phi$$

where $|\phi\rangle = U \otimes I |\omega\rangle$ for some unitary U acting on $L_1 R_1$ and where $|\omega\rangle$ is a maximally entangled state between $L_1 R_1$ and $L'_1 R'_1$. In that case $|\phi\rangle$ is a maximally entangled state between $L_1 R_1$ and $L_2 R_2$ and

$$H(L'_1)_\phi = H(L_1)_\phi$$
$$H(R'_1)_\phi = H(R_1)_\phi$$
$$H(L'_1 R_1)_\phi = H(L_1 R'_1)_\phi.$$

The first two equalities hold because ϕ is maximally entangled, and the third equality holds because ϕ is pure. Thus we see that

$$I(L'_1 : R_1)_\phi = H(L'_1)_\phi + H(R_1)_\phi - H(L'_1 R_1)_\phi$$
$$= H(L_1)_\phi + H(R'_1)_\phi - H(L_1 R'_1)_\phi$$
$$= I(L_1 : R'_1)_\phi. \qquad \blacksquare$$

The expression of the index in (8.11) is intuitive: $I(L' : R)_\phi$ and $I(L : R')_\phi$ measure the flow of information to the right and left respectively. Notice that depending on the choice of cut $I(L' : R)_\phi$ and $I(L : R')_\phi$ can vary individually, but the total *flux* as defined by (8.11) is invariant. One reason this expression for the index is useful is that, contrary to the original definition, it is plausibly well-defined for automorphisms which are not strictly local (or for channels which are not automorphisms). In Theorem 9.18 we will show that taking the limit of the finite subalgebras in (8.10) with increasing radius gives a well-defined and finite limit for any ALPU with appropriately decaying tails, and hence using Proposition 8.7 we conclude that both mutual information terms in (8.11) are finite and (8.11) gives a finite, quantized answer also for an ALPU.

In [GNVW12], a similar numerical expression for the index is provided in terms of overlaps of algebras (their Eq. 45). In fact, their formula (or rather its logarithm) can be interpreted as (8.11) but with the entropies replaced by Rényi-2 entropies,

$$\mathrm{ind}(\alpha) = \frac{1}{2}\left(I_2(L' : R)_\phi - I_2(L : R')_\phi\right),$$

where $I_2(A : B)_\rho := H_2(A)_\rho + H_2(B)_\rho - H_2(AB)_\rho$. While the values of the individual mutual information terms depend on the choice of Rényi-2 or von Neumann entropy, for QCAs, the difference of mutual informations used to define the index does *not* depend on this choice, and in the proof of Proposition 8.8 one can simply replace the entropies H by Rényi entropies H_2. However, the mutual information has better continuity properties with respect to the dimension of the local Hilbert spaces compared to the Rényi-2 mutual information (compare the following with the continuity bound in Lemma 12 of [GNVW12]):

Theorem 8.9 (Continuity of mutual information [AF04, Win16, Wil13])**.** *Suppose ρ, σ are states on $\mathcal{H}_A \otimes \mathcal{H}_B$, and $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\|_1 \leq \varepsilon < 1$. Then*

$$\left|I(A : B)_\rho - I(A : B)_\sigma\right| \leq 3\varepsilon \log(d_A) + 2(1 + \varepsilon)h\left(\frac{\varepsilon}{1 + \varepsilon}\right) \leq 3\varepsilon \log(d_A) + \varepsilon \log\frac{1}{\varepsilon}$$

where $d_A = \dim(\mathcal{H}_A)$ and $h(x) = -x\log(x) - (1 - x)\log(1 - x)$ is the binary entropy.

This continuity is important for the extension to ALPUs, where we need to compute the approximation to the index on a sequence of increasing finite subalgebras. In that case, the indices defined using the Rényi-2 and von Neumann entropies give different answers when restricted to the finite subalgebras. A final remark is that (8.12) can also be rewritten as an entropy difference

$$\begin{aligned}
\mathrm{ind}(\alpha) &= \frac{1}{2}\left(I(L_1' : R_1)_\phi - I(L_1 : R_1')_\phi\right) \\
&= \frac{1}{2}\left(H(L_1 R_1')_\phi - H(L_1' R_1)_\phi\right).
\end{aligned} \tag{8.13}$$

However, the extension of this expression to infinite-dimensional setting is less clear, because both terms diverge.

8.2.2   Robustness of the GNVW index

Because the index can be computed locally, it appears that two QCAs with different index should be easy to distinguish locally. We make this quantitative in Proposition 8.11: two

QCAs which look locally similar must have equal index. We begin with a cruder but more general estimate, describing how the mutual information of the Choi state varies continuously with respect to the automorphism that defines it. This estimate applies to general automorphisms which may not be QCAs, proving useful in the argument for Theorem 9.18.

Let α be an automorphism of $\mathscr{A}_{\mathbb{Z}}$. Even when α is not a QCA, we can mimic the local definition of the index in (8.12) using finite disjoint regions L, R. We denote this quantity $\widetilde{\operatorname{ind}}_{L,R}(\alpha)$ to emphasize α may not be a QCA nor even an ALPU,

$$\widetilde{\operatorname{ind}}_{L,R}(\alpha) = \frac{1}{2}\left(I(L':R)_{\phi} - I(L:R')_{\phi}\right), \tag{8.14}$$

where the mutual information terms are computed with respect to the corresponding subalgebras of $\mathscr{A}_{\mathbb{Z}} \otimes \mathscr{A}_{\mathbb{Z}}$. (As above, primed systems refer to the second copy of $\mathscr{A}_{\mathbb{Z}}$.) Clearly, Eq. (8.14) only depends on the restriction of the Choi state to $\mathscr{A}_X \otimes \mathscr{A}_{X'}$, where $X = L \cup R$, i.e. on the state $\widetilde{\phi}_{XX'} := \phi|_{\mathscr{A}_X \otimes \mathscr{A}_{X'}}$, which is given by

$$\widetilde{\phi}_{XX'}(x) = \omega\left((\alpha \otimes I)(x)\right)$$

for all $x \in \mathscr{A}_X \otimes \mathscr{A}_{X'}$. Then we have the following continuity estimate.

Lemma 8.10. *For two automorphisms α_1 and α_2 of $\mathscr{A}_{\mathbb{Z}}$ with maximum local dimension d, the quantity $\widetilde{\operatorname{ind}}_{L,R}$ in (8.14) obeys*

$$\left|\widetilde{\operatorname{ind}}_{L,R}(\alpha_1) - \widetilde{\operatorname{ind}}_{L,R}(\alpha_2)\right| = \mathcal{O}\left(\varepsilon|X|\log(d) + \varepsilon\log\tfrac{1}{\varepsilon}\right),$$

where $\varepsilon = \|(\alpha_1 - \alpha_2)|_{\mathscr{A}_X}\|$. The same continuity estimate with respect to α_1 and α_2 holds for the individual terms in (8.14).

Proof. First we compare the restricted Choi states $\widetilde{\phi}_{XX',1}$ and $\widetilde{\phi}_{XX',2}$ of α_1 and α_2, respectively. For any $x \in \mathscr{A}_X \otimes \mathscr{A}_{X'}$ with $\|x\| = 1$,

$$\left|\widetilde{\phi}_{XX',1}(x) - \widetilde{\phi}_{XX',2}(x)\right|_1 = \left|\omega((\alpha_1 \otimes I - \alpha_2 \otimes I)(x))\right| \le \|(\alpha_1 \otimes I - \alpha_2 \otimes I)|_{\mathscr{A}_X \otimes \mathscr{A}_{X'}}\|.$$

Thus the trace distance between the two Choi states is bounded by

$$\|\widetilde{\phi}_{XX',1} - \widetilde{\phi}_{XX',2}\|_1 \le \|(\alpha_1 \otimes I - \alpha_2 \otimes I)|_{\mathscr{A}_X \otimes \mathscr{A}_{X'}}\| \le 2\varepsilon + \mathcal{O}(\varepsilon^2)$$

using Lemma 7.7 for the last inequality (with $\mathscr{A}_1 = \mathscr{A}_X \otimes I$, $\mathscr{A}_2 = I \otimes \mathscr{A}_{X'}$ and $\mathscr{A} = \mathscr{A}_X \otimes \mathscr{A}_{X'}$ finite-dimensional and $\mathscr{B} = \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$). The conclusion follows from the continuity of mutual information in Theorem 8.9 with respect to the state, noting the region X has associated Hilbert space of dimension at most $d^{|X|}$. ∎

If α_1 and α_2 are one-dimensional QCAs of radius r, then because the index takes discrete values, there exists ε_0 such that if $\varepsilon \le \frac{\varepsilon_0}{r\log(d)}$ then $\operatorname{ind}(\alpha_1) = \operatorname{ind}(\alpha_2)$. However, we can do better and eliminate the dependence on the local dimension, as a simple application of Theorem 7.6. By blocking sites, we may assume without loss of generality that the QCA is nearest neighbour.

Proposition 8.11 (Robustness of GNVW index for QCAs)**.** *Suppose α_1 and α_2 are two nearest-neighbour QCAs defined on the same quasi-local algebra $\mathscr{A}_{\mathbb{Z}}$ such that*

$$\|(\alpha_1 - \alpha_2)|_{\mathscr{A}_{\{2n,2n+1\}}}\| \leq \varepsilon$$

for some n with $\varepsilon \leq \frac{1}{192}$. Then $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$.

Moreover, the algebras $\mathscr{L}_n^{(1)}$ and $\mathscr{L}_n^{(2)}$ defined by (8.2) *using α_1 and α_2 respectively are isomorphic, with the isomorphism implemented by a unitary $u \in \mathscr{A}_{\{2n-1,2n\}}$ which satisfies $\|u - I\| \leq 36\varepsilon$.*

Note that when working with a coarse-grained QCA, where each site is composed of many smaller sites, the hypotheses like $\|(\alpha_1 - \alpha_2)|_{\mathscr{A}_{\{2n,2n+1\}}}\| \leq \varepsilon$ constraining error on coarse-grained sites may always be replaced by hypotheses constraining the sum of errors on fine-grained sites, using Lemma 7.7. (In other words, upper bounds for errors on small regions control errors on larger regions.)

Proof. By the structure theory for QCAs in Theorem 8.2 there exist algebras $\mathscr{L}_n^{(i)}, \mathscr{R}_{n-1}^{(i)}$ for $i = 1, 2$ defined as in (8.2) that satisfy

$$\mathscr{A}_{\{2n-1,2n\}} = \mathscr{L}_n^{(i)} \otimes \mathscr{R}_{n-1}^{(i)}$$

To prove that $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$, by (8.7) it suffices to show that $\mathscr{L}_n^{(1)}$ and $\mathscr{L}_n^{(2)}$ are isomorphic. To see the isomorphism, take $x \in \mathscr{L}_n^{(1)}$ with $\|x\| = 1$ and let $y = \alpha_2(\alpha_1^{-1}(x))$. Then

$$\|x - y\| = \|\alpha_1(\alpha_1^{-1}(x)) - \alpha_2(\alpha_1^{-1}(x))\| \leq \varepsilon$$

using the assumption $\|(\alpha_1 - \alpha_2)|_{\mathscr{A}_{\{2n,2n+1\}}}\| \leq \varepsilon$ and noting that $\alpha_1^{-1}(x) \in \mathscr{A}_{\{2n,2n+1\}}$ since we assume $x \in \mathscr{L}_n^{(1)}$. Using the conditional expectation from (7.4), define

$$z = \mathbb{E}_{\mathscr{A}'_{\{2n+1,2n+2\}}}(y) = \int_{U(\mathscr{A}_{\{2n+1,2n+2\}})} u y u^* \, \mathrm{d}u$$

such that $z \in \mathscr{L}_n^{(2)}$ by the characterization of $\mathscr{L}_n$ in (8.5). Note

$$\|[a, y]\| = \|[a, y - x]\| \leq 2\varepsilon \|a\|$$

for all $a \in \mathscr{A}_{\{2n+1,2n+2\}}$, so by its definition z satisfies $\|y - z\| \leq 2\varepsilon$, so we can bound

$$\|x - z\| \leq \|x - y\| + \|y - z\| \leq 3\varepsilon.$$

We conclude $\mathscr{L}_n^{(1)} \overset{3\varepsilon}{\subseteq} \mathscr{L}_n^{(2)}$, and by a symmetric argument we see $\mathscr{L}_n^{(2)} \overset{3\varepsilon}{\subseteq} \mathscr{L}_n^{(1)}$. By Theorem 7.6, noting that $3\varepsilon \leq \frac{1}{64}$, we obtain that $\mathscr{L}_n^{(1)}$ and $\mathscr{L}_n^{(2)}$ are isomorphic, and the isomorphism is implemented by a unitary $u \in \mathscr{A}_{\{2n-1,2n\}}$ with $\|u - I\| \leq 36\varepsilon$. ∎

For later use in Section 9.2, below we build on Remark 8.3 to note that Proposition 8.11 also holds for weaker assumptions, by an identical argument.

Remark 8.12. Although in Proposition 8.11 we assumed the automorphisms α_1 and α_2 were QCAs, the only locality properties required to achieve the isomorphism between the algebras $\mathscr{L}_n^{(1)}$ and $\mathscr{L}_n^{(2)}$ are the properties listed in Remark 8.3 as those required to achieve $\mathscr{A}_{\{2n-1,2n\}} = \mathscr{L}_n^{(i)} \otimes \mathscr{R}_{n-1}^{(i)}$ for $i = 1,2$. More explicitly, we only require the inclusions

$$\alpha_i(\mathscr{A}_{\{2n-2,2n-1\}}) \subseteq \mathscr{A}_{\{2n-3,\dots,2n\}}, \qquad \alpha_i(\mathscr{A}_{\{2n,2n+1\}}) \subseteq \mathscr{A}_{\{2n-1,\dots,2n+2\}},$$

as well as the inclusion $\alpha_i^{-1}(\mathscr{A}_{\{2n-1,2n\}}) \subseteq \mathscr{A}_{\{2n-2,\dots,2n+1\}}$ for $i = 1,2$.

This also allows us to confirm the intuition that a one-dimensional QCA which is locally close to the identity can be implemented locally with unitaries close to the identity.

Proposition 8.13. *Suppose α is a one-dimensional QCA with radius R and suppose that for $\varepsilon \le \frac{1}{192}$ we have $\|\alpha(x) - x\| \le \varepsilon \|x\|$ for any $x \in \mathscr{A}_{\mathbb{Z}}$ supported on at most $2R$ sites. Then α can be implemented as a composition of two block partitioned unitaries $u = \prod_n u_n$ and $v = \prod_n v_n$, i.e.*

$$\alpha(x) = v^* u^* x u v$$

with each of the unitaries u_n, v_n acting on $2R$ adjacent sites and satisfying

$$\|u_n - I\| = \mathscr{O}(\varepsilon), \quad \|v_n - I\| = \mathscr{O}(\varepsilon).$$

Proof. By blocking sites in groups of R sites, we may assume without loss of generality that α is nearest neighbour. Let $\alpha_1 = I$ and $\alpha_2 = \alpha$ in Proposition 8.11. Clearly, we have $\mathscr{L}_n^{(1)} = \mathscr{A}_{2n}$ and $\mathscr{R}_{n-1}^{(1)} = \mathscr{A}_{2n-1}$. Proposition 8.11 provides a unitary $v_n \in \mathscr{A}_{\{2n-1,2n\}}$ such that $v_n \mathscr{L}_n^{(2)} v_n^* = \mathscr{A}_{2n}$ with $\|v_n - I\| = \mathscr{O}(\varepsilon)$. It follows that $v_n \mathscr{R}_{n-1}^{(2)} v_n^* = \mathscr{A}_{2n-1}$. We let $v = \prod v_n$ and let $\tilde{\alpha} = v\alpha(x)v^*$. Then $\tilde{\alpha}(\mathscr{A}_{\{2n,2n+1\}}) = \mathscr{A}_{\{2n,2n+1\}}$. Moreover, for all $x \in \mathscr{A}_{\{2n,2n+1\}}$, we estimate

$$\begin{aligned}
\|\tilde{\alpha}(x) - x\| &\le \|v\alpha(x)v^* - \alpha(x)\| + \|\alpha(x) - x\| \\
&\le 2\|v_n \otimes v_{n+1} - I\|\|x\| + \varepsilon\|x\| \\
&\le 2(\|v_n - I\| + \|v_{n+1} - I\|)\|x\| + \varepsilon\|x\| \\
&= \mathscr{O}(\varepsilon)\|x\|.
\end{aligned}$$

Then Proposition 7.10 shows that $\tilde{\alpha}|_{\mathscr{A}_{\{2n,2n+1\}}}$ can be implemented by a unitary u_n on $\mathscr{A}_{\{2n,2n+1\}}$ with $\|u_n - I\| = \mathscr{O}(\varepsilon)$. ∎

Index theory for approximately locality preserving unitaries

One of the reasons to study QCAs is that many physical quantum dynamics preserve locality in some form. However, the locality in Definition 8.1 is very stringent, and one the most important classes of automorphisms violates strict locality, while preserving a form of approximate locality: evolution by a geometrically local Hamiltonian. The locality of these evolutions is expressed by so-called *Lieb-Robinson bounds* [LR72]. In Section 9.1 we discuss the Lieb-Robinson bounds and explain how they lead us to define ALPUs as a natural generalization of QCAs. The main results are in Section 9.2 where we show that any ALPU can be approximated by a sequence of QCAs and that this allows us to extend the index theory of Chapter 8.

9.1 Lieb-Robinson bounds and approximate locality

We will state a fairly general form of the Lieb-Robinson bounds which also holds for Hamiltonians which are not strictly local, but have a sufficiently fast decay, following e.g. [NSY19] or [Has10]. Suppose that Γ is a lattice with a metric d. Then a monotonically decreasing function $F\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is called *reproducing* (implying fast decay) if there exists a constant $C > 0$ such that for all $n, m \in \Gamma$,

$$\sum_l F(d(n,l))F(d(l,m)) \leq CF(d(n,m)),$$

$$\sup_y \sum_x F(d(x,y)) < \infty.$$

These conditions are related to a convolution and integral, respectively. For $\Gamma = \mathbb{Z}^D$ with the Euclidean distance, the function $F(r) = (1+r)^{-(D+\varepsilon)}$ is reproducing for any $\varepsilon > 0$. Note the reproducing property is not strictly a measure of fast decay: an exponential decay alone is not reproducing, despite having faster decay than the previous power law, because it fails the first inequality. Meanwhile, $F(r) = (1+r)^{-(D+\varepsilon)}e^{-ar}$ for any $a > 0$ is again reproducing ([NSY19], Appendix 8.2).

Now we consider the automorphism α on the quasi-local algebra $\mathscr{A}_\Gamma$ which is generated by time evolution for some fixed time T by a Hamiltonian

$$H = \sum_{n \in \Gamma} H_n + \sum_{X \subseteq \Gamma} H_X.$$

The terms H_n act only on site n, and the terms H_X act on the sites in X. Then, if the interaction terms H_X have sufficient decay, we have the following bounds on the decay of $\alpha(x) = e^{iHt} x e^{-iHt}$.[1] We state them without the dependence on the time t, which only affects the constant C below, and which is irrelevant for our purposes:

Theorem 9.1 (Lieb-Robinson [NSY19])**.** *For $\alpha(x) = e^{iHt} x e^{-iHt}$ as above, if F is reproducing and*

$$\sup_{n,m\in\Gamma} \sum_{\substack{X\subseteq\Gamma \\ s.t.\ n,m\in X}} \frac{\|H_X\|}{F(d(n,m))} \leq \infty \tag{9.1}$$

then there exists a constant $C > 0$ such that for all $X, Y \subseteq \Gamma$ and for all $x \in \mathscr{A}_X$, $y \in \mathscr{A}_Y$ we have

$$\|[\alpha(x), y]\| \leq C\|x\|\|y\| \sum_{n\in X}\sum_{m\in Y} F(d(n,m)). \tag{9.2}$$

Here, the Hamiltonian is also allowed to be time-dependent, as long as (9.1) holds uniformly. See [NSY19] for a proof and extensive discussion.

We are particularly interested in the one-dimensional case, where $\Gamma = \mathbb{Z}$ and the metric $d(x, y) = |x - y|$ for $x, y \in \mathbb{Z}$. In that setting we consider the case where X is an *interval* (a finite or infinite sequence of consecutive sites) and Y has bounded distance away from X. A consequence of the Lieb-Robinson bounds in (9.2) is that certain algebras form near inclusions.

Lemma 9.2. *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}$ and suppose there exists a monotonically decreasing function $F\colon \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that for all $X, Y \subseteq \Gamma$,*

$$\|[\alpha(x), y]\| \leq \|x\|\|y\| \sum_{n\in X}\sum_{m\in Y} F(|n - m|). \tag{9.3}$$

and suppose $\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} F(n + m) < \infty$. Then for any (finite or infinite) interval $X \subseteq \mathbb{Z}$, we have

$$\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)},$$

where

$$f(r) = 4 \sum_{n,m=0}^{\infty} F(n + m + r + 1).$$

Proof. We first prove the near inclusion for finite X. By Lemma 7.4, it suffices to show that for any $x \in \mathscr{A}_X$ and any $y \in \mathscr{A}_{B(X,r)^c}$ we have

$$\|[\alpha(x), y]\| \leq f(r)\|x\|\|y\|$$

as in that case $\mathscr{A}_{B(X,r)}^{\mathrm{vN}} = \mathscr{A}_{B(X,r)}$. By Eq. (9.3), we know that

$$\|[\alpha(x), y]\| \leq \|x\|\|y\| \sum_{n\in X}\sum_{m\in B(X,r)^c} F(|n - m|)$$

---

[1]In fact, one generally needs these bounds to prove that the time evolution defines a dynamics on the quasi-local algebra, i.e. that time-evolved quasi-local operators are still quasi-local [NSY19].

Let

$$X_k = \{n \in X \text{ such that } d(n, X^c) = k\},$$
$$Y_l = \{m \in X^c \text{ such that } d(m, X) = r + l\},$$

using the notation $d(n, X) = \min_{x \in X} |n - x|$. Since X is an interval the size of each of these sets is upper bounded by 2. We can therefore estimate

$$\sum_{n \in X} \sum_{m \in B(X,r)^c} F(|n - m|) \leq \sum_{k \geq 1} \sum_{l \geq 1} \sum_{n \in X_k} \sum_{m \in Y_l} F(k + l + r - 1)$$
$$\leq 4 \sum_{k \geq 1} \sum_{l \geq 1} F(k + l + r - 1)$$
$$= f(r).$$

We conclude that $\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}$ for any finite interval X.

If X is infinite and $x \in \mathscr{A}_X$, we can take a sequence x_i such that $\lim_i x_i = x$ in norm and each x_i is supported on a finite interval inside X. By what we showed above, for each i there exists some $y_i \in \mathscr{A}_{B(X,r)}$ such that $\|\alpha(x_i) - y_i\| \leq f(r)\|x\|$. Then

$$\inf_{y \in \mathscr{A}_{B(X,r)}} \|\alpha(x) - y\| \leq \liminf_i \|\alpha(x) - y_i\|$$
$$\leq \liminf_i \left(\|\alpha(x) - \alpha(x_i)\| + \|\alpha(x_i) - y_i\| \right)$$
$$\leq f(r)\|x\|. \qquad \blacksquare$$

For instance, if $F(r) = \frac{1}{r^4}$, then $f(r) = \mathcal{O}(\frac{1}{r^2})$; if $F(r) = e^{-ar}\frac{1}{r^2}$ for $a > 0$, then we have $f(r) = \mathcal{O}(e^{-ar})$. As a side note, we observe that one can use Lemma 7.12 on simultaneous near inclusions to show that (in any dimension) Lieb-Robinson type bounds for single-site operators imply bounds for operators on arbitrary sets (which has already been remarked upon in a more restricted setting in [WW20]):

Lemma 9.3. *Suppose α is an automorphism of the quasi-local algebra $\mathscr{A}_\Gamma$ and suppose there exists a function $G \colon \Gamma \times \Gamma \to \mathbb{R}_{\geq 0}$ such that for any $n, m \in \Gamma$, $x \in \mathscr{A}_n$ and $y \in \mathscr{A}_m$*

$$\|[\alpha(x), y]\| \leq \|x\|\|y\|G(n, m).$$

Then for any finite sets $X, Y \subseteq \Gamma$ and $x \in \mathscr{A}_X$, $y \in \mathscr{A}_Y$,

$$\|[\alpha(x), y]\| \leq 128\|x\|\|y\| \sum_{n \in X} \sum_{m \in Y} G(n, m).$$

Proof. By assumption and Lemma 7.4 with $\mathscr{M} = \mathscr{A}_\Gamma^{\mathrm{vN}}$ we have

$$\alpha(\mathscr{A}_n) \overset{G(n,m)}{\subseteq} \mathscr{A}_m' \cap \mathscr{A}_\Gamma^{\mathrm{vN}} = \mathscr{A}_{\Gamma \setminus \{m\}}^{\mathrm{vN}}$$

for all $m, n \in \Gamma$. Applying Lemma 7.12 we find that

$$\alpha(\mathscr{A}_X) \overset{4 \sum_{n \in X} G(n,m)}{\subseteq} \mathscr{A}_{\Gamma \setminus \{m\}}^{\mathrm{vN}}$$

for all $m \in \Gamma$. Lemma 7.5 with $\mathscr{M} = \mathscr{A}_\Gamma^{\mathrm{vN}}$ shows that

$$\mathscr{A}_m \overset{8 \sum_{n \in X} G(n,m)}{\subseteq} \alpha(\mathscr{A}_X)' \cap \mathscr{A}_\Gamma^{\mathrm{vN}} = \alpha(\mathscr{A}_X' \cap \mathscr{A}_\Gamma^{\mathrm{vN}}) = \alpha(\mathscr{A}_{\Gamma \setminus X}^{\mathrm{vN}}).$$

Again applying Lemma 7.12 and Lemma 7.5 as above yields

$$\alpha(\mathscr{A}_X) \overset{64\sum_{n\in X, m\in Y} G(n,m)}{\subseteq} \mathscr{A}_{\Gamma\backslash Y}^{\mathrm{vN}}$$

which implies the desired commutator bound by Lemma 7.4.   ∎

Following [Has13] we would like to generalize the notion of a QCA to the case where the automorphism does not preserve strict locality, but only approximate locality. Such an automorphism is often called quasi-local. There are various choices of definition that require different decays or dependence on support size; see for instance [NSY19]. For our purpose, the definition should at least include Hamiltonian evolutions satisfying Lieb-Robinson bounds. We will restrict to the one-dimensional case, where Theorem 9.1 and Lemma 9.2 inspire the following definition:

Definition 9.4 (ALPU in one dimension). An automorphism α of the quasi-local algebra $\mathscr{A}_{\mathbb{Z}}$ is called an *approximately locality-preserving unitary* (ALPU) if for all (possibly infinite) intervals $X \subseteq \mathbb{Z}$ and for all $r \geq 0$ we have

$$\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}$$

for some positive function $f(r)$ with $\lim_{r\to\infty} f(r) = 0$. Here we use the notation in Definition 7.2.

We say α has $f(r)$-tails when it satisfies the above, or $\mathscr{O}(g(r))$-tails if $f(r) = \mathscr{O}(g(r))$. We will always assume, without loss of generality, that $f(r)$ is non-increasing.

Note that by definition, if α has $f(r)$-tails, it also has $h(r)$-tails for any function $h(r)$ with $h(r) \geq f(r)$ for all r, i.e., $f(r)$ only serves as an upper bound on the spread of α. Furthermore, note that any ALPU has $o(1)$-tails, by definition.

It suffices to check the conditions in Definition 9.4 either for all finite or for all infinite intervals (see Lemmas 9.8 and 9.9 below). If the above conditions on α are satisfied for all intervals X of some fixed size (and arbitrary $r \geq 0$), but $f(r)$ decays exponentially, then in fact α is an ALPU with $\mathscr{O}(f(r))$-tails by Lemma 9.3. We note an equivalent definition of ALPUs when passing to von Neumann algebras in Remark 9.7.

Remark 9.5. In [Has13] what we call an ALPU is simply called a *locality-preserving unitary* (LPU). Moreover, there it is said that an automorphism is a *locally generated unitary* (LGU) if it arises from time evolution by some time-dependent Hamiltonian. We have chosen the more explicit term ALPU instead of LPU, since in the literature the latter has also been used as a synonym for QCA (e.g. [ŞSBC18]).

We note that to call such automorphisms "unitary" is perhaps slightly misleading: there need not be a unitary $u \in \mathscr{A}_{\mathbb{Z}}$ such that $\alpha(x) = u^* x u$ (but there will be a unique unitary implementing α on the GNS Hilbert space with respect to the tracial state, as discussed in Section 7.0.3).

Example. Lemma 9.2 states that for the class of local Hamiltonians in Theorem 9.1 (Lieb-Robinson), the automorphism $\alpha(x) = e^{iHt} x e^{-iHt}$ is an ALPU at fixed t. It turns out that if the Hamiltonian has *exponentially decaying tails* in the sense that the local Hamiltonian terms decay as $\|H_X\| = \mathscr{O}(e^{-k|X|})$ decays exponentially with the size of the support X, then for any $k' < k$ we may take $f(r) = \mathscr{O}(e^{-k'r})$ and α has $\mathscr{O}(e^{-k'r})$-tails [NSY19, Has10]. Such evolutions composed with translations are also ALPUs.

To use Theorem 7.6, we would like to work in the von Neumann algebra $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$. However, in the definition of an ALPU we consider an automorphism of $\mathscr{A}_{\mathbb{Z}}$, not $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$. We therefore prove some results allowing us to translate between the tails for automorphisms of $\mathscr{A}_{\mathbb{Z}}$ versus $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$.

Lemma 9.6.

(i) *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ and $\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_Y^{\mathrm{vN}}$ for some $X, Y \subseteq \mathbb{Z}$ and $\varepsilon \geq 0$. Then $\alpha(\mathscr{A}_X^{\mathrm{vN}}) \overset{\varepsilon}{\subseteq} \mathscr{A}_Y^{\mathrm{vN}}$. In particular, any ALPU with $f(r)$-tails extends to an automorphism α of $\mathscr{A}_X^{\mathrm{vN}}$ such that $\alpha(\mathscr{A}_X^{\mathrm{vN}}) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}$ for any interval X and any $r \geq 0$.*

(ii) *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ such that $\alpha(\mathscr{A}_X^{\mathrm{vN}}) \overset{\varepsilon}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}$ for any interval X and some fixed $r, \varepsilon \geq 0$. Then $\alpha^{-1}(\mathscr{A}_X^{\mathrm{vN}}) \overset{4\varepsilon}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}$ for any interval X.*

(iii) *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ such that $\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}$ for any interval X and any $r \geq 0$, where $f(r)$ is a function with $\lim_{r \to \infty} f(r) = 0$. Then α restricts to an ALPU with $f(r)$-tails.*

(iv) *If α is an ALPU with $f(r)$-tails, then α^{-1} is an ALPU with $4f(r)$-tails.*

Proof. (i) Let $x \in \mathscr{A}_X^{\mathrm{vN}}$. Using the Kaplansky density theorem, choose a net $x_i \in \mathscr{A}_X$ with $\|x_i\| \leq \|x\|$, converging to x in the weak operator topology, hence also in the weak-$*$ topology (since these topologies are the same on bounded subsets). Because α is weak-$*$ continuous, $\alpha(x_i)$ converges to $\alpha(x)$ in the weak-$*$ and hence also in the weak operator topology. Meanwhile, by our assumption that $\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_Y^{\mathrm{vN}}$, there exist $y_i \in \mathscr{A}_Y^{\mathrm{vN}}$ such that $\|\alpha(x_i) - y_i\| \leq \varepsilon \|x\|$. In particular, $\|y_i\| \leq (1 + \varepsilon)\|x\|$. Hence the net y_i is bounded in norm. Since norm balls are compact in the weak operator topology, this implies there must be a converging subnet, which we also denote by y_i. Denoting the limit of y_i by y, then $y \in \mathscr{A}_Y^{\mathrm{vN}}$, and by lower semi-continuity of the norm in the weak operator topology,

$$\|y - \alpha(x)\| \leq \liminf_i \|y_i - \alpha(x_i)\| \leq \varepsilon \|x\|.$$

This shows that $\alpha(x) \overset{\varepsilon}{\in} \mathscr{A}_Y^{\mathrm{vN}}$, proving (i).

(ii) Note that $\mathbb{Z} \setminus B(X, r)$ is a disjoint union of at most two intervals Y_1 and Y_2, and we have $B(Y_i, r) \subseteq \mathbb{Z} \setminus X$ for $i = 1, 2$, so $\alpha(\mathscr{A}_{Y_i}^{\mathrm{vN}}) \overset{\varepsilon}{\subseteq} \mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}}$. Then applying (7.17) in Lemma 7.12 to these two near inclusions with $\mathscr{M} = \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$,

$$\mathscr{A}_X^{\mathrm{vN}} = (\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} \overset{4\varepsilon}{\subseteq} \left(\alpha(\mathscr{A}_{Y_1}^{\mathrm{vN}}) \cup \alpha(\mathscr{A}_{Y_2}^{\mathrm{vN}})\right)' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} = (\alpha(\mathscr{A}_{Y_1}^{\mathrm{vN}}))' \cap (\alpha(\mathscr{A}_{Y_2}^{\mathrm{vN}}))' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$$
$$= \alpha((\mathscr{A}_{Y_1}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}) \cap \alpha((\mathscr{A}_{Y_2}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}})$$
$$= \alpha((\mathscr{A}_{Y_1}^{\mathrm{vN}})' \cap (\mathscr{A}_{Y_2}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}})$$
$$= \alpha(\mathscr{A}_{\mathbb{Z} \setminus Y_1}^{\mathrm{vN}} \cap \mathscr{A}_{\mathbb{Z} \setminus Y_2}^{\mathrm{vN}}) = \alpha(\mathscr{A}_{B(X,r)}^{\mathrm{vN}})$$

and the conclusion follows by applying α^{-1}.

(iii) We need to show that if $x \in \mathscr{A}_{\mathbb{Z}}$, then $\alpha(x) \in \mathscr{A}_{\mathbb{Z}}$. First consider x strictly local, on some finite interval X. Then by assumption there is a sequence $y_r \in \mathscr{A}_{B(X,r)}^{\mathrm{vN}} = \mathscr{A}_{B(X,r)}$

such that $\|\alpha(x) - y_r\| \le f(r)\|x\|$. Hence, y_r is a sequence of strictly local operators converging in norm to $\alpha(x)$ and hence $\alpha(x) \in \mathscr{A}_{\mathbb{Z}}$. If $x \in \mathscr{A}_{\mathbb{Z}}$ is not strictly local, let x_i be a sequence of strictly local operators converging in norm to x. Then $\alpha(x_i) \in \mathscr{A}_{\mathbb{Z}}$ and $\alpha(x_i)$ converges in norm to $\alpha(x)$. Similarly, α^{-1} maps $\mathscr{A}_{\mathbb{Z}}$ into $\mathscr{A}_{\mathbb{Z}}$ (using that (ii) implies locality bounds for α^{-1}) and hence we conclude that α restricts to an automorphism of $\mathscr{A}_{\mathbb{Z}}$.

This implies the desired result, as then

$$\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}^{\mathrm{vN}}_{B(X,r)} \cap \mathscr{A}_{\mathbb{Z}} = \mathscr{A}_{B(X,r)}$$

using the general fact that $\mathscr{A}^{\mathrm{vN}}_Y \cap \mathscr{A}_{\mathbb{Z}} = \mathscr{A}_Y$ for any $Y \subseteq \mathbb{Z}$.

(iv) By (i), α extends to an automorphism of $\mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$ with $f(r)$-tails, by (ii) the inverse of this extension has $4f(r)$-tails, and by (iii) the restriction of the latter is an ALPU with $4f(r)$-tails. ∎

Recall that any automorphism α of the quasi-local algebra $\mathscr{A}_{\mathbb{Z}}$ extends uniquely to an automorphism of the von Neumann algebra $\mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$, which we denote by the same symbol α. Then Lemma 9.6(i) and (iii) together allow an equivalent definition of ALPUs with $f(r)$-tails using the von Neumann algebra, rather than using the quasi-local algebra as in Definition 9.4. We summarize below.

Remark 9.7. Any automorphism $\alpha : \mathscr{A}_{\mathbb{Z}} \to \mathscr{A}_{\mathbb{Z}}$ with $f(r)$-tails, i.e., which is such that for all intervals X and $r \ge 0$ we have

$$\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)},$$

uniquely extends to an automorphism $\alpha : \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}} \to \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$ that has the same tails, i.e. that satisfies $\alpha(\mathscr{A}^{\mathrm{vN}}_X) \overset{f(r)}{\subseteq} \mathscr{A}^{\mathrm{vN}}_{B(X,r)}$. Conversely, any $\alpha : \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}} \to \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$ satisfying the latter (for all X and r) restricts to an ALPU $\alpha : \mathscr{A}_{\mathbb{Z}} \to \mathscr{A}_{\mathbb{Z}}$ with $f(r)$-tails. Hence we may identify an ALPU α with its extension to $\mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$ and refer to the latter also as an "ALPU with $f(r)$-tails."

We can use Lemma 9.6 to show that in Definition 9.4 we may in fact restrict to either only finite intervals or only half-infinite intervals, as shown by the following lemmas.

Lemma 9.8. *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}$ such that $\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}$ for any finite interval $X \subseteq \mathbb{Z}$ and any $r \ge 0$, where $f(r)$ is a positive function with $\lim_{r \to \infty} f(r) = 0$. Then α is an ALPU with $f(r)$-tails.*

Proof. As explained earlier, we can extend α to an automorphism of $\mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$ (denoted again by α). Let $X \subseteq \mathbb{Z}$ be an infinite interval, $r > 0$, and $0 \ne x \in \mathscr{A}_X$. We first show that for any $\delta > 0$,

$$\alpha(x) \overset{(1+\delta)f(r)}{\in} \mathscr{A}_{B(X,r)}. \tag{9.4}$$

By definition of $\mathscr{A}_{\mathbb{Z}}$, we can approximate x with a sequence $x_i \to x$ converging in norm, with $x_i \in \mathscr{A}_{X_i}$, where each $X_i \subseteq X$ is a finite interval. Then $\alpha(x_i) \to \alpha(x)$ converges in norm as well, and

$$\inf_{y \in \mathscr{A}_{B(X,r)}} \|y - \alpha(x)\| \le \liminf_i \left(\inf_{y \in \mathscr{A}_{B(X,r)}} \|y - \alpha(x_i)\| + \|\alpha(x) - \alpha(x_i)\| \right)$$

$$\le \liminf_i \left(f(r)\|x_i\| + \|\alpha(x) - \alpha(x_i)\| \right) = f(r)\|x\|,$$

so for any $\delta' > 0$ there exists some $y \in \mathscr{A}_{B(X,r)}$ such that $\|y - \alpha(x)\| \le f(x)\|x\| + \delta'$. If we apply this with $\delta' = f(x)\delta\|x\|$ then Eq. (9.4) follows. Next, we claim that

$$\alpha(x) \overset{f(r)}{\in} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}. \tag{9.5}$$

Indeed, by Eq. (9.4) we can for any $n > 0$ take some $y_n \in \mathscr{A}_{B(X,r)}$ such that

$$\|\alpha(x) - y_n\| \le \left(1 + \frac{1}{n}\right)f(r).$$

In particular, y_n is a bounded sequence in $\mathscr{A}_{B(X,r)}^{\mathrm{vN}}$. Since norm balls are compact in the weak operator topology, there is a subsequence y_{n_i} converging to some $y \in \mathscr{A}_{B(X,r)}^{\mathrm{vN}}$, and

$$\|\alpha(x) - y\| \le \liminf_i \|\alpha(x) - y_{n_i}\| \le \liminf_i \left(1 + \tfrac{1}{n_i}\right)f(r) = f(r).$$

Thus we have proved Eq. (9.5). As a consequence, we have for any interval $X \subseteq \mathbb{Z}$ and any $r \ge 0$,

$$\alpha(\mathscr{A}_X) \overset{f(r)}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}.$$

Now the lemma follows from Lemma 9.6(iii). ∎

Lemma 9.9. *Suppose α is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ such that $\alpha(\mathscr{A}_{\le n}) \overset{f(r)}{\subseteq} \mathscr{A}_{\le n+r}^{\mathrm{vN}}$ and $\alpha(\mathscr{A}_{\ge n}) \overset{f(r)}{\subseteq} \mathscr{A}_{\ge n-r}^{\mathrm{vN}}$ for any $n \in \mathbb{Z}$ and $r \ge 0$, where $f(r)$ is a positive function with $\lim_{r \to \infty} f(r) = 0$. Then α restricts to an ALPU with $8f(r)$-tails.*

Proof. By (iii) of Lemma 9.6 we only need to show that for any finite interval

$$X = \{n, n+1, \dots, n+m\}$$

it holds that

$$\alpha(\mathscr{A}_X) \overset{8f(r)}{\subseteq} \mathscr{A}_{B(X,r)}^{\mathrm{vN}}.$$

Now, by (i) of Lemma 9.6 we have

$$\alpha(\mathscr{A}_{\le n+m}^{\mathrm{vN}}) \overset{f(r)}{\subseteq} \mathscr{A}_{\le n+m+r}^{\mathrm{vN}},$$

$$\alpha(\mathscr{A}_{\ge n}^{\mathrm{vN}}) \overset{f(r)}{\subseteq} \mathscr{A}_{\ge n-r}^{\mathrm{vN}},$$

hence we obtain by taking commutants and applying Lemma 7.5 with $\mathscr{M} = \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ that

$$\mathscr{A}_{\ge n+m+r+1}^{\mathrm{vN}} \overset{2f(r)}{\subseteq} \alpha(\mathscr{A}_{\le n+m}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} = \alpha(\mathscr{A}_{\ge n+m+1}^{\mathrm{vN}}) \subseteq \alpha(\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}}),$$

$$\mathscr{A}_{\le n-r-1}^{\mathrm{vN}} \overset{2f(r)}{\subseteq} \alpha(\mathscr{A}_{\ge n}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} = \alpha(\mathscr{A}_{\le n-1}^{\mathrm{vN}}) \subseteq \alpha(\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}}).$$

By Eq. (7.17) of Lemma 7.12 it follows that

$$\alpha(\mathscr{A}_X) = \alpha(\mathscr{A}_X^{\mathrm{vN}}) = \alpha(\mathscr{A}_{\mathbb{Z} \setminus X}^{\mathrm{vN}})' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} \overset{8f(r)}{\subseteq} \left(\mathscr{A}_{\ge n+m+r+1}^{\mathrm{vN}} \cup \mathscr{A}_{\le n-r-1}^{\mathrm{vN}}\right)' \cap \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}} = \mathscr{A}_{B(X,r)}^{\mathrm{vN}}.$$

∎

If we consider an ALPU, we may coarse-grain the lattice by grouping together (or 'blocking') sites. This yields again an ALPU, but with faster decaying tails. In particular, for any fixed $\varepsilon > 0$, we can always coarse-grain by sufficiently large blocks of sites so that on the coarse-grained lattice, $\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_{B(X,1)}$ for any interval X. This motivates the following definition:

Definition 9.10 (ε-nearest neighbour automorphism in one dimension)**.** An automorphism α of $\mathscr{A}_{\mathbb{Z}}$ is called ε-*nearest neighbour* for some $\varepsilon \geq 0$ if for any (finite or infinite) interval $X \subseteq \mathbb{Z}$ we have

$$\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_{B(X,1)}. \tag{9.6}$$

If α is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ we instead require the weaker condition that

$$\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_{B(X,1)}^{\mathrm{vN}} \tag{9.7}$$

for all intervals $X \subseteq \mathbb{Z}$. Note that the condition in Eq. (9.7) is equivalent to the condition $\alpha(\mathscr{A}_X^{\mathrm{vN}}) \overset{\varepsilon}{\subseteq} \mathscr{A}_{B(X,1)}^{\mathrm{vN}}$ by Lemma 9.6(i).

If an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ extends an automorphism of $\mathscr{A}_{\mathbb{Z}}$, as will usually be the case for us, then Eqs. (9.6) and (9.7) are equivalent, since $\mathscr{A}_{B(X,1)}^{\mathrm{vN}} \cap \mathscr{A}_{\mathbb{Z}} = \mathscr{A}_{B(X,1)}$ for any $X \subseteq \mathbb{Z}$. As such, Definition 9.10 is unambiguous.

9.2   Index theory of one-dimensional ALPUs

In this section we develop the index theory of ALPUs in one dimension. Just like in the rest of the paper, all ALPUs will be one-dimensional.

For a general ALPU α, we show in Theorems 9.16 and 9.18 that there always exist an approximation of α by a sequence of QCAs β_j. We can use the limit of the indices of the latter (which become stationary for large j) as the definition of the index of α. If α has $\mathcal{O}(r^{-(1+\delta)})$-tails for some $\delta > 0$, we further show that this index can be computed as a difference of mutual informations,

$$\mathrm{ind}(\alpha) = \frac{1}{2} \left(I(L' : R)_\phi - I(L : R')_\phi \right), \tag{9.8}$$

with both terms being finite, just like we saw in Eq. (8.11) for QCAs. The local computation of the index in (8.12) does not yield the exact index for ALPUs. However, the exact index can still be computed locally; we show that on sufficiently large regions, the local index computation gives the exact answer when rounded to the nearest value in the fixed set of discrete index values.

In the remainder of the section, we discuss the properties of this index. We find that once circuits are replaced by evolutions by time-dependent Hamiltonians, the results of [GNVW12] stated in Theorem 8.5 generalize in a natural way. Our results are summarized in Theorem 9.25.

9.2.1   Approximating an ALPU by a QCA

We sketch the general strategy for approximating an ALPU α by a QCA. We first develop a method for deforming α into an ALPU α_n that behaves as a QCA with a strict causal cone in the proximity of the site n, exhibited by Proposition 9.14 and Fig. 9.2. In Proposition 9.15 we then we stitch the different α_n together into a QCA β using the structure theory for one-dimensional QCAs, obtaining a QCA approximation to α. If we apply this result to increasingly coarse-grained lattices, in Theorem 9.16 we obtain a sequence of QCAs of increasing radius that approximate α with increasing accuracy.

   To achieve Proposition 9.14 localizing an ALPU α on a local patch, we compose α with a sequence of unitary rotations. Some individual rotation steps are described by Lemma 9.11 and Lemma 9.13, with proof illustrated in Fig. 9.1. Each step uses Theorem 7.6 to rotate nearby subalgebras, e.g. rotating an algebra $\alpha(\mathscr{A}_X) \overset{\varepsilon}{\subseteq} \mathscr{A}_Y$ to obtain an exact inclusion. We start with these two lemmas. Lemma 9.11, Lemma 9.13 and Proposition 9.14 are each divided into two parts, (i) and (ii). In each case, part (i) is valid for ε-nearest neighbour automorphisms (which need not be ALPUs), while part (ii) gives a more refined statement when assuming an ALPU as input. For the majority of the further development in this paper, in fact only the parts (i) will be necessary, and so the first-time reader may wish to skip part (ii) of these results, as well as the supporting Lemma 9.12. Those parts will only be necessary for later results about blending, following Definition 9.23.

Lemma 9.11.

(i) *There exist universal constants $C_0, \varepsilon_0 > 0$ such that if α is an ε-nearest neighbour automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ with $\varepsilon \le \varepsilon_0$ and*

$$\alpha(\mathscr{A}_{\ge n}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\ge n-1}^{\mathrm{vN}}$$

for some site $n \in \mathbb{Z}$, then there exists an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$, which is of the form $\tilde{\alpha}(x) = u^* \alpha(x) u$ for some unitary $u \in \mathscr{A}_{\ge n-1}^{\mathrm{vN}}$ with $\|u - I\| \le C_0 \varepsilon$ and

$$\tilde{\alpha}(\mathscr{A}_{\le n-1}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\le n}^{\mathrm{vN}}, \tag{9.9}$$

$$\tilde{\alpha}(\mathscr{A}_{\ge n}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\ge n-1}^{\mathrm{vN}}. \tag{9.10}$$

(ii) *If additionally α is an ALPU with $f(r)$-tails, we can take u such that $\tilde{\alpha}$ is an ALPU with $\mathscr{O}(f(r-1))$-tails and such that we have, for $r \to \infty$,*

$$\|(\alpha - \tilde{\alpha})|_{\mathscr{A}_{\le n-r-1}^{\mathrm{vN}}}\| = \mathscr{O}(f(r)), \tag{9.11}$$

$$\|(\alpha - \tilde{\alpha})|_{\mathscr{A}_{\ge n+r}^{\mathrm{vN}}}\| = \mathscr{O}(f(r-1)) \tag{9.12}$$

and, for all $x \in \mathscr{A}_{\ge n+r+1}$,

$$\|u^* x u - x\| = \mathscr{O}(f(r)\|x\|). \tag{9.13}$$

Proof. (i) Note α^{-1} is 4ε-nearest neighbour by (ii) of Lemma 9.6. Thus

$$\alpha^{-1}(\mathscr{A}_{\ge n+1}^{\mathrm{vN}}) \overset{4\varepsilon}{\subseteq} \mathscr{A}_{\ge n}^{\mathrm{vN}}$$

and then $\mathscr{A}^{\mathrm{vN}}_{\geq n+1} \overset{4\varepsilon}{\subseteq} \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})$. By Theorem 7.6 with $\mathscr{A}_0 = \mathscr{A}_{\geq n+1}$, $\mathscr{A} = \mathscr{A}''_0 = \mathscr{A}^{\mathrm{vN}}_{\geq n+1}$ and with $\mathscr{B} = \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})$, provided that $\varepsilon \leq \frac{1}{256}$, there exists a unitary $u \in (\mathscr{A}^{\mathrm{vN}}_{\geq n+1} \cup \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}))''$ such that

$$u\mathscr{A}^{\mathrm{vN}}_{\geq n+1}u^* \subseteq \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})$$

and $\|u - I\| \leq 48\varepsilon$. Because $\alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}) \subseteq \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$, we also have $u \in \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$. We define a new automorphism $\tilde{\alpha}(x) = u^*\alpha(x)u$ that satisfies $\tilde{\alpha}(\mathscr{A}^{\mathrm{vN}}_{\geq n}) \supseteq \mathscr{A}^{\mathrm{vN}}_{\geq n+1}$, and then satisfies (9.9) by taking commutants. Moreover,

$$\tilde{\alpha}(\mathscr{A}^{\mathrm{vN}}_{\geq n}) = u^*\alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})u \subseteq u^*\mathscr{A}^{\mathrm{vN}}_{\geq n-1}u = \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$$

using the assumption $\alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}) \subseteq \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$ and the fact $u \in \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$. Then $\tilde{\alpha}$ also satisfies (9.10).

(ii) Now we further assume α is an ALPU with $f(r)$-tails and show (9.11). By our use of Theorem 7.6 to construct $u \in (\mathscr{A}^{\mathrm{vN}}_{\geq n+1} \cup \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}))''$, we know that for $x \in \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$,

$$\|[x, y]\| \leq \delta\|x\|\|y\| \quad \forall y \in \mathscr{A}^{\mathrm{vN}}_{\geq n+1} \cup \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}) \implies \|u^*xu - x\| = \mathcal{O}(\delta\|x\|).$$

For $r \geq 0$ and $x \in \alpha(\mathscr{A}^{\mathrm{vN}}_{\leq n-r-1})$, the above condition is satisfied for $\delta = 2f(r+1)$ because we have $x \overset{f(r+1)}{\in} \mathscr{A}^{\mathrm{vN}}_{\leq n}$ (using Lemma 7.4) and $x \in \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})'$, so

$$\|(\alpha - \tilde{\alpha})|_{\mathscr{A}^{\mathrm{vN}}_{\leq n-r-1}}\| = \mathcal{O}(f(r+1))$$

and Eq. (9.11) follows.

Our application of Theorem 7.6 also implies that for $x \in \mathscr{A}^{\mathrm{vN}}_{\mathbb{Z}}$,

$$x \overset{\delta}{\in} \mathscr{A}_{\geq n+1} \text{ and } x \overset{\delta}{\in} \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n}) \implies \|u^*xu - x\| = \mathcal{O}(\delta\|x\|). \tag{9.14}$$

For $r \geq 1$ and $x \in \alpha(\mathscr{A}_{\geq n+r})$, those conditions are satisfied for $\delta = f(r-1)$ because we have $x \overset{f(r-1)}{\in} \mathscr{A}_{\geq n+1}$ and $x \in \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})$, so

$$\|(\alpha - \tilde{\alpha})|_{\mathscr{A}_{\geq n+r}}\| = \mathcal{O}(f(r-1))$$

and hence Eq. (9.12) follows.

Next, we prove Eq. (9.13). Recall that by Lemma 9.6(iv), α^{-1} is also an ALPU, with $\mathcal{O}(f(r))$-tails. Therefore, for any $r \geq 0$ and $x \in \mathscr{A}_{\geq n+r+1}$ we have $\alpha^{-1}(x) \overset{f(r+1)}{\in} \mathscr{A}_{\geq n}$, hence $x \overset{f(r+1)}{\in} \alpha(\mathscr{A}^{\mathrm{vN}}_{\geq n})$, and now Eq. (9.14) shows that $\|u^*xu - x\| = \mathcal{O}(f(r+1)\|x\|)$ and Eq. (9.13) follows.

Finally we show that $\tilde{\alpha}$ is an ALPU with $\mathcal{O}(f(r-1))$-tails. To this end we apply Lemma 9.12 below and the fact that

$$\|uxu^* - x\| = \mathcal{O}(f(r-1)\|x\|)$$

holds for the following x and all $r \geq 0$: for $x \in \mathscr{A}_{\leq n-r-2}$ since $u \in \mathscr{A}^{\mathrm{vN}}_{\geq n-1}$, for $x \in \mathscr{A}_{\geq n+r+1}$ by Eq. (9.13), for $x \in \alpha(\mathscr{A}_{\leq n-r-1})$ by Eq. (9.11) and for $x \in \alpha(\mathscr{A}_{\geq n+r})$ by Eq. (9.12). ∎

The following lemma is used in the proof above (and in similar proofs below) that the construction gives rise to an ALPU when the input is an ALPU.
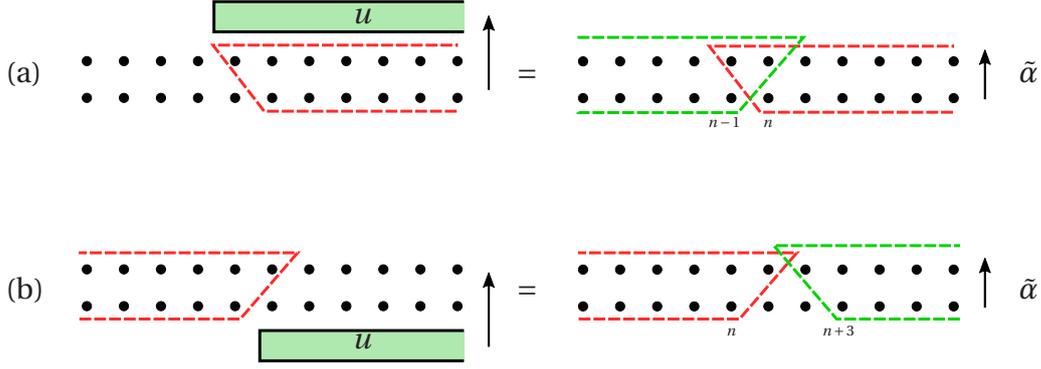
Figure 9.1: (a) Illustration of the construction in Lemma 9.11. The dashed lines indicate causal cones. (b) Analogous illustration of Lemma 9.13.

Lemma 9.12. *Suppose that α is an ALPU with $f(r)$-tails and $\tilde{\alpha}$ is an automorphism of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ of the form $\tilde{\alpha}(x) = u^* \alpha(x) u$ for some $u \in \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ and all $x \in \mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$, which satisfies*

$$\tilde{\alpha}(\mathscr{A}_{\leq n-1}) \subseteq \mathscr{A}_{\leq n}^{\mathrm{vN}},$$
$$\tilde{\alpha}(\mathscr{A}_{\geq n}) \subseteq \mathscr{A}_{\geq n-1}^{\mathrm{vN}} \tag{9.15}$$

for some site $n \in \mathbb{Z}$. If for any $r \geq 0$ and $x \in \mathscr{A}_{\leq n-r-2} \cup \mathscr{A}_{\geq n+r+1} \cup \alpha(\mathscr{A}_{\leq n-r-1}) \cup \alpha(\mathscr{A}_{\geq n+r})$ we have

$$\|u^* x u - x\| \leq g(r)\|x\|,$$

where $g(r)$ is non-increasing with $\lim_{r \to \infty} g(r) = 0$, then $\tilde{\alpha}$ is an ALPU with $\mathcal{O}(f(r)+g(r))$-tails.

Proof. We abbreviate $h(r) = f(r) + g(r)$. By Lemma 9.9, it suffices to show

$$\tilde{\alpha}(\mathscr{A}_{\leq m}) \overset{\mathcal{O}(h(r))}{\subseteq} \mathscr{A}_{\leq m+r}^{\mathrm{vN}} \quad \text{and} \quad \tilde{\alpha}(\mathscr{A}_{\geq m}) \overset{\mathcal{O}(h(r))}{\subseteq} \mathscr{A}_{\geq m-r}^{\mathrm{vN}}$$

for all $m \in \mathbb{Z}$ and $r \geq 0$. We only prove the former, since the proof of the latter proceeds analogously. We distinguish two cases:

- $m < n$: Then $m = n - k - 1$ for some $k \geq 0$. By assumption, $\tilde{\alpha}(\mathscr{A}_{\leq n-k-1}) \subseteq \mathscr{A}_{\leq n}^{\mathrm{vN}}$, so it remains to show

$$\tilde{\alpha}(\mathscr{A}_{\leq n-k-1}) \overset{\mathcal{O}(h(r))}{\subseteq} \mathscr{A}_{\leq n-k-1+r}^{\mathrm{vN}}.$$

for $0 \leq r \leq k$. This holds since, by assumption, $\|u^* x u - x\| \leq g(k)\|x\|$ for all $x \in \alpha(\mathscr{A}_{\leq n-k-1})$, and hence $\|(\tilde{\alpha} - \alpha)|_{\mathscr{A}_{\leq n-k-1}}\| \leq g(k) \leq g(r)$ for any $0 \leq r \leq k$.

- $m \geq n$: Then $m = n + k$ for some $k \geq 0$. To prove that

$$\tilde{\alpha}(\mathscr{A}_{\leq n+k}) \overset{\mathcal{O}(h(r))}{\subseteq} \mathscr{A}_{\leq n+k+r}^{\mathrm{vN}}$$

by Lemma 7.4 it suffices to show that for all $x \in \mathscr{A}_{\leq n+k}$ and for all $y \in \mathscr{A}_{\geq n+k+r+1}^{\mathrm{vN}}$,

$$\|[\tilde{\alpha}(x), y]\| = \mathcal{O}(h(r))\|x\|\|y\|.$$

By assumption

$$\|uyu^* - y\| = \|u^* yu - y\| \le g(k+r)\|y\| \le g(r)\|y\|$$

for all $y \in \mathscr{A}_{\ge n+k+r+1}$ and thus for all $\mathscr{A}^{vN}_{\ge n+k+r+1}$, and hence we indeed have that

$$\|[\tilde{\alpha}(x), y]\| = \|[\alpha(x), uyu^*]\| \le \|[\alpha(x), y]\| + \|[\alpha(x), uyu^* - y]\| = O(h(r))\|x\|\|y\|,$$

using that $\|[\alpha(x), y]\| \le 2f(r)\|x\|\|y\|$ by Lemma 7.4, since α is an ALPU with $f(r)$-tails. ∎

Lemma 9.13.

(i) *There exist universal constants $C_0', \varepsilon_0' > 0$ such that if α is an ε-nearest neighbour automorphism of $\mathscr{A}^{vN}_{\mathbb{Z}}$ with $\varepsilon \le \varepsilon_0'$ and*

$$\alpha(\mathscr{A}^{vN}_{\le n}) \subseteq \mathscr{A}^{vN}_{\le n+1}$$

for some site $n \in \mathbb{Z}$, then there exists an automorphism of $\mathscr{A}^{vN}_{\mathbb{Z}}$ of the form $\tilde{\alpha}(x) = \alpha(uxu^*)$ for some unitary $u \in \mathscr{A}^{vN}_{\ge n+1}$ with $\|u - I\| \le C_0'\varepsilon$ and

$$\tilde{\alpha}(\mathscr{A}^{vN}_{\ge n+3}) \subseteq \mathscr{A}^{vN}_{\ge n+2}, \tag{9.16}$$
$$\tilde{\alpha}(\mathscr{A}^{vN}_{\le n}) \subseteq \mathscr{A}^{vN}_{\le n+1}. \tag{9.17}$$

In fact, it holds that $\alpha|_{\mathscr{A}^{vN}_{\le n}} = \tilde{\alpha}|_{\mathscr{A}^{vN}_{\le n}}$.

(ii) *If additionally α is an ALPU with $f(r)$-tails, we can take u such that $\tilde{\alpha}$ is an ALPU with $\mathscr{O}(f(r-1))$-tails and such that for $r \to \infty$,*

$$\|(\alpha - \tilde{\alpha})|_{\mathscr{A}^{vN}_{\ge n+r+3}}\| = \mathscr{O}(f(r)). \tag{9.18}$$

Proof. (i) This follows by application of Lemma 9.11 to $\beta = \alpha^{-1}$. Here we use that if α is an ε-nearest neighbour automorphism of $\mathscr{A}^{vN}_{\mathbb{Z}}$, then β is a 4ε-nearest neighbour automorphism by Lemma 9.6(ii). Now, $\alpha(\mathscr{A}^{vN}_{\le n}) \subseteq \mathscr{A}^{vN}_{\le n+1}$ implies that $\mathscr{A}^{vN}_{\le n} \subseteq \beta(\mathscr{A}^{vN}_{\le n+1})$ and hence $\beta(\mathscr{A}^{vN}_{\ge n+2}) \subseteq \mathscr{A}^{vN}_{\ge n+1}$. Let $\varepsilon_0' := \varepsilon_0/4$ and $C_0' = 4C_0$ for $C_0, \varepsilon_0 > 0$ the constants from Lemma 9.11. Thus we may apply Lemma 9.11 to β (with $n+2$ in place of n) to find an automorphism $\tilde{\beta}$ of $\mathscr{A}^{vN}_{\mathbb{Z}}$ that is of the form $\tilde{\beta}(x) = u^*\beta(x)u$ for a unitary $u \in \mathscr{A}^{vN}_{\ge n+1}$ with $\|u - I\| \le 4C_0\varepsilon = C_0'\varepsilon$ and which satisfies

$$\tilde{\beta}(\mathscr{A}^{vN}_{\le n+1}) \subseteq \mathscr{A}^{vN}_{\le n+2},$$
$$\tilde{\beta}(\mathscr{A}^{vN}_{\ge n+2}) \subseteq \mathscr{A}^{vN}_{\ge n+1}.$$

We then see that $\tilde{\alpha} = \tilde{\beta}^{-1}$ is given by $\tilde{\alpha}(x) = \alpha(uxu^*)$ and satisfies the desired properties in (i). In particular, note that $u \in \mathscr{A}^{vN}_{\ge n+1}$ immediately implies that $\alpha|_{\mathscr{A}^{vN}_{\le n}} = \tilde{\alpha}|_{\mathscr{A}^{vN}_{\le n}}$.

(ii) If α is an ALPU with $f(r)$-tails, then by Lemma 9.6(iv) β is an ALPU with $\mathscr{O}(f(r))$-tails. Hence by part (ii) of Lemma 9.11, $\tilde{\beta}$ is an ALPU with $\mathscr{O}(f(r-1))$-tails and thus the same is true for $\tilde{\alpha}$, again by Lemma 9.6(iv). Eq. (9.18) follows since by Eq. (9.13) we have

$$\|uxu^* - x\| = \|u^* xu - x\| = \mathscr{O}(f(r)\|x\|)$$

for all $x \in \mathscr{A}_{\ge n+r+3}$. ∎
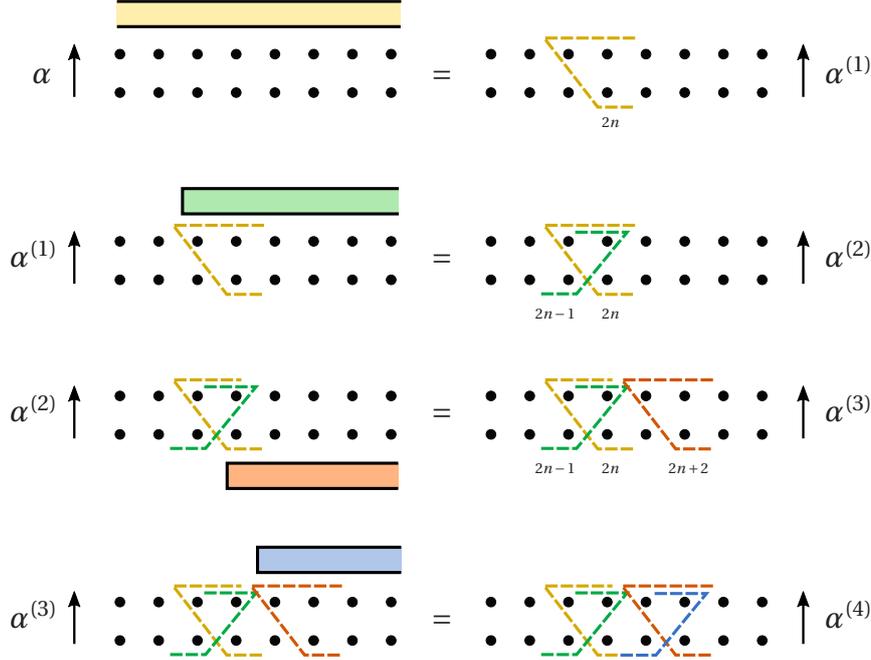
Figure 9.2: Illustration of the construction of $\alpha^{(i)}$ for $i = 1, 2, 3, 4$ in the proof of Proposition 9.14. Each row depicts an equation, and the solid strips on the left depict applications of unitary operators supported on those regions. The second, third, and fourth row use Lemma 9.11, Lemma 9.13, and Lemma 9.11 again respectively. Dashed lines indicate causal cones.

We iteratively apply Lemma 9.11 and Lemma 9.13 to show that for an ε-nearest neighbour automorphism, for any small patch, one can find a nearby $\mathcal{O}(\varepsilon)$-nearest neighbour automorphism that is strictly local on that patch. Below we work with a patch near site $2n$, and the modified automorphism is denoted α_n.

Proposition 9.14.

(i) *There exist universal constants $C_1, \varepsilon_1 > 0$ such that for any ε-nearest neighbour automorphism α of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ with $\varepsilon \leq \varepsilon_1$ and for any site $n \in \mathbb{Z}$, there exists an automorphism α_n of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ such that for $k \in \{0, 1, 2, 3\}$,*

$$\alpha_n(\mathscr{A}_{\leq 2n+2k-1}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\leq 2n+2k}^{\mathrm{vN}},$$
$$\alpha_n(\mathscr{A}_{\geq 2n+2k}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\geq 2n+2k-1}^{\mathrm{vN}},$$
$$\|\alpha_n - \alpha\| \leq C_1 \varepsilon.$$

In particular, denoting $\mathscr{B}_m = \mathscr{A}_{\{2m, 2m+1\}}$ and $\mathscr{C}_m = \mathscr{A}_{\{2m-1, 2m\}}$ as in Eq. (8.1), we have

$$\alpha_n(\mathscr{B}_m) \subseteq \mathscr{C}_m \otimes \mathscr{C}_{m+1} \quad \text{for } m \in \{n, n+1, n+2\},$$
$$\alpha_n^{-1}(\mathscr{C}_m) \subseteq \mathscr{B}_{m-1} \otimes \mathscr{B}_m \quad \text{for } m \in \{n+1, n+2\}.$$

(ii) *Moreover, if α is an ALPU with $f(r)$-tails, we may take α_n to be an ALPU with $\mathcal{O}(f(r-7))$-tails and such that, for $r \to \infty$,*

$$\|(\alpha - \alpha_n)|_{\mathscr{A}_{\leq 2n-r-1}^{\mathrm{vN}}}\| = \mathcal{O}(f(r-1)), \tag{9.19}$$
$$\|(\alpha - \alpha_n)|_{\mathscr{A}_{\geq 2n+r+6}^{\mathrm{vN}}}\| = \mathcal{O}(f(r-7)). \tag{9.20}$$

Proof. (i) We define a sequence of automorphisms $\alpha_n^{(i)}$, $i = 1, \ldots, 8$ to obtain $\alpha_n := \alpha_n^{(8)}$ with the desired properties. To begin,

$$\alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) \overset{\varepsilon}{\subseteq} \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$$

so by Theorem 7.6, where we let $\mathscr{A}_0 = \alpha(\mathscr{A}_{\geq 2n})$, $\mathscr{A} = \mathscr{A}_0'' = \alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$ and $\mathscr{B} = \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$, there exists $u_1 \in (\alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) \cup \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}})''$ such that

$$u_1^* \alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) u_1 \subseteq \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$$

with $\|u_1^* - I\| \leq 12\varepsilon$. We define $\alpha_n^{(1)}(x) = u_1^* \alpha(x) u_1$, which by construction satisfies

$$\alpha_n^{(1)}(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}},$$
$$\|\alpha - \alpha_n^{(1)}\| = \mathscr{O}(\varepsilon).$$

Then $\alpha_n^{(1)}$ is an $\mathscr{O}(\varepsilon)$-nearest neighbour automorphism by the above, and we are in a situation where we can apply Lemma 9.11 (but replacing n with $2n$) to obtain an automorphism $\alpha_n^{(2)}(x) = u_2^* \alpha_n^{(1)}(x) u_2$ for unitary $u_2 \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$, such that

$$\alpha_n^{(2)}(\mathscr{A}_{\leq 2n-1}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\leq 2n}^{\mathrm{vN}},$$
$$\alpha_n^{(2)}(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}},$$
$$\|\alpha_n^{(1)} - \alpha_n^{(2)}\| = \mathscr{O}(\varepsilon).$$

Then $\alpha_n^{(2)}$ is again an $\mathscr{O}(\varepsilon)$-nearest neighbour automorphism, and hence we can apply Lemma 9.13 (but replacing n with $2n-1$) to obtain a new automorphism $\alpha_n^{(3)}$ defined by $\alpha_n^{(3)}(x) = \alpha_n^{(2)}(u_3 x u_3^*)$ for unitary $u_3 \in \mathscr{A}_{\geq 2n}^{\mathrm{vN}}$, such that

$$\alpha_n^{(3)}(\mathscr{A}_{\geq 2n+2}^{\mathrm{vN}}) \subseteq \mathscr{A}_{\geq 2n+1}^{\mathrm{vN}},$$
$$\|\alpha_n^{(3)} - \alpha_n^{(2)}\| = \mathscr{O}(\varepsilon).$$

Since $u_3 \in \mathscr{A}_{\geq 2n}^{\mathrm{vN}}$, $\alpha_n^{(3)}$ also satisfies the locality properties listed for $\alpha_n^{(2)}$ above. See Fig. 9.2 for an illustration of the construction.

We continue to apply Lemma 9.11 and Lemma 9.13 alternatingly. Explicitly, we apply Lemma 9.11 (with $n \to 2n+2$) to define $\alpha_n^{(4)}$, as illustrated in the figure, and then Lemma 9.13 (with $n \to 2n+1$) to define $\alpha_n^{(5)}$, followed by Lemma 9.11 (with $n \to 2n+4$) to define $\alpha_n^{(6)}$ and Lemma 9.13 (with $n \to 2n+3$) to define $\alpha_n^{(7)}$. Finally we use Lemma 9.11 (with $n \to 2n+6$) to obtain $\alpha_n^{(8)}$. We take $\alpha_n := \alpha_n^{(8)}$; then α_n has the desired locality properties in the proposition statement. We must assume ε is sufficiently small to meet the conditions of these lemmas at each step, determining the universal constant ε_1 in the proposition statement.

(ii) Now we further assume α is an ALPU with $f(r)$-tails, to demonstrate Eqs. (9.19) and (9.20) and prove that α_n is an ALPU.

We first show that $\alpha_n^{(2)}$ is an ALPU with $\mathscr{O}(f(r-1))$ tails, using Lemma 9.12 (where we take $n \to 2n$). Note that $\alpha_n^{(2)}(x) = v^* \alpha(x) v$ for $v = u_1 u_2$, and $\alpha_n^{(2)}$ satisfies the necessary locality properties in Eq. (9.15) (unlike $\alpha_n^{(1)}$!), so in order to apply the lemma we only need to show that

$$\|v^* x v - x\| = \mathscr{O}(f(r-1))\|x\|, \tag{9.21}$$

for all $r \geq 0$ and $x \in \mathscr{A}_{\leq 2n-r-2} \cup \mathscr{A}_{\geq 2n+r+1} \cup \alpha(\mathscr{A}_{\leq 2n-r-1}) \cup \alpha(\mathscr{A}_{\geq 2n+r})$. To this end, recall that for u_1 we applied Theorem 7.6 using $\mathscr{A}_0 = \alpha(\mathscr{A}_{\geq 2n})$, $\mathscr{A} = \mathscr{A}_0'' = \alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$ and with $\mathscr{B} = \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$, and in the construction of $u_2 \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$ in Lemma 9.11 (with $n \mapsto 2n$) we applied Theorem 7.6 with $\mathscr{A}_0 = \mathscr{A}_{\geq 2n+1}$, $\mathscr{A} = \mathscr{A}_0'' = \mathscr{A}_{\geq 2n+1}^{\mathrm{vN}}$ and $\mathscr{B} = \alpha_n^{(1)}(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$.

First consider $x \in \mathscr{A}_{\leq 2n-r-2}$. As $\alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}}) \overset{f(r+1)}{\subseteq} \mathscr{A}_{\geq 2n-r-1}^{\mathrm{vN}}$, we have

$$\|[x,y]\| = 2f(r+1)\|x\|\|y\|$$

for all $y \in \alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$ by Lemma 7.4. Since moreover $[x,y] = 0$ for all $y \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$, Theorem 7.6(ii) shows that $\|u_1^* x u_1 - x\| = \mathscr{O}(f(r+1)\|x\|)$. In addition, we have $u_2^* x u_2 = x$ since $u_2 \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$. Together we find that

$$\|v^* x v - x\| = \mathscr{O}(f(r+1)\|x\|).$$

Next consider $x \in \mathscr{A}_{\geq 2n+r+1}$. By Lemma 9.6, α^{-1} is an ALPU with $\mathscr{O}(f(r))$-tails, so we have

$$\alpha^{-1}(x) \overset{\mathscr{O}(f(r+1))}{\in} \mathscr{A}_{\geq 2n} \tag{9.22}$$

and hence $x \overset{\mathscr{O}(f(r+1))}{\in} \alpha(\mathscr{A}_{\geq 2n})$. Since moreover $x \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$, Theorem 7.6(iii) shows that $\|u_1^* x u_1 - x\| = \mathscr{O}(f(r+1)\|x\|)$. Since $(\alpha_n^{(1)})^{-1}(x) = \alpha^{-1}(u_1 x u_1^*)$, the latter along with Eq. (9.22) in turn implies that $x \overset{\mathscr{O}(f(r+1))}{\in} \alpha_n^{(1)}(\mathscr{A}_{\geq 2n})$. Also, $x \in \mathscr{A}_{\geq 2n+1}$, hence we obtain $\|u_2^* x u_2 - x\| = \mathscr{O}(f(r+1))$, again by Theorem 7.6(iii). Together we find that $\|v^* x v - x\| = \mathscr{O}(f(r+1)\|x\|)$.

Now consider $x \in \alpha(\mathscr{A}_{\leq 2n-r-1})$, i.e., $x = \alpha(z)$ for some $z \in \mathscr{A}_{\leq 2n-r-1}$. Then x commutes with $\alpha(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$. Moreover, $x \overset{f(r-1)}{\in} \mathscr{A}_{\leq 2n-2}$, hence $\|[x,y]\| \leq 2f(r-1)\|x\|\|y\|$ for all $y \in \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$. Thus we obtain $\|u_1^* x u_1 - x\| = \mathscr{O}(f(r-1)\|x\|)$ by Theorem 7.6(ii). The preceding in turn implies that for all $y \in \mathscr{A}_{\geq 2n+1}^{\mathrm{vN}}$,

$$\|[\alpha_n^{(1)}(z), y]\| = \|[u_1^* x u_1, y]\| \leq 2\|u_1^* x u_1 - x\|\|y\| + \|[x, y]\| = \mathscr{O}(f(r-1)\|x\|\|y\|).$$

Also, $\alpha_n^{(1)}(z)$ commutes with $\alpha_n^{(1)}(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$. Therefore, again by Theorem 7.6(ii) we see that

$$\|v^* x v - u_1^* x u_1\| = \|u_2^* \alpha_n^{(1)}(z) u_2 - \alpha_n^{(1)}(z)\| = \mathscr{O}(f(r-1)\|x\|).$$

We conclude that $\|v^* x v - x\| = \mathscr{O}(f(r-1)\|x\|)$.

Finally, let $x \in \alpha(\mathscr{A}_{\geq 2n+r})$, i.e., $x = \alpha(z)$ for some $z \in \mathscr{A}_{\geq 2n+r}$. Then $x \in \alpha(\mathscr{A}_{\geq 2n})$ and $x \overset{f(r+1)}{\in} \mathscr{A}_{\geq 2n-1}^{\mathrm{vN}}$. So, by Theorem 7.6(iii) we find that

$$\|u_1^* x u_1 - x\| = \mathscr{O}(f(r+1)\|x\|).$$

Using the latter, as well as $\|\alpha_n^{(1)}(z) - \alpha(z)\| = \|u_1^* x u_1 - x\|$ and $\alpha(z) \overset{f(r-1)}{\in} \mathscr{A}_{\geq 2n+1}$, we find that $\alpha_n^{(1)}(z) \overset{\mathscr{O}(f(r-1))}{\in} \mathscr{A}_{\geq 2n+1}$. Moreover, $\alpha_n^{(1)}(z) \in \alpha_n^{(1)}(\mathscr{A}_{\geq 2n}^{\mathrm{vN}})$, so by Theorem 7.6(iii) we obtain that

$$\|v^* x v - u_1^* x u_1\| = \|u_2^* \alpha_n^{(1)}(z) u_2 - \alpha_n^{(1)}(z)\| = \mathscr{O}(f(r-1)\|x\|),$$

and hence $\|v^* x v - x\| = \mathcal{O}(f(r-1)\|x\|)$. Altogether we have verified that Eq. (9.21) holds for all $r \geq 0$ and $x \in \mathscr{A}_{\leq 2n-r-2} \cup \mathscr{A}_{\geq 2n+r+1} \cup \alpha(\mathscr{A}_{\leq 2n-r-1}) \cup \alpha(\mathscr{A}_{\geq 2n+r})$. We may therefore apply Lemma 9.12 and conclude that $\alpha_n^{(2)}$ is an ALPU with $\mathcal{O}(f(r-1))$-tails.

For $i = 3, \dots, 8$, we simply observe that by our applications of Lemma 9.11 and Lemma 9.13, the automorphisms $\alpha_n^{(i)}$ are guaranteed to be APLUs with $\mathcal{O}(f(r+1-i))$-tails.

To see that Eq. (9.19) holds, note that Eq. (9.21) implies that

$$\|\alpha_n^{(2)}(x) - \alpha(x)\| = \mathcal{O}(f(r-1)\|x\|)$$

for all $x \in \mathscr{A}_{\leq 2n-r-1}$. Moreover, $\alpha_n(x) = \alpha_n^{(2)}(x)$ for such x, since $\alpha_n = \alpha_n^{(8)}$ is obtained from $\alpha_n^{(2)}$ by conjugating the input with unitaries in $\mathscr{A}_{\geq 2n}^{\mathrm{vN}}$ (leaving x unchanged) and the output by unitaries in $\mathscr{A}_{\geq 2n+1}^{\mathrm{vN}}$ (leaving $\alpha_n^{(2)}(x) \in \mathscr{A}_{\leq 2n}^{\mathrm{vN}}$ unchanged). Thus Eq. (9.19) follows.

Finally, Eq. (9.20) follows since the $\alpha_n^{(i)}$ for $i = 3, \dots, 8$ satisfy analogs of Eqs. (9.12) and (9.18) and we have $\|\alpha_n^{(2)}(x) - \alpha(x)\| \leq \mathcal{O}(f(r-1)\|x\|)$ for all $x \in \mathscr{A}_{\geq 2n+r}$, again by Eq. (9.21). ∎

Proposition 9.15 (QCA approximation of ε-nearest neighbour automorphism)**.** *There exists a universal constant $\varepsilon_2 > 0$ such that if α is an ε-nearest neighbour automorphism of $\mathscr{A}_\mathbb{Z}$ with $\varepsilon \leq \varepsilon_2$, then there exists a QCA β with radius 2 such that*

$$\|(\alpha - \beta)|_{\mathscr{A}_X}\| = \mathcal{O}(\varepsilon|X|)$$

for all regions X with $|X|$ sites.

Proof. Recall that α can be extended to a ε-nearest neighbour automorphism of $\mathscr{A}_\mathbb{Z}^{\mathrm{vN}}$ by Lemma 9.6, which we will denote by the same symbol. Let C_1 and ε_1 be the constants from Proposition 9.14, and take $\varepsilon_2 := \min\{\frac{\varepsilon_1}{2}, \frac{1}{384}C_1\}$. As usual, we write $\mathscr{B}_n = \mathscr{A}_{\{2n,2n+1\}}$ and $\mathscr{C}_n = \mathscr{A}_{\{2n-1,2n\}}$. Now apply part (i) of Proposition 9.14 to find a collection of automorphisms α_m for each $m \in \mathbb{Z}$, which satisfy the locality properties

$$\alpha_m(\mathscr{B}_n) \subseteq \mathscr{C}_n \otimes \mathscr{C}_{n+1}$$

for $n \in \{m, m+1, m+2\}$ as well as $\alpha_m^{-1}(\mathscr{C}_n) \subseteq \mathscr{B}_{n-1} \otimes \mathscr{B}_n$ for $n \in \{m+1, m+2\}$. Then by Theorem 8.2 and the subsequent Remark 8.3, we can define

$$\mathscr{L}_n^{(m)} = \alpha_m(\mathscr{B}_n) \cap \mathscr{C}_n,$$
$$\mathscr{R}_{n-1}^{(m)} = \alpha_m(\mathscr{B}_{n-1}) \cap \mathscr{C}_n$$

such that, for $m \in \{n-1, n-2\}$,

$$\mathscr{C}_n = \mathscr{L}_n^{(m)} \otimes \mathscr{R}_{n-1}^{(m)}. \tag{9.23}$$

Moreover, again by Theorem 8.2 and Remark 8.3, we have

$$\mathscr{B}_n = \alpha_{n-1}^{-1}(\mathscr{L}_n^{(n-1)}) \otimes \alpha_{n-1}^{-1}(\mathscr{R}_n^{(n-1)}), \tag{9.24}$$

which we will use below.

Note that $\|\alpha_{n-1} - \alpha_{n-2}\| \leq \|\alpha_{n-1} - \alpha\| + \|\alpha_{n-2} - \alpha\| \leq 2C_1\varepsilon \leq \frac{1}{192}$. Because α_{n-1} and α_{n-2} are nearby ALPUs with locality properties satisfying Remark 8.12, we can apply

the argument from Proposition 8.11 to α_{n-1} and α_{n-2}, finding that $\mathscr{L}_n^{(n-2)}$ and $\mathscr{L}_n^{(n-1)}$ are related by a unitary $u_n \in \mathscr{C}_n$, i.e. $u_n \mathscr{L}_n^{(n-1)} u_n^* = \mathscr{L}_n^{(n-2)}$, with $\|u_n - I\| = \mathcal{O}(\varepsilon)$. Finally we define

$$\beta_n \colon \mathscr{B}_n \to \mathscr{C}_n \otimes \mathscr{C}_{n+1}, \quad \beta_n(x) = u_n \alpha_{n-1}(x) u_n^*.$$

Each β_n is an injective homomorphism and by (9.24) we obtain

$$\beta_n(\mathscr{B}_n) = u_n \left(\mathscr{L}_n^{(n-1)} \otimes \mathscr{R}_n^{(n-1)} \right) u_n^* = \mathscr{L}_n^{(n-2)} \otimes \mathscr{R}_n^{(n-1)}, \tag{9.25}$$

where the second equality holds because $u_n \in \mathscr{C}_n$ and $\mathscr{R}_n^{(n-1)} \subseteq \mathscr{C}_{n+1}$ commute. From Eq. (9.25) we conclude that $\beta_n(\mathscr{B}_n)$ and $\beta_m(\mathscr{B}_m)$ commute for $n \neq m$. Hence we can define a global injective homomorphism β that acts as β_n on each $\mathscr{B}_n$. By (9.23) and (9.25), this homomorphism is surjective. Indeed,

$$\beta_{n-1}(\mathscr{B}_{n-1}) \otimes \beta_n(\mathscr{B}_n) \supseteq \mathscr{R}_{n-1}^{(n-2)} \otimes \mathscr{L}_n^{(n-2)} = \mathscr{C}_n$$

for all n. Thus the map β is an automorphism. By construction it is clear that this automorphism is a QCA with radius 2. For any single site operator x we have that $x \in \mathscr{B}_n$ for some n, so

$$\|\beta(x) - \alpha(x)\| \leq 2\|u_n - I\| + \|\alpha - \alpha_{n-1}\|$$
$$\leq \mathcal{O}(\varepsilon).$$

We showed $\|(\beta - \alpha)|_{\mathscr{A}_n}\| = \mathcal{O}(\varepsilon)$ for all single sites n, and the desired result holds by Lemma 7.7. ∎

By Proposition 9.15 and coarse-graining, we obtain the main result of this section, which shows that any ALPU in one dimensions can be approximated by a sequence of QCAs.

Theorem 9.16 (QCA approximations)**.** *If α is a one-dimensional ALPU with $f(r)$-tails, then there exists a sequence of QCAs $\{\beta_j\}_{j=1}^\infty$ of radius $2j$, such that for any finite $X \subset \mathbb{Z}$,*

$$\|(\alpha - \beta_j)|_{\mathscr{A}_X}\| = \mathcal{O}\left(f(j) \min\left\{|X|, \left\lceil \tfrac{\operatorname{diam}(X)}{j} \right\rceil \right\}\right). \tag{9.26}$$

Moreover, there is a constant $C_f > 0$, depending only on $f(r)$, such that the following holds for all j and finite $X \subset \mathbb{Z}$:

$$\|(\alpha - \beta_j)|_{\mathscr{A}_X}\| \leq C_f f(j) \min\left\{|X|, \left\lceil \tfrac{\operatorname{diam}(X)}{j} \right\rceil \right\}. \tag{9.27}$$

In particular, the β_j converge strongly to α, meaning that $\lim_{j \to \infty} \|\alpha(x) - \beta_j(x)\| = 0$ for all $x \in \mathscr{A}_{\mathbb{Z}}$.

Proof. By blocking j sites we obtain an ε_j-nearest neighbour QCA on the coarse-grained lattice where $\varepsilon_j = f(j)$. For $j > j_0$ sufficiently large, we can apply Proposition 9.15 to obtain a QCA β_j of radius 2 on the coarse-grained lattice satisfying

$$\|(\alpha - \beta)|_{\mathscr{A}_X}\| = \mathcal{O}(f(j)m)$$

for all regions X composed of m coarse-grained sites. If we now consider β_j as a QCA of radius $2j$ on the original lattice (before coarse-graining), we arrive at (9.26). To obtain (9.27), for smaller $j \leq j_0$, we may choose some arbitrary QCA β_j and we may use that $\|\alpha - \beta_j\| \leq 2$ at the expense of incurring a tails-dependent constant $C_f > 0$.

We now show that the sequence of QCAs β_j converges strongly to α. For $x \in \mathscr{A}_\mathbb{Z}$ arbitrary, let x_n be a sequence of strictly local operators, where x_n is supported on n contiguous sites, such that $\lim_{n\to\infty} x_n = x$ converges in norm. Then,

$$\limsup_{j\to\infty} \|\alpha(x) - \beta_j(x)\| \leq \limsup_{j\to\infty} \Big(\|\alpha(x) - \alpha(x_n)\| + \|\alpha(x_n) - \beta_j(x_n)\| + \|\beta_j(x_n) - \beta_j(x)\| \Big)$$

$$\leq 2\|x - x_n\| + \limsup_{j\to\infty} \|\alpha(x_n) - \beta_j(x_n)\| = 2\|x - x_n\|.$$

The second inequality holds since α and the β_j are $*$-homomorphisms; the final equality follows by (9.27). Since the above holds for all n, we conclude that

$$\lim_{j\to\infty} \|\alpha(x) - \beta_j(x)\| = 0.$$

■

9.2.2 Definition of the index for ALPUs

We now use the QCA approximations developed in the preceding to define an index for general ALPUs. In addition, we give two alternative ways of computing the index for ALPUs with appropriately decaying tails, and we prove that the index is stable also for ALPUs.

Definition 9.17 (Index for ALPUs). Let α be a one-dimensional ALPU with $f(r)$-tails and let β_j be a sequence of QCAs of radius at most $2j$ such that for any finite subset $X \subset \mathbb{Z}$,

$$\|(\alpha - \beta_j)|_{\mathscr{A}_X}\| \leq C_f f(j) \left\lceil \frac{\mathrm{diam}(X)}{j} \right\rceil, \tag{9.28}$$

where $C_f > 0$ is the constant from Theorem 9.16. We define the *index* of α by

$$\mathrm{ind}(\alpha) := \lim_{j\to\infty} \mathrm{ind}(\beta_j). \tag{9.29}$$

Note that by Theorem 9.16 such a sequence β_j always exists. The following theorem shows that the index is a well-defined, finite quantity.

Theorem 9.18 (Index for ALPUs). *Let α be a one-dimensional ALPU with $f(r)$-tails and let β_j be a sequence of QCAs as in Definition 9.17. Then the following hold:*

(i) There exists j_0, depending only on $f(r)$, such that $\mathrm{ind}(\beta_j)$ is constant for $j \geq j_0$. Accordingly, the limit (9.29) exists and is in $\mathbb{Z}[\{\log(p_i)\}]$, where the p_i are the finitely many prime factors of the local Hilbert space dimensions d_n, and $\mathbb{Z}[\cdot]$ denotes integer linear combinations. Moreover, this limit does not depend on the choice of sequence β_j. Thus, $\mathrm{ind}(\alpha)$ is well-defined by (9.29).

(ii) *There is a constant r_1, depending only on $f(r)$, with the following property: Let α' be another one-dimensional ALPU with $f(r)$-tails. Then, for any interval X with diameter $|X| \geq r_1$,*

$$\| (\alpha - \alpha')|_{\mathscr{A}_X} \| \leq \frac{1}{384} \implies \mathrm{ind}(\alpha) = \mathrm{ind}(\alpha').$$

In particular, the index is completely determined by $\alpha|_{\mathscr{A}_X}$ for any such X.

(iii) *If $f(r) = o(\frac{1}{r})$ then there exist a constant r_2, depending only on $f(r)$ and the local Hilbert space dimensions d_n, such that the index may also be computed locally as in (8.12),*

$$\mathrm{ind}(\alpha) = \mathrm{round}_{\mathbb{Z}[\{\log(p_i)\}]} \frac{1}{2} \big(I(L_1' : R_1)_\phi - I(L_1 : R_1')_\phi \big),$$

where ϕ denotes the Choi state, the intervals L_1, R_1, L_1', R_1' must be of size at least r_2, and the notation means that we round to the nearest value in $\mathbb{Z}[\{\log(p_i)\}]$.

(iv) *If $f(r) = \mathcal{O}(\frac{1}{r^{1+\delta}})$ for some $\delta > 0$, then the index can also be computed as in (9.8), by*

$$\mathrm{ind}(\alpha) = \frac{1}{2} \big(I(L' : R)_\phi - I(L : R')_\phi \big),$$

where both $I(L' : R)_\phi$ and $I(L : R')_\phi$ are finite.

In both calculations (iii) and (iv) of the index, the cut defining the regions L, R may be chosen anywhere on the chain.

Proof. Throughout this proof, the implicit constants in the $\mathcal{O}$ notation are allowed to depend on the tails $f(r)$.

(i) and (ii): To see that $\mathrm{ind}(\beta_j)$ stabilizes at large j and hence the limit (9.29) exists, consider β_j and β_{j+1}. After coarse-graining by blocking $2(j+1)$ sites, both β_j and β_{j+1} are nearest neighbour. Moreover, on any subset X_j that consists of two neighbouring coarse-grained sites,

$$\| (\beta_j - \beta_{j+1})|_{\mathscr{A}_{X_j}} \| \leq \| (\beta_j - \alpha)|_{\mathscr{A}_{X_j}} \| + \| (\alpha - \beta_{j+1})|_{\mathscr{A}_{X_j}} \| = \mathcal{O}(f(j)) \qquad (9.30)$$

by Eq. (9.28). Since $f(r) = o(1)$ this implies that $\| (\beta_j - \beta_{j+1})|_{\mathscr{A}_{X_j}} \|$ approaches zero as $j \to \infty$. By Proposition 8.11 this implies that $\mathrm{ind}(\beta_j) = \mathrm{ind}(\beta_{j+1})$ for sufficiently large $j \geq j_0$, where the constant j_0 can be taken as the minimum j such that the right-hand side of Eq. (9.30) remains below $\frac{1}{192}$. Thus we conclude that the limit (9.29) exists and equals $\mathrm{ind}(\beta_j)$ for $j \geq j_0$. Moreover, $\mathrm{ind}(\alpha) \in \mathbb{Z}[\{\log(p_i)\}]$, since the same is true for the index of the QCAs β_j.

To conclude the proof of (i), we still need to argue that the index is well-defined. We will demonstrate this together with (ii). Consider an ALPU α' that also has $f(r)$-tails, and let β_j' be a corresponding sequence of QCAs as in Definition 9.17. Note that $\mathrm{ind}(\beta_j)$ and $\mathrm{ind}(\beta_j')$ stabilize for $j \geq j_0$, with the same constant j_0. We claim that $\mathrm{ind}(\beta_j) = \mathrm{ind}(\beta_j')$ for some (and hence for all) $j \geq j_0$. To see this, we consider β_j and β_j' as nearest-neighbour QCAs on a coarse-grained lattice obtained by blocking $2j$ sites. Then by Proposition 8.11, it is sufficient to show $\| (\beta_j - \beta_j')|_{\mathscr{A}_Y} \| \leq \frac{1}{192}$ for a region Y consisting of

two neighbouring coarse-grained sites. Note Y then consists of $4j$ sites on the original lattice. Now,

$$\|(\beta_j - \beta'_j)|_{\mathscr{A}_Y}\| \leq \|(\beta_j - \alpha)|_{\mathscr{A}_Y}\| + \|(\alpha - \alpha')|_{\mathscr{A}_Y}\| + \|(\alpha' - \beta'_j)|_{\mathscr{A}_Y}\|$$
$$\leq \mathcal{O}(f(j)) + \|(\alpha - \alpha')|_{\mathscr{A}_Y}\|.$$

Since $f(r) = o(1)$, we can find $j_1 \geq j_0$ large enough such that the $\mathcal{O}(f(j))$ term is smaller than $\frac{1}{384}$. Take $r_1 := 8j_1$ to ensure that any interval X with r_1 sites contains two neighbouring sites of the coarse-grained lattice, so that $\|(\alpha - \alpha')|_{\mathscr{A}_Y}\| \leq \frac{1}{384}$ by assumption. Then, $\|(\beta_{j_1} - \beta'_{j_1})|_{\mathscr{A}_Y}\| \leq \frac{1}{192}$, and Proposition 8.11 implies that $\mathrm{ind}(\beta_j) = \mathrm{ind}(\beta'_j)$ for $j = j_1$ and hence for all $j \geq j_0$. This implies that the index is well-defined (take $\alpha = \alpha'$), concluding the proof of (i), and it also establishes (ii).

(iii) Let $L_j = \{-2j+1, \ldots, 0\}$ and $R_j = \{1, \ldots, 2j\}$. Since β_j is a QCA of radius $2j$, by Proposition 8.8 we can compute

$$\mathrm{ind}(\beta_j) = \frac{1}{2}\left(I(L'_j : R_j)_{\phi_j} - I(L_j : R'_j)_{\phi_j}\right) \tag{9.31}$$

where $\phi_j = (\beta_j^\dagger \otimes I)(\omega)$, with ω a maximally entangled state on $\mathscr{A}_{\mathbb{Z}} \otimes \mathscr{A}_{\mathbb{Z}}$. We let

$$\widetilde{\mathrm{ind}}_j(\alpha) := \widetilde{\mathrm{ind}}_{L_j, R_j}(\alpha) = \frac{1}{2}\left(I(L'_j : R_j)_\phi - I(L_j : R'_j)_\phi\right) \tag{9.32}$$

as in (8.14), where $\phi = (\alpha^\dagger \otimes I)(\omega)$. By Eq. (9.28), $\|(\alpha - \beta_j)|_{A_{X_j}}\| = \mathcal{O}(f(j))$, where we let $X_j = L_j \cup R_j$. Thus Lemma 8.10 shows that

$$\left|\mathrm{ind}(\beta_j) - \widetilde{\mathrm{ind}}_j(\alpha)\right| = \mathcal{O}\left(jf(j)\log(d) + f(j)\log\frac{1}{f(j)}\right) \tag{9.33}$$

where $d = \max_n d_n$ is the maximum of the local Hilbert space dimensions associated to $\mathscr{A}_{\mathbb{Z}}$. Assuming that $f(j) = o(\frac{1}{j})$ the above approaches zero as $j \to \infty$. Because the sequence $\mathrm{ind}(\beta_j)$ stabilizes to $\mathrm{ind}(\alpha)$ by definition in (9.29), this implies that

$$\lim_{j \to \infty} \widetilde{\mathrm{ind}}_j(\alpha) = \mathrm{ind}(\alpha). \tag{9.34}$$

Since $\mathrm{ind}(\alpha)$ takes values in the nowhere dense set $\mathbb{Z}[\{\log(p_i)\}]$, rounding $\widetilde{\mathrm{ind}}_j(\alpha)$ must yield $\mathrm{ind}(\alpha)$ for sufficiently large j, proving (iii).

(iv) Even though the quantities in Eqs. (9.31) and (9.32) converge with j, we have not yet shown that the individual mutual information terms converge. We will show this next, assuming that $f(r) = \mathcal{O}(\frac{1}{r^{1+\delta}})$ for some $\delta > 0$. We consider the subsequence $\{\beta_{2^k}\}$. Then, by Eq. (9.28)

$$\|(\beta_{2^k} - \alpha)|_{\mathscr{A}_{X_{2^{k+1}}}}\| = \mathcal{O}(f(2^k)), \tag{9.35}$$

and thus

$$\|(\beta_{2^k} - \beta_{2^{k+1}})|_{\mathscr{A}_{X_{2^{k+1}}}}\| = \mathcal{O}(f(2^k)).$$

Hence by Lemma 8.10, noting that $I(L'_{2^{k+1}} : R_{2^{k+1}})_{\phi_{2^k}} = I(L'_{2^k} : R_{2^k})_{\phi_{2^k}}$ since β_{2^k} has radius 2^{k+1}, as similarly observed in the proof of Proposition 8.8, this implies

$$|I(L'_{2^k} : R_{2^k})_{\phi_{2^k}} - I(L'_{2^{k+1}} : R_{2^{k+1}})_{\phi_{2^{k+1}}}| = \mathcal{O}\left(2^k f(2^k)\log(d) + f(2^k)\log\frac{1}{f(2^k)}\right)$$
$$= \mathcal{O}(2^{-\delta k}).$$

Thus $I(L'_{2^k} : R_{2^k})_{\phi_{2^k}}$ is a Cauchy sequence and hence converges. Moreover, using Lemma 8.10, Eq. (9.35) also implies that

$$|I(L'_{2^k} : R_{2^k})_\phi - I(L'_{2^k} : R_{2^k})_{\phi_{2^k}}| = \mathcal{O}\left(2^k f(2^k) \log(d) + f(2^k) \log \frac{1}{f(2^k)}\right)$$
$$= \mathcal{O}(2^{-\delta k}).$$

Thus $I(L'_{2^k} : R_{2^k})_\phi$ also converges, with the same limit as $I(L'_{2^k} : R_{2^k})_{\phi_{2^k}}$. Then using Proposition 8.7, this implies that

$$I(L' : R)_\phi = \lim_{k\to\infty} I(L'_{2^k} : R_{2^k})_\phi = \lim_{k\to\infty} I(L'_{2^k} : R_{2^k})_{\phi_{2^k}}$$

is finite. A similar argument shows that $I(L : R')_\phi$ is finite and can be computed as

$$I(L' : R)_\phi = \lim_{k\to\infty} I(L_{2^k} : R'_{2^k})_\phi = \lim_{k\to\infty} I(L_{2^k} : R'_{2^k})_{\phi_{2^k}}.$$

It follows that

$$\mathrm{ind}(\alpha) = \frac{1}{2}\left(I(L' : R)_\phi - I(L : R')_\phi\right),$$

as a consequence either of Eq. (9.34) or of Eq. (9.29).

In parts (iii) and (iv) we took the cut between L and R to be at $n = 0$, but the index may be calculated using regions translated anywhere along the chain, which follows from the same fact for the QCAs β_j. ∎

The proof of Theorem 9.18 also shows that in part (iv), the two mutual information quantities can be computed as limits of corresponding mutual information quantities for finite intervals.

9.2.3 Properties of the index for ALPUs

In this section we will show that the index for ALPUs defined in Theorem 9.18 inherits essentially all properties of the GNVW index for QCAs stated in Theorem 8.5.

We first use Theorem 9.16 to construct a path between any ALPU α with $\mathrm{ind}(\alpha) = 0$ and the identity automorphism I, using a one-parameter family of ALPUs $\beta[t]$ for time $t \in [0, 1]$, with $\beta[0] = I$ and $\beta[1] = \alpha$. The path will be *strongly continuous*, in the sense that for all $x \in \mathscr{A}_\mathbb{Z}$, $t_0 \in [0, 1]$,

$$\lim_{t\to t_0} \|\alpha[t](x) - \alpha[t_0](x)\| = 0. \tag{9.36}$$

Theorem 9.19 (Continuous deformations)**.** *If α is a one-dimensional ALPU with $f(r)$-tails with $\mathrm{ind}(\alpha) = 0$, then there exists a strongly continuous path $\alpha[t]$ starting at $\alpha[0] = I$ and $\alpha[1] = \alpha$ such that $\alpha[t]$ has $g(r)$-tails for all t, for some $g(r) = \mathcal{O}(f(Cr))$ and some universal constant $C > 0$. Moreover, this path may be given by a time evolution using a time-dependent Hamiltonian $H(t)$ evolving for unit time. For every $t < 1$ there exists l such that the Hamiltonian $H(t)$ has only terms H_X on (nonoverlapping) sets X of diameter at most $16l$, and it holds that $\|H_X(t)\| = \mathcal{O}(f(l)\log(l))$.*
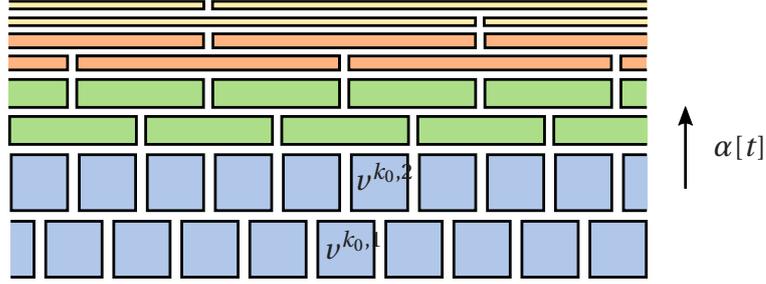
Figure 9.3: Illustration of the construction of $\alpha[t]$ in Theorem 9.19. The evolution consists of successive evolutions by different time-independent Hamiltonians, depicted as successive layers, with interaction terms increasing in diameter but decreasing in strength.

The above $H(t)$ is constructed as piecewise-constant for $t \in [0,1)$. The idea of the proof is that we continuously interpolate between consecutive QCAs β_{2j} as constructed in Theorem 9.16. For large j we need to use a Hamiltonian with a correspondingly large support to interpolate between β_{2j} and β_{2j+1}, but on the other hand β_{2j} and β_{2j+1} are locally close, so the interaction strength is small. As we increase j, we "speed up" the interpolation, so we get to α in unit time. In particular, the Hamiltonian is piecewise constant on time intervals that decrease in size as t goes to 1, and the support of the Hamiltonian increases as t goes to 1. This procedure is illustrated in Fig. 9.3 and leads to the given bound on the terms $H_X(t)$ of the Hamiltonian. For $f(r) = \mathcal{O}((\log r)^{-1})$ the norms $\|H_X(t)\|$ are uniformly bounded as $t \to 1$; more generally for $f(r) = o(1)$ the terms may diverge in norm as $t \to 1$, but nonetheless the path $H(t)$ is strongly continuous on the interval $[0,1]$. Of course, the path $\alpha(t)$ and associated Hamiltonian $H(t)$ are not unique; we just provide one particular construction.

Proof. We apply Theorem 9.16 to obtain a sequence of QCA approximations β_j with radius $2j$ with error $\|(\alpha - \beta_j)|_{\mathscr{A}_X}\| = \mathcal{O}(f(j)\lceil\frac{\text{diam}(X)}{j}\rceil)$ as $j \to \infty$. Therefore

$$\|(\beta_{2j} - \beta_j)|_{\mathscr{A}_X}\| \le \|(\alpha - \beta_j)|_{\mathscr{A}_X}\| + \|(\alpha - \beta_{2j})|_{\mathscr{A}_X}\| = \mathcal{O}(f(j)\lceil\tfrac{\text{diam}(X)}{j}\rceil),$$

having used that f is non-increasing, and hence

$$\|(\beta_{2j}\beta_j^{-1} - I)|_{\mathscr{A}_X}\| \le \|(\beta_{2j} - \beta_j)|_{\mathscr{A}_{B(X,2j)}}\| = \mathcal{O}(f(j)\lceil\tfrac{\text{diam}(X)}{j}\rceil)$$

We can therefore define QCAs

$$\gamma_k = \beta_{2^{k+1}}\beta_{2^k}^{-1}$$

which have at most radius $R_k = 2^{k+3}$ and satisfy $\|(\gamma_k - I)|_{\mathscr{A}_X}\| = \mathcal{O}(f(2^k))$ for on $X \subset \mathbb{Z}$ with $\text{diam}(X) \le 2R_k = 2^{k+4}$. For sufficiently large $k \ge k_0$, $\text{ind}(\beta_{2^k}) = \text{ind}(\alpha) = 0$, and hence $\text{ind}(\gamma_k) = 0$.

By Theorem 8.2 and Theorem 8.5, any index-0 QCA of radius R can be decomposed as a two-layer circuit with unitaries on blocks of diameter $2R$. If the QCA is ε-near the identity when restricted to intervals of size $2R$, the individual unitaries in the circuit are $\mathcal{O}(\varepsilon)$-near the identity by Proposition 8.13. Therefore γ_k may be implemented by a two-layer unitary circuit for $k \ge k_0$. We proceed to describe this circuit as a Hamiltonian

evolution, with a different time-independent Hamiltonian generating each layer, in the straightforward way. To be precise, from Proposition 8.13 we obtain that

$$\gamma_k(x) = (v^{(k,2)})^*(v^{(k,1)})^* x v^{(k,1)} v^{(k,2)}$$

where for each layer $a \in \{1, 2\}$

$$v^{(k,a)} = \prod_n v_n^{(k,a)}$$

and where the $\{v_n^{(k,a)}\}_n$ are unitary gates acting on disjoint regions with a diameter of $2R_k = 2^{k+4}$. Moreover, each gate satisfies $\|v_n^{(k,a)} - I\| = \mathcal{O}(f(2^k))$. Each gate is generated by a Hamiltonian $H_n^{(k,a)} = -i \log(v_n^{(k,a)})$, defined using the principal logarithm, with bounded norm $\|H_n^{(k,a)}\| = \mathcal{O}(f(2^k))$. Let $H^{(k,a)} = \sum_n H_n^{(k,a)}$ denote the total Hamiltonian generating the a-th layer. Then we can define a Hamiltonian evolution $\gamma_k[t]$ for $t \in [0, 1]$ with $\gamma_k[0] = I$, $\gamma_k[1] = \gamma_k$:

$$\gamma_k[t](x) = e^{2iH^{(k,1)}t}(x) e^{-2iH^{(k,1)}t}$$

for $t \in [0, \frac{1}{2}]$ and

$$\gamma_k[t](x) = e^{2iH^{(k,2)}(t-\frac{1}{2})} e^{iH^{(k,1)}}(x) e^{-iH^{(k,1)}} e^{-2iH^{(k,2)}(t-\frac{1}{2})}$$

for $t \in (\frac{1}{2}, 1]$. Note that the gates implementing $\gamma_k[t]$ are all of the form $(v_n^{(k,a)})^s$ for some $s \in [0, 1]$. From this it is clear that $\gamma_k[t]$ defines a strongly continuous path and the evolution is gentle in the sense that $\gamma_k[t]$ never strays far from I:

$$\|(\gamma_k[t] - I)|_{\mathscr{A}_X}\| = \mathcal{O}(f(2^k)).$$

for $\mathrm{diam}(X) \le 2R_k$. By construction, $\gamma_k[t]$ is a QCA with radius at most $3R_k$ for every time $t \in [0, 1]$.

We let $\alpha_{k+1}[t] := \gamma_k[t]\beta_{2^k}$, which is a strongly continuous path with $\alpha_{k+1}[0] = \beta_{2^k}$ and $\alpha_{k+1}[1] = \beta_{2^{k+1}}$. For all $t \in [0, 1]$

$$\|(\alpha_{k+1}[t] - \alpha)|_{\mathscr{A}_X}\| \le \|(\gamma_k - I)|_{\mathscr{A}_{B(X, 2^{k+1})}}\| + \|(\alpha - \beta_{2^k})|_{\mathscr{A}_X}\| = \mathcal{O}(f(2^k)) \tag{9.37}$$

for $\mathrm{diam}(X) \le R_k$. Moreover $\alpha_{k+1}[t]$ is a QCA with radius $3R_k + 2^{k+1} \le 4R_k$.

We defined $\alpha_k[t]$ only for $k > k_0$. Let $\alpha_{k_0}[t]$ be the Hamiltonian evolution implementing the index-0 QCA $\beta_{2^{k_0}}$ for $t \in [0, 1]$, in the same way we defined $\gamma_k[t]$. Let

$$T = \sum_{k=0}^{\infty} \frac{1}{1 + k^2}, \qquad t_k = \sum_{l=0}^{k-1} \frac{1}{T(1 + l^2)}.$$

We define $\alpha[t]$ by gluing together the $\alpha_k[t]$, "speeding up" α_{k_0+k} by a factor $T(k^2 + 1)$ in order to make this a unit time evolution:

$$\alpha[t] = \alpha_{k_0+k}\left[\frac{t - t_k}{T(k^2 + 1)}\right] \text{ if } t \in (t_k, t_{k+1})$$

for $t \in [0, 1)$ and $\alpha[1] = \alpha$. The construction of the path $\alpha[t]$ is illustrated in Fig. 9.3. Going through γ_{k_0+k} faster by a factor $T(k^2+1)$ is equivalent to rescaling the Hamiltonian

by $T(k^2 + 1)$, and is still strongly continuous. Hence $\alpha[t]$ is strongly continuous for at any $t \in [0, 1)$. The strong continuity at $t = 1$ follows from the fact that the sequence β_{2^k} converges strongly to α. Indeed, let $x \in \mathscr{A}_X$ for finite X. Then consider k such that $\mathrm{diam}(X) \le R_{k_0+k}$, then we see that for $l \ge k_0 + k \|\alpha_{l+1}[s](x) - \alpha(x)\| = \mathcal{O}(f(2^l))$ for $s \in [0, 1]$. Hence, $\|\alpha[t](x) - \alpha(x)\|$ goes to zero as $t \to 1$. As in the proof of Theorem 9.16 we see that the same holds for general $x \in \mathscr{A}_{\mathbb{Z}}$. Moreover, we see that at each point in time the Hamiltonian will have terms H_X with support of diameter $16l = 2^{k+4}$ for some k with $\|H_X\| = \mathcal{O}(f(2^k)k^2) = \mathcal{O}(f(l)\log(l))$.

Finally, we need to show that $\alpha[t]$ has uniform tail bounds for $t \in (0, 1)$. (We already have tail bounds at the initial and final time.) Let $X \subseteq \mathbb{Z}$ be an arbitrary (finite or infinite) interval. Take some $r > 4R_{k_0} = r_0$. There will be some k such that $4R_k \le r < 4R_{k+1}$, and there will be some l and $s \in [0, 1]$ such that $\alpha[t] = \alpha_{l+1}[s]$. If $k \ge l$, by construction we have an inclusion $\alpha[t](\mathscr{A}_X) \subseteq \mathscr{A}_{B(X,4R_l)} \subseteq \mathscr{A}_{B(X,r)}$. On the other hand, suppose that $k < l$. Write $X = X_1 \cup X_2$ where X_1 is the (possibly empty set) of elements with distance from the boundary larger than $4R_l$. Then $\alpha_{l+1}[s](\mathscr{A}_{X_1}) \subset \mathscr{A}_X$. Moreover, since the set X_2 consists of at most two intervals of size $4R_l$ we have, using Lemma 7.7 and (9.37) that $\|(\alpha - \alpha_{l+1}[s])|_{\mathscr{A}_{X_2}}\| = \mathcal{O}(f(2^l))$. Since α has $f(r)$-tails,

$$\alpha(\mathscr{A}_{X_2}) \overset{\mathcal{O}(f(r))}{\subseteq} \mathscr{A}_{B(X,r)},$$

and since $r < 4R_l = 2^{l+5}$ we see that

$$\alpha_{l+1}[s](\mathscr{A}_{X_2}) \overset{\mathcal{O}(f(r)+f(2^l))}{\subseteq} \mathscr{A}_{B(X,r)}$$

$$\alpha_{l+1}[s](\mathscr{A}_{X_2}) \overset{\mathcal{O}(f(\frac{r}{32}))}{\subseteq} \mathscr{A}_{B(X,r)}.$$

Lemma 7.12 allows us to conclude that

$$\alpha[t](\mathscr{A}_X) = \alpha_{l+1}[s](\mathscr{A}_X) \overset{\mathcal{O}(f(\frac{r}{32}))}{\subseteq} \mathscr{A}_{B(X,r)}.$$

$\blacksquare$

Remark 9.20. If α has $\mathcal{O}(\frac{1}{r^{1+a}})$-tails for $a > 0$, then for $0 < b < a$ and with reproducing function $F(r) = \frac{1}{(1+r)^{1+b}}$ the Hamiltonian constructed in Theorem 9.19 satisfies the hypotheses in Theorem 9.1 (Lieb-Robinson). However, notice that the locality estimates you get from applying the Lieb-Robinson bounds to these bounds are weaker than the original locality bounds on $\alpha[t]$.

Remark 9.21. The Hamiltonian evolution constructed in Theorem 9.19 cannot always be approximated by a 2-local quantum circuit of constant depth. Likewise, even QCAs of radius r may have circuit complexity exponential in r when using 2-local gates.

Remark 9.22. Finally, we observe that in Theorem 9.19 if we have exponential tails, which decay as $f(r) = \mathcal{O}(e^{-ar})$, one obtains that $\alpha[t]$ has $\mathcal{O}(e^{-aCr})$-tails. This is not entirely optimal, and for exponential tails one can slightly change the proof, by considering the sequence β_k rather than β_{2^k} and correspondingly $\gamma_k = \beta_{k+1}\beta_k^{-1}$ instead of $\gamma_k = \beta_{2^{k+1}}\beta_{2^k}^{-1}$. The same arguments as in the proof of Theorem 9.19 then lead to a path $\alpha[t]$ where $\alpha[t]$ has $\mathcal{O}(f(r+C)) = \mathcal{O}(e^{-ar})$-tails, which is implemented by a Hamiltonian $H(t)$. In this case the Hamiltonian is such that for every t, there exists k such that $H(t)$ has only terms H_X on (nonoverlapping) sets X of diameter at most k, with $\|H_X(t)\| = \mathcal{O}(k^2 e^{-ak})$.

Next we discuss blending. We need a slightly weaker notion than for QCAs.

Definition 9.23. Two ALPUs α_1 and α_2 in one dimension can be *blended* (at the origin) if there exists an ALPU β on some $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ such that

$$\lim_{r \to \infty} \|(\beta - \alpha_1)|_{\mathscr{A}_{\leq -r}}\| = 0,$$
$$\lim_{r \to \infty} \|(\beta - \alpha_2)|_{\mathscr{A}_{\geq r}}\| = 0.$$

Proposition 9.24. *Two ALPUs α_1, α_2 can be blended if and only if* $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$.

When $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$ and both ALPUs have $f(r)$-tails, the approximation requirement of the blending as defined in Definition 9.23 can be refined as in (9.40) as discussed in the proof. The blending proceeds similarly to the construction in Proposition 9.15.

Proof. First we assume α_1 and α_2 can be blended and show $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$. Consider a blended ALPU β as in Definition 9.23. By Theorem 9.18(ii), one may compute $\mathrm{ind}(\beta)$ locally on either half of the blended chain. Both calculations must yield the same index, which does not depend on where it is locally calculated. By (ii) of Theorem 9.18, the index computed locally at the sufficiently far left must be $\mathrm{ind}(\alpha_1)$, and the index computed at the far right must be $\mathrm{ind}(\alpha_2)$. Thus, $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$.

Next we show that if $\mathrm{ind}(\alpha_1) = \mathrm{ind}(\alpha_2)$, then α_1 and α_2 can be blended. We assume both ALPUs are defined on the same $\mathscr{A}_{\mathbb{Z}}$ (i.e., the local dimensions are the same) and address the general case afterward. Both ALPUs extend to automorphisms of $\mathscr{A}_{\mathbb{Z}}^{\mathrm{vN}}$ as in Remark 9.7. Coarse-grain the lattice until both α_1 and α_2 are ε-nearest neighbour ALPUs, with ε smaller than a universal constant determined by the remainder of the proof. Then we can apply Proposition 9.14 (if $\varepsilon \leq \varepsilon_1$) separately to α_1 and α_2 at site $n = 0$. Denote the ALPUs resulting from Proposition 9.14 as $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$, respectively. Then by construction $\|\tilde{\alpha}_i - \alpha_i\| \leq C_1 \varepsilon$ for $i = 1, 2$. Moreover by Theorem 9.18(ii), we can take ε small enough that $\mathrm{ind}(\alpha_i) = \mathrm{ind}(\tilde{\alpha}_i)$ for $i = 1, 2$, hence $\mathrm{ind}(\tilde{\alpha}_1) = \mathrm{ind}(\tilde{\alpha}_2)$.

As usual, we write $\mathscr{B}_n = \mathscr{A}_{\{2n,2n+1\}}$ and $\mathscr{C}_n = \mathscr{A}_{\{2n-1,2n\}}$. Then by their construction, $\tilde{\alpha}_i$ for $i = 1, 2$ both satisfy the locality properties $\tilde{\alpha}_i(\mathscr{B}_n) \subseteq \mathscr{C}_n \otimes \mathscr{C}_{n+1}$ for $n = 0, 1, 2$, as well as $\tilde{\alpha}_i^{-1}(\mathscr{C}_n) \subseteq \mathscr{B}_{n-1} \otimes \mathscr{B}_n$ for $n = 1, 2$. Then by Theorem 8.2 and subsequent Remark 8.3, for each $i = 0, 1$ and $n = 1, 2$ we can define

$$\mathscr{L}_n^{(i)} = \tilde{\alpha}_i(\mathscr{B}_n) \cap \mathscr{C}_n,$$
$$\mathscr{R}_{n-1}^{(i)} = \tilde{\alpha}_i(\mathscr{B}_{n-1}) \cap \mathscr{C}_n$$

such that

$$\mathscr{C}_n = \mathscr{L}_n^{(i)} \otimes \mathscr{R}_{n-1}^{(i)} \tag{9.38}$$

and, for $n = 1$,

$$\mathscr{B}_n = \tilde{\alpha}_i^{-1}(\mathscr{L}_n^{(i)}) \otimes \tilde{\alpha}_i^{-1}(\mathscr{R}_n^{(i)}). \tag{9.39}$$

Following the structure theory of QCAs in Theorem 8.2, one can for each $i = 1, 2$ find

a QCA β_i of radius 2 such that $\beta_i|_{\mathscr{A}_{\{0,\ldots,5\}}} = \tilde{\alpha}_i|_{\mathscr{A}_{\{0,\ldots,5\}}}$.[2] By the latter condition, Theorem 9.18, (ii) implies (if we have sufficiently coarse-grained in our initial step) that $\mathrm{ind}(\tilde{\alpha}_i) = \mathrm{ind}(\beta_i)$. Recalling that $\mathrm{ind}(\tilde{\alpha}_1) = \mathrm{ind}(\tilde{\alpha}_2)$, we then have $\mathrm{ind}(\beta_1) = \mathrm{ind}(\beta_2)$, so from Eq. (8.7) we conclude that $\mathscr{R}_0^{(1)}$ and $\mathscr{R}_0^{(2)}$ have the same dimension and hence are isomorphic finite-dimensional subalgebras of $\mathscr{C}_1$. Hence there exists a unitary $u \in \mathscr{C}_1$ such that $u\mathscr{R}_0^{(1)}u^* = \mathscr{R}_0^{(2)}$.

Now we are in position to define the blended ALPU β. For $x \in \mathscr{A}_{\leq 1}$ we define β as $\beta(x) = u\tilde{\alpha}_1(x)u^*$. We let $\beta|_{\mathscr{A}_{\geq 2}} = \tilde{\alpha}_2|_{\mathscr{A}_{\geq 2}}$. Then

$$\beta(\mathscr{A}_{\leq 1}) = (\mathscr{A}_{\leq 0} \cup \mathscr{R}_0^{(2)})''$$

and

$$\beta(\mathscr{A}_{\geq 2}) = (\mathscr{L}_1^{(2)} \cup \mathscr{A}_{\geq 3})''$$

commute by construction, so β is a well-defined injective unital $*$-homomorphism. Moreover $\mathscr{C}_1 = \mathscr{L}_1^{(2)} \otimes \mathscr{R}_0^{(2)}$ from (9.38), so β is surjective, hence a well-defined ALPU. By construction of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ using Proposition 9.14, for $r \to \infty$,

$$\begin{aligned}
\|(\beta - \alpha_1)|_{\mathscr{A}_{\leq -r-1}}\| &= \mathcal{O}(f(r-1)), \\
\|(\beta - \alpha_2)|_{\mathscr{A}_{\geq r+6}}\| &= \mathcal{O}(f(r-7)).
\end{aligned} \tag{9.40}$$

Above we assumed both α_1 and α_2 were defined on the same $\mathscr{A}_{\mathbb{Z}}$ (i.e., that both chains use algebras $\mathscr{A}_n$ of the same dimensions). If α_1 and α_2 have different local dimensions, then in the region where we blend them above, we can first pad them with extra tensor factors so that they have identical local dimensions within that region. ∎

The following theorem extends all properties in Theorem 8.5 for QCAs to ALPUs, replacing the role of circuits by Hamiltonian evolutions, and allowing strongly continuous paths through the space of ALPUs with uniform tail bounds.

Theorem 9.25 (Properties of index for ALPUs)**.** *Suppose α and β are ALPUs in one dimension. Then:*

(i) $\mathrm{ind}(\alpha \otimes \beta) = \mathrm{ind}(\alpha) + \mathrm{ind}(\beta)$.

(ii) *If α and β are defined on the same algebra,* $\mathrm{ind}(\alpha\beta) = \mathrm{ind}(\alpha) + \mathrm{ind}(\beta)$.

(iii) *The following are equivalent:*

  (a) $\mathrm{ind}(\alpha) = \mathrm{ind}(\beta)$.

  (b) *α and β may be blended.*

---

[2] Indeed, β_i can be constructed as follows. First we define $\beta_i|_{\mathscr{A}_{\{0,\ldots,5\}}} = \tilde{\alpha}_i|_{\mathscr{A}_{\{0,\ldots,5\}}}$ and then we define the action of β_i on the remainder of $\mathscr{A}_{\mathbb{Z}}$ as follows. We focus on defining β_i for $\mathscr{A}_{\leq -1}$; the definition for $\mathscr{A}_{\geq 6}$ is directly analogous. An easy argument shows that $\mathscr{L}_0^{(i)} := \tilde{\alpha}_i(\mathscr{B}_0) \cap \mathscr{C}_0$ is a factor and moreover Eq. (9.39) also holds for $n = 0$. Then Eq. (9.38) will also hold for $n = 0$ if we define $\mathscr{R}_{-1}^{(i)}$ alternatively as the complementary factor to $\mathscr{L}_0^{(i)} \subset \mathscr{C}_0$. For each $n \leq -1$, choose an arbitrary factorization $\mathscr{C}_n = \mathscr{L}_n^{(i)} \otimes \mathscr{R}_{n-1}^{(i)}$ with $\mathscr{L}_n^{(i)} \cong \mathscr{L}_1^{(i)}$ and $\mathscr{R}_{n-1}^{(i)} \cong \mathscr{R}_{-1}^{(i)} \cong \mathscr{R}_0^{(i)} \cong \mathscr{R}_1^{(i)}$ (using that, by assumption, all local dimensions are the same). For each $n \leq -1$, choose an arbitrary factorization $\mathscr{B}_n = \tilde{\mathscr{L}}_n^{(i)} \otimes \tilde{\mathscr{R}}_n^{(i)}$ into factors isomorphic to those used for $\mathscr{B}_1$ in Eq. (9.39) for $n = 1$. Then we have $\tilde{\mathscr{L}}_n^{(i)} \cong \mathscr{L}_n^{(i)}$ and $\tilde{\mathscr{R}}_n^{(i)} \cong \mathscr{R}_n^{(i)}$, so we can define β_i to act as $\beta_i(\tilde{\mathscr{L}}_n^{(i)}) = \mathscr{L}_n^{(i)}$ and $\beta_i(\tilde{\mathscr{R}}_n^{(i)}) = \mathscr{R}_n^{(i)}$ for $n \leq -1$. This completes the definition of β_i for $\mathscr{A}_{\leq -1}$.

      (c) There exists an index-0 ALPU γ such that $\alpha = \beta\gamma$.

      (d) There exists $g(r) = o(1)$ and a strongly continuous path $\alpha[t]$ through the space of ALPUs with $g(r)$-tails such that $\alpha[0] = \alpha$ and $\alpha[1] = \beta$.

In (d), if α and β have $f(r)$-tails, we may take $g(r) = \mathcal{O}(f(Cr))$ for a universal constant C. If they have $\mathcal{O}(e^{-ar})$-tails, we may take $g(r) = \mathcal{O}(e^{-ar})$. In general, the path in (d) may be implemented by composing α (or β) with a Hamiltonian evolution with time-dependent Hamiltonian $H(t)$, with interactions bounded as in Theorem 9.19 and Remark 9.22.

In (c), if α and β do not have the same local dimensions, the statement only holds after separately tensoring α and β with appropriate identity automorphisms, such that α and β then have the same local dimensions. The analogous modification is needed for (d).

Proof. If α and β are ALPUs with approximating sequences α_n and β_n as in Theorems 9.16 and 9.18, then $\alpha_n \otimes \beta_n$ and $\alpha_{2n}\beta_n$ approximate $\alpha \otimes \beta$ and $\alpha\beta$ respectively. Then (i) and (ii) follow from the corresponding property for QCAs (Theorem 8.5). For (iii) the equivalence (a) $\Leftrightarrow$ (b) is stated by Proposition 9.24. The equivalence (a) $\Leftrightarrow$ (c) follows from $\mathrm{ind}(\beta\alpha^{-1}) = \mathrm{ind}(\beta) - \mathrm{ind}(\alpha)$, using property (ii). The implication (a) $\Rightarrow$ (d) follows from Theorem 9.19 applied to $\beta\alpha^{-1}$. The comment about exponential tails follows from the remark after Theorem 9.19. Next we show (d) $\Rightarrow$ (a), i.e. that the index must remain constant along a strongly continuous path. Because all ALPUs in the path are assumed to have $g(r)$-tails for some fixed $g(r)$, by Theorem 9.18(ii) there exists a finite interval X such that any two ALPUs γ and γ' with $\|(\gamma - \gamma')|_{\mathscr{A}_X}\|$ sufficiently small must have $\mathrm{ind}(\gamma) = \mathrm{ind}(\gamma')$. By the strong continuity (9.36) of the path, the index must then be constant along the path. ∎

In the terminology of [Has13], Theorem 9.25 shows that an (A)LPU is an LGU (locally generated unitary) if and only if it has index zero. We can also interpret Theorem 9.25 as a converse to the Lieb-Robinson bounds in one dimension. Again, recall that Lieb-Robinson bounds demonstrate that local Hamiltonian evolution exhibits an approximate causal cone, quantified by the bound. Conversely, we ask whether evolutions that satisfy Lieb-Robinson-type bounds (i.e. ALPUs) can be generated by some time-dependent Hamiltonian. We find the following converse, emphasized below.

Corollary 9.26 (Converse to Lieb-Robinson bounds)**.** *Suppose α is an ALPU in one dimension with $f(r)$-tails. If (and only if) $\mathrm{ind}(\alpha) = 0$, α can be implemented by a strongly continuous path $\alpha[t]$ generated by some time-dependent Hamiltonian $H(t)$, such that $\alpha[0] = I$, $\alpha[1] = \alpha$, and $\alpha[t]$ has $g(r)$-tails for all t, for some $g(r) = o(1)$. If α has $f(r)$-tails, we may take $g(r) = \mathcal{O}(f(Cr))$ for a universal constant C. If the ALPU α has $\mathcal{O}(e^{-ar})$-tails, we may take $g(r) = \mathcal{O}(e^{-ar})$. The Hamiltonian $H(t)$ can be taken to have interactions bounded as in Theorem 9.19 and Remark 9.22.*

    More generally, every ALPU in one dimension is a composition of a shift and a Hamiltonian evolution as above.

Proof. The equivalence follows immediately from (iii) in Theorem 9.25. The final statement follows by letting σ be a shift with $\mathrm{ind}(\sigma) = \mathrm{ind}(\alpha)$, then $\mathrm{ind}(\alpha\sigma^{-1}) = 0$ by (ii) in Theorem 9.25 so there exists a Hamiltonian evolution γ such that $\gamma = \alpha\sigma^{-1}$ and hence we have $\alpha = \gamma\sigma$. ∎

The use of time-dependent rather than time-independent Hamiltonians is necessary, as [ZFFM20] shows that there exist QCAs with index 0 that cannot be implemented by any time-independent local Hamiltonian.

9.2.4   Finite chains

We developed the above structure theory of ALPUs on the infinite one-dimensional lattice. The statements are easily be adapted to the case of a finite one-dimensional chain with non-periodic ("open") boundary conditions. The statements as well as the proofs essentially hold unchanged, but we make some clarifying remarks. In summary, the theorems only become nontrivial when the length $|\Gamma|$ of the chain is taken larger than some finite threshold, but this threshold depends only on the tails and local dimensions of the ALPU. Meanwhile, the index is always zero.

We work with the algebra $\mathscr{A}_\Gamma$, where Γ is now a finite interval $\Gamma \subset \mathbb{Z}$. By non-periodic boundary conditions, we mean that Γ is considered as an interval rather than a circle, i.e. Γ inherits the metric from $\mathbb{Z}$, and the sites at either end of the interval are not considered neighbours. We again consider ALPUs on $\mathscr{A}_\Gamma$ with $f(r)$ tails, where $f(r)$ is only meaningful for $r < |\Gamma|$. In our arguments, $\mathscr{A}_{\leq n}$ becomes the finite-dimensional algebra corresponding to all sites left of $n+1$, and so on.

With this modification, Lemma 9.11 holds as stated, and the proof is identical. Importantly, all unspecified constants appearing as $\mathscr{O}(\cdot)$ in e.g. (9.11) are independent of the chain length $|\Gamma|$.

We then arrive at Theorem 9.16 for finite one-dimensional lattices, describing QCA approximations to ALPUs. Given ALPU α with $f(r)$ tails, the theorem describes an increasing sequence of QCA approximations β_j of radius j. For finite Γ, we restrict attention to $j \leq |\Gamma|$, so that the notion of a QCA of radius j remains meaningful. Recall the QCA approximations β_j were only guaranteed to have the listed properties in Theorem 9.18 for $j > j_0$, with j_0 chosen such that $f(j_0)$ is smaller than some universal constant independent of $|\Gamma|$. Then we only need $|\Gamma| > j_0$ for Theorem 9.18 to yield nontrivial QCA approximations, and this threshold size is determined only by the tails $f(r)$. Finally, the assumption $f(r) = o(\frac{1}{r})$ used for the latter claims of Theorem 9.18 may expressed more explicitly as the assumption that $f(j_0)j_0$ is smaller than some constant depending only on the local dimensions d_n of $\mathscr{A}_\Gamma$. This assumption then increases the minimum length $|\Gamma|$ for the theorem to become nontrivial, but with the minimum depending only on the tails and local dimensions, rather the details of α.

While Theorem 9.18 holds as written for finite Γ, it also reduces to a special case: the index is always zero. Calculating the index as the entropy difference in (8.13), we see the entropies correspond to complementary regions of a pure state, yielding zero. In fact, the trivial index was inevitable. On the infinite lattice, ALPUs with nonzero index implement shifts, and these shifts have no analog on the finite interval with non-periodic boundary conditions.

We can therefore apply Theorem 9.19 about Hamiltonian evolutions to every ALPU on finite $\Gamma \subset \mathbb{Z}$. As above, the theorem becomes nontrivial lattices of a certain size, using the same threshold discussed above. We then obtain a local Hamiltonian evolution generating the ALPU, with locality as specified by Theorem 9.19.

While finite chains with non-periodic boundary conditions descend as a special case from the infinite lattice, the case of periodic boundary conditions (i.e. Γ inherits the

metric of a circle) appears more difficult. Many of the tools we develop appear useful there, but the key Lemma 9.11 has no obvious analog. Therefore we cannot offer a rigorous index theory of ALPUs on finite chains with periodic boundary conditions. The question is nonetheless important, and perhaps crucial for a generalization to higher dimensions. We leave the question to future work.

9.2.5 Translations cannot be implemented by local Hamiltonians

In this section, we discuss how our result answers the following natural question: do local 'momentum densities' on the one-dimensional lattice exist? In quantum many-body systems, local conserved quantities dramatically influence dynamics. For instance, under local Hamiltonian evolution, energy itself is a local conserved quantity, and after the system has locally equilibrated, the dynamics are often governed by energy diffusion. More generally, when a system admits more local conserved quantities in addition to energy, the near-equilibrium dynamics are often governed by the hydrodynamics of these quantities [LMMR14, DNBD19, BMEK17]. For translation-invariant systems, one expects momentum is also a local conserved quantity. For instance, in scalar quantum field theory, the i'th component of the total momentum operator may be expressed as $P^i = \int \pi(x) \partial_i \phi(x) \mathrm{d}x$ which is manifestly local, with local momentum density $\pi(x) \partial_i \phi(x)$.

The long-wavelength, low-energy regime of a lattice system like a spin chain is often described by a field theory, and a local momentum density is well-defined under this approximation. However, we might also ask for a local momentum operator $P = \sum_x p_x$ on the spin chain that generates translations, yielding $U = e^{iP}$ as the one-site translation operator. If P were constructed with local terms p_x, and if P commuted with some translation-invariant Hamiltonian, this exactly conserved momentum density might play an important role in dynamics.

The existence of such a local P is precisely the question of whether the shift QCA can be generated by a local "Hamiltonian," referring now to P as a Hamiltonian. We show such a local Hamiltonian cannot exist. In particular, on the infinite one-dimensional chain, it is impossible to implement the translation operator by time evolution using any time-dependent Hamiltonian satisfying Lieb-Robinson bounds, if the Lieb-Robinson bounds lead to an ALPU with $o(1)$-tails. This follows immediately from Theorem 9.25(iii)(d). For instance, we have:

Corollary 9.27 (No-go for local momentum densities)**.** *If P is a local Hamiltonian*

$$P = \sum_{X \subseteq \mathbb{Z}} P_X$$

on an infinite one-dimensional spin chain which has decaying interactions such that for all $n \in \mathbb{Z}$,

$$\sum_{\substack{X \subseteq \mathbb{Z} \\ s.t.\ n \in X}} \|P_X\| = \mathcal{O}\left(\mathrm{diam}(X)^{-(2+\varepsilon)}\right)$$

for some $\varepsilon > 0$, then e^{iP} cannot be the unitary lattice translation operator that translates by one site.

Recall our notation that e.g. P_X is a term local to region $X \subset \mathbb{Z}$ on the lattice $\mathbb{Z}$. The no-go result is also robust: Theorem 9.18 constrains how well e^{iP} can approximate the translation operator locally. (Note that while [GNVW12] already demonstrated that finite-depth circuits cannot achieve translations, their statements about circuits cannot be easily re-cast as claims about Hamiltonian evolution, at least not without further robustness results such as those developed here.)

Given that the translation operator cannot be generated by a finite-depth circuits, our analogous for claim for sufficiently local Hamiltonians might seem in intuitive. However the claim is not obvious, as demonstrated by the following example: if we allow evolution generated by Hamiltonians with $\frac{1}{r}$-decaying interaction terms (which then violate Lieb-Robinson bounds), we *can* implement a translation. The example involves a chain of qubits; we only sketch the construction but the details are easily verified. A Jordan-Wigner transformation maps the chain of qubits to a chain of fermions (or formally, it maps the quasi-local algebra to the CAR-algebra). Let $c_n^\dagger$ and c_n be the fermionic creation and annihilation operators at site $n \in \mathbb{Z}$. The Jordan-Wigner transform of the translation automorphism T is again the translation automorphism, $T(c_n) = c_{n-1}$. Taking a Fourier transform we see that

$$T(\hat{c}_k) = e^{ik}\hat{c}_k$$

Hence time evolution for time $t = 1$ using Hamiltonian

$$H = \int_{-\pi}^{\pi} k\hat{c}_k^\dagger \hat{c}_k \mathrm{d}k$$

implements T. In real space

$$H = \sum_{n,m \in \mathbb{Z}} h_{n-m} c_n^\dagger c_m$$

where the coefficients h_r (of which the precise form is not important) have magnitude $\frac{1}{r}$. Of course, we can also take the inverse Jordan-Wigner transform of this Hamiltonian to obtain a Hamiltonian on the spin chain

$$\tilde{H} = \sum_{n,m} h_{n-m} \sigma_{n,m}$$

where $\sigma_{n,m}$ is a Pauli operator supported on sites $\min\{n,m\}, \ldots, \max\{n,m\}$. In this way we can construct a Hamiltonian *not* satisfying Lieb-Robinson bounds which does implement T. This shows that our demand that the ALPUs have $o(1)$-tails in our construction of the index is not arbitrary; the classification by index collapses once we allow evolutions such as those generated by $\tilde{H}$ above with $\frac{1}{r}$-decaying interactions. In fact, by Theorem 9.25 we conclude that $e^{-i\tilde{H}t}$ cannot have $o(1)$-tails.

For the case of a single-particle Hamiltonian (i.e. a quantum walk), the obstruction to generating the translation operator with a local Hamiltonian hinges on the non-trivial winding of the dispersion relation [GNVW12]. It has been observed that for quadratic fermion Hamiltonians, every such Hamiltonian that implements the translation operator will need to have a discontinuity in its dispersion relation (in our example at $k = \pm\pi$) and hence at least $\frac{1}{r}$-tails in real space [ZFFM20, WW20]. These single-particle and free fermion results do not permit obvious generalization to the broader many-body case; our results allow us to draw conclusions for all local many-body Hamiltonians satisfying Lieb-Robinson bounds.

Discussion and open questions

In this chapter we have defined and studied the index for approximately locality-preserving unitaries (ALPUs) on spin chains. Various open questions remain, and at this point we will speculate on a few of these.

(i) Our results are restricted to the infinitely extended chain, or an open finite chain as in Section 9.2.4. One could also investigate what happens with a finite *periodic* chain with an ε-nearest neighbour automorphism for small ε. It appears that our proof technique relies on the fact that the chain is infinite (or open), so probably a different strategy is needed for finite periodic chains.

(ii) An obvious question of interest is the generalization to higher dimensions. The index theory generalizes to higher dimensions, but is not known to provide a complete classification in dimensions larger than two [FH20] as discussed in Section 6.1. Although the classification is less well understood, one could still hope that for any ALPU α there exists a sequence of QCAs α_j approximating α as in Theorem 9.16. Our constructions of approximating QCAs for an ALPU rely rather heavily on the structure theory (i.e., the GNVW index theory) of one-dimensional QCAs. Hence, it is not immediately clear how to generalize to higher dimensions. In fact, we have not even given a definition of what an ALPU is in higher dimensions, where some choices exist. For two-dimensional QCAs the notion of a boundary algebra allows one to classify all two-dimensional QCAs (in which any QCA is a composition of a circuit and a generalized shift). Potentially, this structure theory, as developed in [FH20, Haa21] can be used in a similar fashion to construct the α_j. This could involve proving stability results for the notion of a "visibly simple algebra" as introduced in [FH20]. A direct physical application of this would be a rigorous understanding of the index discussed in Section 6.2.1.

(iii) A basic open question is a generalization of Proposition 8.13, that is, to show that it is also true in spatial dimension greater than one that if α is a QCA which is sufficiently close to the identity, it is a circuit.

(iv) Another direction to generalize in is to channels which preserve locality but which are not unitary (i.e. an automorphism), see [PC20] for definitions and a recent discussion. In other words, what happens if the dynamics is slightly noisy? Is the index robust under small amounts of noise? One could hope to show that any locality preserving channel which is almost unitary can be approximated by a QCA.

(v) There is also a notion of *fermionic* QCAs, with a corresponding GNVW index. It should be possible to use similar arguments to extend the index to fermionic ALPUs.

Some of these open questions are closely related. For instance, it is easy to see that for any ALPU one can approximate by a strictly locality preserving channel (which is then not necessarily unitary). Thus, a solution for Item (iv) immediately provides a solution for approximating ALPUs by QCAs and could also be helpful for (i). Secondly, a

natural approach to two-dimensional systems could be a localizing in a radial direction, and approximation results for the periodic chain could be useful here. A final comment is that one method to show that on a finite chain ε-nearest neighbour automorphisms are close to QCAs would be by adapting the proof of Theorem 8.2. For this, consider an automorphism α, then, in the notation of Theorem 8.2, one could define a quantum channel

$$\Phi : \mathcal{B}_n \to \mathcal{C}_n, \qquad \Phi(x) = \mathbb{E}_{\mathcal{C}_n}(\alpha(x))$$

and let $\Psi = \Phi^\dagger \Phi$ where $\Phi^\dagger$ is the adjoint channel of Φ. If α were a nearest neighbour QCA, by the proof of Theorem 8.2 we would find that Ψ is a conditional expectation onto the subalgebra $\mathcal{L}_n$. On the other hand, if α is an ε-nearest neighbour automorphism we find that Ψ is a unital channel which is almost idempotent. It is a known fact that unital idempotent channels are always equivalent to a conditional expectation onto a subalgebra, so if we were able to show that Ψ is actually *close* to a strictly idempotent unital channel we could use this to establish, as in Theorem 8.2, a factorization result and construct a nearby strict QCA. Thus, in this approach the challenge would be to extend stability results and show that if a unital channel Ψ is almost idempotent (which means that $\|\Psi^2(x) - \Psi(x)\| \le \varepsilon \|x\|$), this implies that there exists a nearby unital idempotent channel $\tilde{\Psi}$ with $\|\Psi(x) - \tilde{\Psi}(x)\| \le f(\varepsilon)\|x\|$ for some function f. This could also provide alternative proofs for the results in this chapter which would not rely on infinite dimensional algebras.

Part III

Entanglement in random tensor networks

Introduction to quantum information in quantum gravity

Since the discovery of the AdS/CFT correspondence and holography there has been a fruitful interaction between quantum gravity and quantum information. In this part of the dissertation we contribute to this interaction. We start by giving a brief introduction to the role of quantum information theory in quantum gravity. One of the basic results is the computation of CFT entanglement entropies by areas of minimal surfaces. In Chapter 11 we introduce a basic tool for computing entropies in quantum field theories and in holography: the replica trick. We review recent work which computes the entanglement entropy in situations where there are multiple relevant minimal surfaces. The material in the current introduction and in Chapter 11 gives a very brisk review of holographic quantum gravity and the computation of entropies and is not mathematically rigorous, but serves to explain the relevance of the precise computations in Chapter 12, Chapter 13 and Chapter 14. The results and exposition follow based on [CPWW].

In Chapter 12 we introduce the toy model we will study in this part of the dissertation: random tensor network states. We explain how a similar replica trick can be used to compute entanglement entropies for this model as well. We introduce a generalization of the random tensor network model, allowing arbitrary link states, which we will see reproduces holographic theories more realistically.

Our main contributions can be found in Chapter 13 and Chapter 14, where we show that random tensor network states reproduce the entropy computations reviewed in Chapter 11. In this case we can make these computations completely rigorous. The methods are related to the theory of free probability and to one-shot quantum information theory.

10.1   The black hole information paradox

Black holes provide one of the main motivations for studying quantum gravity. What happens to an object when it disappears into a black hole? In classical general relativity one can more or less avoid answering this question, since the object can never be retrieved from the black hole. A famous calculation by Hawking shows that an analysis of quantum field theory near the black hole horizon implies that the black hole is a radiating object (sending out *Hawking radiation*) at some nonzero temperature [Haw75]. This calculation can be performed without knowledge of a complete theory of quantum

gravity, since the gravitational force at the black hole horizon is still sufficiently weak to treat semiclassically. Mysteriously, this radiation appears to be completely thermal, in the sense that it contains no information about what has formed the black hole. By emitting Hawking radiation the black hole loses mass, and in the end the black hole will completely evaporate. This poses serious conceptual difficulties. If one believes that any quantum theory of a closed system must be unitary (that is, reversible), this is in conflict with black hole evaporation as described above: the process of matter falling into the black hole, and returning as Hawking radiation, appears to be irreversible, and there seems to be a loss of information. This forms the basis for what is known as the *black hole information paradox*. Another perspective on information in the process of black hole evaporation is the *Page curve*. This is the conjectured development of the entanglement entropy between the emitted radiation and the black hole interior. Initially, while the black hole is emitting Hawking radiation, as this radiation is purely thermal (so one can roughly think of the state of the radiation as maximally mixed), the entanglement entropy increases (and one can think of the radiation as being maximally entangled with the black hole interior). However, if we have unitary evolution, after evaporation of the black hole the entanglement entropy must be zero, so if one assumes unitarity there will be a point at which the entanglement entropy between the black hole and the radiation starts to decrease. The point at which this happens is essentially when the number of degrees of freedom of the radiation system becomes larger than the number of degrees of freedom in the black hole interior, see Fig. 10.1a. At this point, we generally still do not need a theory of quantum gravity near the black hole horizon, which is still sufficiently far from the singularity for semiclassical quantum gravity to apply; this makes the expected decrease of entropy puzzling. A sharp version of the black hole information paradox was formulated as the AMPS paradox in [AMPS13]. The black hole information paradox has been the starting point for a large body of research on widely differing potential resolutions [Pre92, HM04, HP07, ST08, AMPS13, HH13, PR13, PR14]. See [Har16] for a review of quantum information and black hole physics, and quantitative statements of the black hole information paradox. Below we will discuss one specific approach based on insights from holographic quantum gravity.

An important insight is that the entropy of a black hole is proportional to the *area* of its horizon. To be precise, the thermodynamic entropy is give, in $c = \hbar = 1$ units, by

$$H_{\text{BH}} = \frac{A}{4G_N} \tag{10.1}$$

where A is the area of the black hole horizon. Initially, this expression for the entropy was based on analogies to thermodynamics and from the computation of the temperature of the Hawking radiation. Later, it was also shown that if one counts the number of microstates corresponding to a black hole in certain string theories one recovers Eq. (10.1) [SV96], providing strong evidence that Eq. (10.1) can really be seen as the number of microstates corresponding to a macroscopic black hole of a prescribed area.

10.1.1   Random unitary model for black hole evaporation

To make sense of the black hole information paradox we will consider very basic toy models for black hole evaporation which show qualitatively similar behavior based on random unitaries [Pag93b, HP07]. We start with a toy model which reproduces the Page

(a) A sketch of the Page curve.

(b) A random unitary model which reproduces the Page curve, and the set-up for an old black hole in the Hayden-Preskill model.
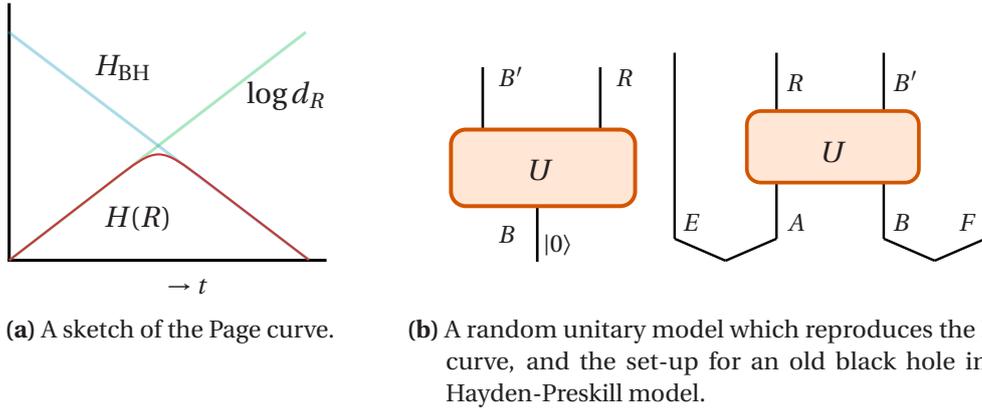
Figure 10.1: The entanglement entropy of a radiating black hole develops along the Page curve: initially, the entanglement entropy of the black hole equals the number of radiation qubits $\log d_R$, but at some point the total number of degrees of freedom of the black hole, given by Eq. (10.1) is smaller, and gives an upper bound on the entanglement entropy.

curve. If we assume unitary time evolution, and we assume that the emitted Hawking radiation is approximately maximally mixed, then during the evaporation process the entanglement entropy between the radiation and the black hole interior has to grow linearly. However, during the evaporation process, the black hole shrinks, and therefore the total entropy of the black hole (which is proportional to its horizon area) decreases. At some point (the *Page time*), the black hole entropy will be smaller than the linearly increasing entropy of the radiation system. This is an upper bound on the entanglement entropy as well (as it is an upper bound on the system size). The Page curve is the conjecture that the actual entanglement entropy between the radiation and the black hole interior is given by the maximal possible entanglement entropy: it is the minimum between the (increasing) size of the radiation entropy and the (decreasing) size of the black hole entropy.

A very basic model reproducing this behavior is the following. Let B be a finite-dimensional quantum system with Hilbert space $\mathcal{H}_B$ of dimension d_B and consider a fixed initial state $|0\rangle$. Then we 'model' the black hole evaporation process by applying a *Haar-random unitary U* to the initial state, and splitting up the system $B = B'R$, where R is a system of dimension d_R, representing the emitted radiation and B' is a system of dimension $d_{B'}$ representing the remainder of the black hole. The result is a Haar-random state $|\psi\rangle = U|0\rangle$ (corresponding to a uniformly random unit vector in $\mathcal{H}_B$), see Fig. 10.1b. It is a well-known result that such state has, with high probability, almost maximal entanglement [Pag93a, HLW06]. That is, with high probability

$$H(R)_\psi \approx \min\{\log d_R, \log d_{B'}\} \tag{10.2}$$

which corresponds to the Page curve: the entropy grows linearly with the number of quanta of Hawking radiation which are emitted, up to the point where the black hole entropy is too small.

Let us also consider a slightly more elaborate model, proposed by [HP07]. Here we investigate the following question: suppose we have an already existing black hole, and we throw in some object, when can we recover the object from the radiation? First we must explain what we mean by 'recovering'. Suppose we have an observer Alice,

who has a quantum system system A of dimension d_A. We let $|\phi_{AE}^+\rangle$ be a maximally entangled state between Alice and a reference system E of the same dimension. Then, after disposing the A system into the black hole, and collecting radiation R, we say that we can recover A if R is again maximally entangled with E. This is a standard notion of recoverability in quantum information theory. It implies that we can now apply a channel to the radiation to get back the original state ϕ_{AE}, which would recover any initial state of the A system, preserving correlations with a reference system E An approximately maximally entangled state suffices for approximate recovery. Let us apply this to set up a toy model for black hole evaporation. Let B denote the system of the original black hole. We consider two different scenarios. The first scenario is where we have a 'young' black hole; in this case the B system is initially in a state $|0\rangle$. The second situation models an 'old' black hole, where we have already collected a large amount of Hawking radiation in a system F, which we assume to be maximally entangled with B, see Fig. 10.1b. So, the initial state is given by $|\phi_{AE}^+\rangle \otimes |0_B\rangle$ for a young black hole or $|\phi_{AE}^+\rangle \otimes |\phi_{BF}^+\rangle$ for an old black hole. Then one can show, using *decoupling* techniques from quantum information theory that for a young black hole, and has to wait until just after the Page time before you can (with high probability) recover A. On the other hand, for an old black hole, one can recover A from the total radiation system RF almost immediately after A falls into the black hole! That is, one only needs $\mathcal{O}(1)$ qubits of radiation R to recover. For this reason, [HP07] called old black holes 'information mirrors'.

We may say that after the Page time the black hole interior is partially encoded in the radiation system. This conflicts with usual notions of *locality* for quantum field theory. However, the actual process of recovery can be very complicated; in fact, this may be so complicated that for all practical purposes recovery is not possible. See [HH13, KTP20] for arguments that the decoding process (after the Page time) may have exponential complexity, and that for this reason observers with polynomial computational resources will not be able to observe potential breakdowns of locality.

While clearly a very crude model, we would like to comment on why using random unitaries makes sense as a toy model for a black hole. It has been argued that black holes should have highly chaotic dynamics [SS08, SS14, MSS16]. The intuition for this is roughly as follows: if matter falls into a stationary black hole, there is a perturbation of the black hole which is damped exponentially. After a very short time, the black hole is again in a stationary state, and the black hole has equilibrated. This fast equilibration should mean that the black hole is highly chaotic (in some sense, maximally chaotic, as argued in [MSS16]), and the *scrambling time* of the black hole as a quantum system should correspond to the equilibration time as computed gravitationally. For times longer than the scrambling time we can therefore model the black hole dynamics by a random unitary, as chaotic quantum dynamics behave similar to random dynamics, and one can actually use random matrix theory to understand black hole dynamics [CGAH⁺17].

The random unitary toy models should be seen as an indication of what will happen to quantum information provided we assume the black hole dynamics are unitary. They do not provide sufficient detail to 'solve' the black hole information paradox: they do not propose a mechanism to correct the entropy computations for the Hawking radiation. In this dissertation we will study a more sophisticated random matrix theory model for quantum gravity, and we will make extensive use of similar decoupling results. We will

investigate what happens in this model at the Page transition, where $\log d_R \approx \log d_{B'}$ in Eq. (10.2). We will do so in the context of *holographic quantum gravity.*

10.2 Quantum information in holography

More than twenty years after its discovery, the AdS/CFT correspondence [Mal99] remains arguably the only known example of a theory of quantum gravity; indeed it is unique in being a nonperturbatively defined theory for which we have strong evidence for the existence of a semiclassical limit consisting of Einstein gravity coupled to quantum field theory. Holography states that there is an equivalence between two theories: on the one hand a quantum gravity theory on a asymptotically Ant-de Sitter space (the *bulk*), and on the other hand a conformal field theory on its conformal *boundary*. This phenomenon is also known as holography [Hoo93, Sus95] and is inspired by the fact that the black hole entropy scales with the *area* as we saw in Eq. (10.1). The strongest statement of the AdS/CFT correspondence is the conjecture that there is an equivalence of partition functions. Suppose that there exist bulk fields ϕ (amongst which the gravitational metric), then we let

$$Z_{\text{AdS}}[\phi_0] = \int_{\phi|_{\partial M} = \phi_0} D\phi \, e^{-S_{\text{AdS}}(\phi)}$$

be the path integral over the bulk fields, where $S_{\text{AdS}}(\phi)$ is the (quantum) gravity action, and we have set boundary conditions at the conformal boundary ∂M. On the other hand, we may consider the CFT partition function

$$Z_{\text{CFT}}[\phi_0] = \int D\psi \, e^{-S_{\text{CFT}}(\psi) - \int_{\partial M} \phi_0(x) O(x)}$$

where we perform the path integral over the CFT fields, and we have added operators O using the (classical) fields ϕ_0 as sources. Then one statement of the AdS/CFT correspondence is that these partition functions coincide:

$$Z_{\text{AdS}}[\phi_0] = Z_{\text{CFT}}[\phi_0].$$

This allows one to relate various quantities in the gravitational and CFT theories, a correspondence known as the *AdS/CFT dictionary*. Known examples of holographic duality are typically such that the bulk theory is a string theory and the boundary theory is a supersymmetric model. The boundary CFT has a central charge c, which effectively counts the number of degrees of freedom of the theory. For example, for the most well-known example of $\mathcal{N} = 4$ supersymmetric Yang-Mills theory with $SU(N)$ gauge fields, which is dual to a type IIB string theory, we have $c = N^2$. It turns out that correspondingly on the gravitational side $G_N \sim c^{-1}$. This means that weakly interacting gravity corresponds to large c, or large N. In this limit we may approximate the gravitational path integral

$$Z_{\text{AdS}}[\phi_0] \sim e^{-I(\phi)} + \dots \tag{10.3}$$

by its saddle point, corresponding to the classical solution ϕ of the gravitational equations of motion with boundary conditions ϕ_0, and where $I(\phi)$ is the action of the classical

limit of the gravitational theory (for instance, this is a supergravity action in the case of type IIB string theory). The vacuum state of the CFT is dual to empty AdS space. As we turn on sources in the boundary, we find states which are dual to other semiclassical geometries. Below we will assume that we are in the semiclassical regime (with $c \gg 1$) and we consider states ρ in the CFT which are dual to semiclassical geometries in the bulk.

10.2.1 Computing entropies

A crucial feature of the AdS/CFT correspondence is that the emergence of a (classical) space-time is closely related to the entanglement structure of the boundary theory. One of the ways to see this is by computing entanglement entropies for the boundary theory. An extensive introduction to entanglement entropies in holography can be found in [RT17]. We first consider the case where we have a (pure) CFT state ρ which is dual to a static space-time (so the space-time looks like $\mathbb{R} \times \Sigma$ for some fixed spatial manifold Σ). We then consider the boundary CFT on the conformal boundary $\partial\Sigma$, and we let A be a subregion. What is the entanglement entropy $H(\rho_A)$? At this point we comment that strictly speaking it is not very clear what the 'reduced density matrix' ρ_A should be: in a quantum field theory we do not have a factorization of Hilbert spaces with respect to a spatial region and its complement. Also, the entropy of the A-system will be infinite. We implicitly place a UV cut-off on the theory and throughout this chapter and Chapter 11 pretend that we have finite dimensional Hilbert spaces, a useful and not too harmful fiction.

To give a prescription for $H(\rho_A)$, consider the set $C(A)$ of all co-dimension 1 surfaces γ_A in Σ which are homologous to A; in particular this means that $\partial\gamma_A = \partial A$. Then, the *Ryu-Takayanagi (RT) formula* [RT06a, LM13] states that

$$H(\rho_A) = \min_{\gamma_A \in C(A)} \frac{|\gamma_A|}{4G_N} + \mathcal{O}(1) \tag{10.4}$$

where $|\gamma_A|$ is the *area* of the surface γ_A, and the corrections are of constant order with respect to G_N. The expression in Eq. (10.4) has suppressed the dependence on the cut-off. Formally, both $H(\rho_A)$ and $|\gamma_A|$ are infinite, and we regulated ρ_A by a UV cut-off, and we may similarly regulate the area of the surface by cutting off γ_A some distance from the boundary. This can be done such that the regulated expressions match. The RT formula is strongly reminiscent of the area formula for the black hole entropy in Eq. (10.1).

Various generalizations exist. One can allow for non-stationary space-times; in this case, we consider a subsystem A of a spatial slice of the boundary, and we consider the set $X(A)$ of all co-dimension 2 surfaces γ_A which are homologous to A (but now not restricted to a fixed spatial slice) and which are an *extremal* point of the area functional. Then the *Hubeny-Rangamani-Takayanagi (HRT) formula* [HRT07] extends the RT-formula by expressing the entropy as

$$H(\rho_A) = \min_{\gamma_A \in X(A)} \frac{|\gamma_A|}{4G_N} + \mathcal{O}(1). \tag{10.5}$$

Another possibility is to allow the bulk space to have 'entropy' itself, for instance in the form of matter or a black hole, in which case the dual boundary state need not be pure. For this, given an extremal surface γ_A on a spatial slice, let Γ_A be the bulk region region

enclosed by γ_A and A. This is known as the *entanglement wedge*. We denote by $S_{\text{bulk}}(\Gamma_A)$ the entropy of the bulk state reduced to Γ_A. It has been conjectured [EW15] that we can now compute the entropy of ρ_A by the *quantum extremal surface formula*:

$$H(\rho_A) = \min_{\gamma_A \in X(A)} \left\{ \frac{|\gamma_A|}{4G_N} + S_{\text{bulk}}(\Gamma_A) \right\} + \mathcal{O}(1). \tag{10.6}$$

Given γ_A, let $H_{\text{QES}}(\gamma_A) = \frac{|\gamma_A|}{4G_N} + S_{\text{bulk}}(\Gamma_A)$ The quantum extremal surface formula is highly relevant for the black hole information paradox, as was shown in [Pen20, AEMM19]. To see this, consider a state ρ which is dual to a black hole. We collect the Hawking radiation R at the boundary system, and we would like to reproduce the Page curve by computing the entropy of the radiation $H(\rho_R)$. One can show that there are two potential quantum extremal surfaces [Pen20]. One candidate is the empty surface. In this case Γ_A is the complete spatial slice, and $H_{\text{QES}}(\gamma_A)$ equals the entropy of the black hole, which equals the entropy of the Hawking radiation collected and grows linearly. The other option is a surface γ_A just inside the black hole horizon. For this surface, $H_{\text{QES}}(\gamma_A)$ is given by (approximately) the black hole horizon area. Taking the minimum of these two options we indeed find the Page curve. This idea has lead to a better understanding of the black hole information paradox. In this derivation, the assumption in the paradox which is broken is a locality assumption: after the Page time the black hole interior is partially encoded in the radiation. Here, the derivation of the Page curve relies on Eq. (10.6), which is still of a conjectural nature. However, Eq. (10.6) can be verified in certain restricted situations [PSSY19, AHM$^+$20], and is closely related to so-called replica wormholes (which we will discuss in Chapter 11). The surprising feature of this approach to the black hole information paradox is that is that the encoding of the black hole interior in the radiation can be realized with an essentially semiclassical computation. A review of these recent developments can be found in [AHM$^+$21].

An interesting subtlety in the quantum extremal surface formula is the precise conditions that ρ has to satisfy. We demanded that ρ should be dual to a classical geometry (in order for the notion of a minimal surface to make sense). However, if the $S_{\text{bulk}}(\Gamma_A)$ term becomes relevant in the minimization problem we also need the bulk matter to be sufficiently well-behaved, as was pointed out in [AP20], where the application of *one-shot quantum information theory* was introduced to understand the conditions under which the quantum extremal surface is valid.

10.2.2 Recovery and entanglement wedge reconstruction

An extension of the entropy computations by means of minimal surfaces is *subregion-subregion duality*, or *entanglement wedge reconstruction*. Informally speaking, this states that the reduced state ρ_A on a boundary subsystem A is dual to the entanglement wedge Γ_A, which is the bulk region bounded by A and the minimal surface γ_A for A. What does it mean for Γ_A to be 'dual' to ρ_A? The idea is that if we act with a local operator in Γ_A, this is equivalent to acting with a local operator on A. However, we should be careful: if we act with an operator which actually deforms the geometry, or creates a superposition of different geometries this becomes ambiguous. For this reason we consider a *code subspace S* of bulk states, which all have the same bulk geometry, with some low-energy excitations. Then, each of these states is dual to a boundary state. As in our discussion of the black hole information paradox, let ϕ_{SE} be a maximally

entangled state between the code subspace S and a reference system E. Then, the state ϕ is dual to a state ρ which is a maximally entangled state between E and the boundary, from which we can completely recover the bulk state ϕ_{SE}. If we restrict to a boundary subsystem A, so we have the reduced state ρ_{AE}, what can we still recover? The answer turns out to be $\phi_{\Gamma_A E}$, the reduced bulk state on the entanglement wedge. This can be framed in the language of *error correcting codes*: we have quantum information encoded in the bulk theory, and the information which is encoded in the entanglement wedge Γ_A is protected against erasure error on the complement of A. See [CHP$^+$19] for an information-theoretic perspective on entanglement wedge reconstruction. A good introduction, with an emphasis on the relation to notions on locality in the bulk, can be found in [Har18]. Finally we comment that for entanglement wedge reconstruction the role of one-shot information theory has also been investigated in [AP20, AP22].
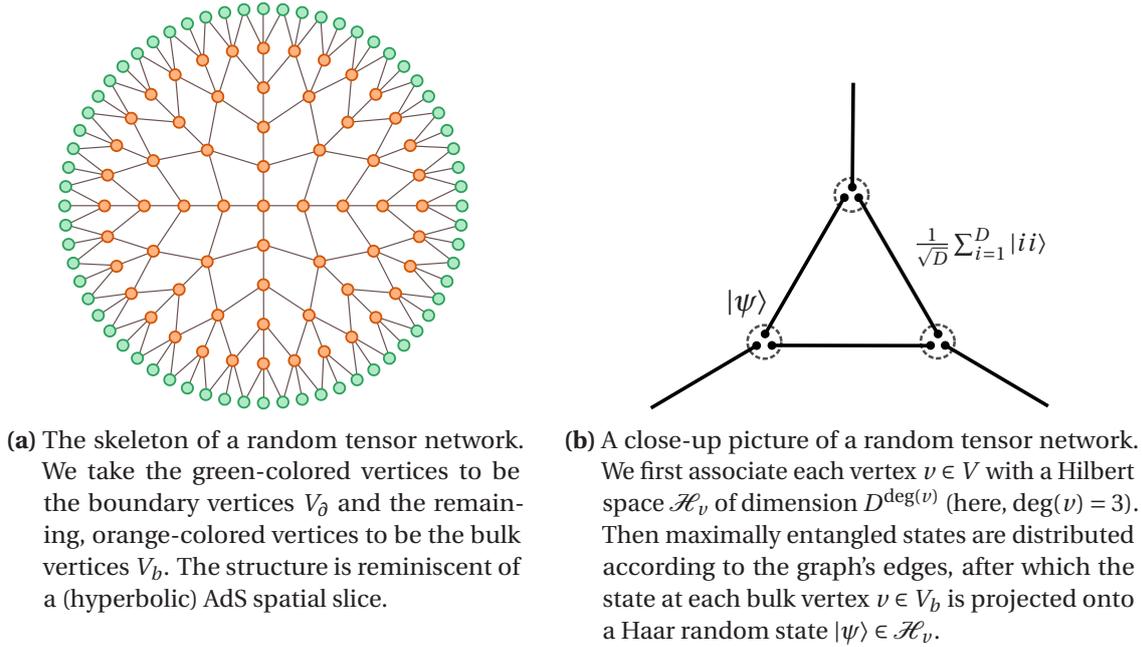
10.2.3 Holography from random tensor networks

A careful study of quantum information principles in holography is complicated by the fact that one has to deal with complicated strongly interacting field theories. It has been useful to construct *toy models* which are not themselves realistic models of quantum gravity, but nevertheless capture certain structural aspects of holography. One of the most powerful such models are *random tensor network states*, which are the topic of this part of the dissertation. Random tensors and tensor networks also arise in a number of other fields of physics, including quantum information, where they have been used to explore generic entanglement properties of quantum states [HLW06, CNŻ10, Aub12, ASY12, AN12, CGGPG13, CNŻ13, CN16, AS17, Has17, NW20, WW21a, MB21, LPG21] and condensed matter physics, e.g. in the study of random circuits and measurements [VPYL19, LPWV20, NRSR21, MVS21, YLFC21, LC21, LVFL21].

A random tensor network can be seen as a PEPS state where the choice of tensors is random. The most basic version of a random tensor network is characterized by a choice of bond dimension D and a graph $G = (V, E)$, where the vertices $V = V_b \sqcup V_\partial$ of G are partitioned into "bulk" vertices V_b and "boundary" vertices V_∂. To each edge $e \in E$, we associate a maximally entangled state

$$\frac{1}{\sqrt{D}} \sum_{i=1}^{D} |ii\rangle \tag{10.7}$$

on two D-dimensional Hilbert spaces, one of which is associated to each endpoint of e; each vertex $v \in V$ is therefore associated with a Hilbert space $\mathcal{H}_v$ of dimension $D^{\deg(v)}$. Finally, we project each *bulk* vertex $y \in V_b$ into a Haar random state $|\psi_y\rangle \in \mathcal{H}_y$. The resulting 'random tensor network state' lives in the Hilbert space $\mathcal{H}_\partial = \otimes_{x \in V_\partial} \mathcal{H}_x$ associated to the boundary vertices $x \in V_\partial$, as shown in Fig. 10.2.

To characterize the typical entanglement structure of random tensor network states, we can compute the von Neumann entropy $H(\rho_A)$ of the reduced density matrix ρ_A on a subset $A \subset V_\partial$ of the boundary vertices. In the limit where the bond dimension D is very large, this entropy can be shown to converge with high probability to $\log(D)|\gamma_A|$, where γ_A is the set of edges crossing the minimal cut (for the moment, assumed to be the unique such cut) in the graph separating A from its boundary complement $V_\partial \setminus A$ (see Fig. 10.5a). This formula is closely analogous to the holographic RT formula. Similarly, if we allow entropy in the bulk, one can show that an analog of the quantum extremal

(a) The skeleton of a random tensor network. We take the green-colored vertices to be the boundary vertices V_{∂} and the remaining, orange-colored vertices to be the bulk vertices V_b. The structure is reminiscent of a (hyperbolic) AdS spatial slice.

(b) A close-up picture of a random tensor network. We first associate each vertex $v \in V$ with a Hilbert space $\mathcal{H}_v$ of dimension $D^{\deg(v)}$ (here, $\deg(v) = 3$). Then maximally entangled states are distributed according to the graph's edges, after which the state at each bulk vertex $v \in V_b$ is projected onto a Haar random state $|\psi\rangle \in \mathcal{H}_v$.

Figure 10.2: The basic structure of a random tensor network.

surface prescription is valid, and prove subregion-subregion duality for random tensor networks.

In quantum field theories it is often convenient to study the k-th Rényi entropies

$$H_k(\rho_A) = \frac{1}{1-k} \log \mathrm{tr}[\rho_A^k].$$

For integer $k > 1$, these are more amenable to direct computation than the von Neumann entropy, and one can often extract the von Neumann entropy by analytic continuation to $k = 1$.

The computation of Rényi entropies in random tensor network models is in fact closely analogous to holographic computations. In both cases, the idea is to use the *replica trick* – essentially this is the observation that $\mathrm{tr}[\rho_A^k] = \mathrm{tr}[\tau \rho_A^{\otimes k}]$ where τ is an operator which permutes the k copies of ρ_A cyclically. In the holographic computation, this can be written as a path integral, on k copies of the theory, glued together in an appropriate way. By the holographic dictionary this path integral can then be computed by the action of a bulk geometry with certain boundary conditions [LM13]. We will review this in Chapter 11. For random tensor networks, one finds that $\mathrm{tr}[\rho_A^k]$ concentrates around its expectation, and that this can be computed as the partition function of a classical spin model on the bulk vertices, with boundary conditions dictated by the choice of boundary subsystem [HNQ+16]. This computation will be explained in detail in Chapter 12.

10.2.4   A single random tensor and non-crossing partitions

For now, let us review some well-known results for the easiest version of this computation, which is the case of a single random tensor. Consider a tensor $|\psi\rangle$ of size $d_A \times d_{\bar{A}}$. We have two boundary systems, A of dimension d_A and $\bar{A}$ of dimension $d_{\bar{A}}$. There are two edges, and the edge with minimal dimension is a 'minimal' cut. We are especially

interested in the case where d_A and $d_{\bar{A}}$ are of the same order of magnitude. In this case there are two competing 'minimal cuts'. This can be thought of as the point of the 'phase transition' in the Page curve in Eq. (10.2). This situation leads to $\mathcal{O}(1)$ corrections to the entropy which is of order $\min\{\log d_A, \log d_{\bar{A}}\}$.

We introduce some notation for permutations $\pi \in S_k$, the symmetric group on k elements. Given $\pi \in S_k$, let $C(\pi)$ be the cycle type of the permutation, i.e. the collection of lengths of the disjoint cycles in π. Then $|C(\pi)|$ is the number of cycles of π. For $\pi_1, \pi_2 \in S_k$ we let $d(\pi_1, \pi_2) = k - |C(\pi_1^{-1}\pi_2)|$, which gives a metric on S_k (see Chapter 13 for details). We let $\tau = (12\ldots k)$ be the full cycle. The replica trick allows one to show that if we let ρ denote the associated random tensor network state (which, in this case, simply is a normalized version of ψ), then

$$\mathbb{E}\operatorname{tr}\left[\rho_A^k\right] = \sum_{\pi \in S_k} d_A^{d(\tau, \pi)} d_{\bar{A}}^{d(\pi, \mathrm{id})}$$

We will derive a generalization of this expression for general random tensor networks in Chapter 12. A first observation is that if $d_A \ll d_{\bar{A}}$, there is a single dominant contribution, given by $\pi = \tau$. Similarly, if $d_A \gg d_{\bar{A}}$, the dominant contribution arises from $\pi = \mathrm{id}$.

We now would like to investigate the regime where $d_A = D$ and $d_{\bar{A}} = \alpha D$, where D is large and α goes to some constant. We need a property of permutations: for any permutation π, $d(\tau, \pi) + d(\pi, \mathrm{id}) \geq k-1$. The permutations for which we have equality in this expression correspond to *non-crossing partitions* $NC(k)$ (again, see Chapter 13 for details). These are permutations which can be 'drawn' in the the plane without crossing. This means that

$$\mathbb{E}\operatorname{tr}\left[\rho_A^k\right] = D^{-(k-1)} \sum_{\pi \in NC(k)} \alpha^{|C(\pi)|-k} + \mathcal{O}(D^{-k})$$
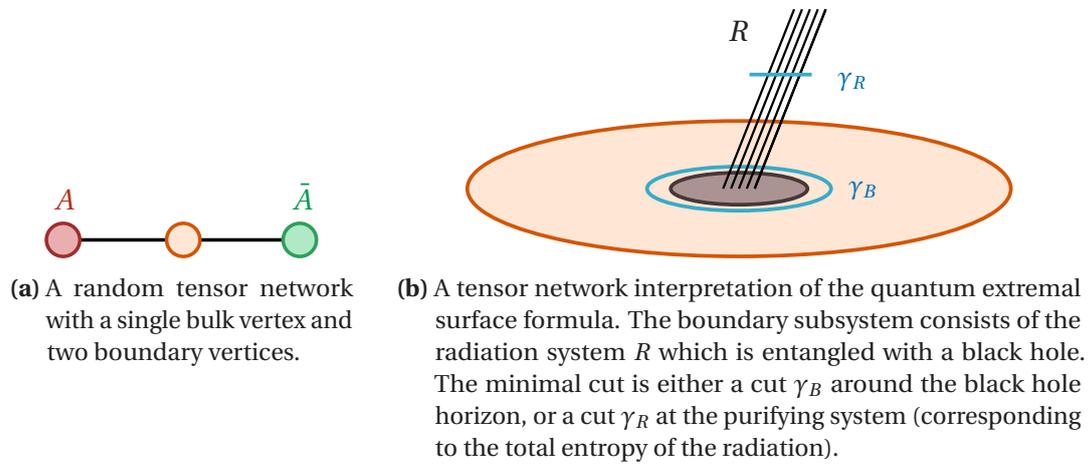
For large D we may ignore the $\mathcal{O}(D^{-k})$ contributions, and use this to show that the spectrum of $D\rho_A$ converges to a *Marchenko-Pastur* contribution $\mathrm{MP}(\alpha)$ (see Chapter 13 for details). In particular, this allows one to compute $\mathcal{O}(1)$ corrections to the $\log(D)$ leading contribution to the entropy. For instance, one can show that, with high probability for large D

$$H(\rho_A) \approx \log(D) - \frac{1}{2\alpha}.$$

One of our goals will be to generalize this result to a wide class of random tensor networks.

10.3   Summary of contributions

In Section 10.2.1 we saw that the quantum extremal surface prescription for holographic entropies leads to a derivation of the Page curve. In particular, at the Page time there is a 'phase transition' in which surface is the quantum extremal surface. A similar phase transition occurs when we consider a boundary region A consisting of two disjoint subregions A_1 and A_2. Then, upon varying the size or location of A_1 and A_2 there can be a change in whether the entanglement wedge for A is connected, or is the disjoint union of the entanglement wedges for A_1 and A_2 separately, see Fig. 10.4. In both cases, we would like to understand what happens at this transition point.

(a) A random tensor network with a single bulk vertex and two boundary vertices.

(b) A tensor network interpretation of the quantum extremal surface formula. The boundary subsystem consists of the radiation system R which is entangled with a black hole. The minimal cut is either a cut γ_B around the black hole horizon, or a cut γ_R at the purifying system (corresponding to the total entropy of the radiation).

Figure 10.3: The most basic example of a random tensor network with a single tensor corresponds to the model in Fig. 10.1b. Tensor network models provide intuition for the quantum extremal surface formula.
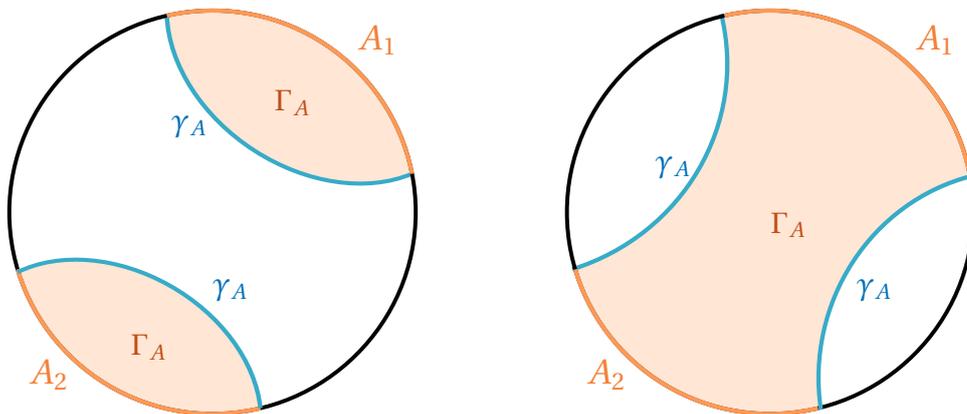
Figure 10.4: If the boundary subsystem A consists of two connected components A_1 and A_2 there are two possibilities for the RT-surfaces. Either the RT surface γ_A for A is the union of the RT surfaces for A_1 and A_2, in which case the entanglement wedge is disconnected, or this is not the case and the entanglement wedge Γ_A is connected. In the first case, there is essentially no entanglement between A_1 and A_2, whereas in the second case there is entanglement between A_1 and A_2.

While there is a strong analogy between random tensor networks and holographic CFT states, from the replica trick computations one finds that holographic CFT states and random tensor network states have Rényi entropies which behave quite differently when $k \neq 1$. For random tensor network states the Rényi entropies are approximately independent of k in the large D limit. This means that their entanglement spectrum is close to 'flat', or, in other words, that the boundary state ρ_A is approximately maximally mixed within a certain subspace.

On the other hand, CFT states that are dual to semiclassical spacetime geometries have Rényi entropies that vary non-trivially with k, as we will see in Chapter 11. It has recently been argued that *fixed-area states* in AdS/CFT do have flat spectra, and more generally have an entanglement structure that closely matches random tensor network states [AR19, DHM19, BPSW19, MWW20, DQW21]. Such states have a well-defined semiclassical geometry associated to a fixed spatial slice; however, thanks to the uncertainty principle, they cannot describe a single semiclassical *spacetime* geometry [BPSW19].

In the random tensor network model, the flatness of the spectrum can be traced to the maximally-entangled states used as 'link states' (see Eq. (10.7) and Fig. 10.2b) on the edges of the graph, which themselves have flat entanglement spectra. To take results about random tensor networks beyond the fixed-area state regime, it is natural – see, e.g., discussion in [HNQ⁺16, BPSW19] – to replace the maximally entangled link states by general states

$$|\phi_e\rangle = \sum_{i=1}^{D} \sqrt{\lambda_i}\, |ii\rangle\,. \tag{10.8}$$

If one considers a random tensor network with non-trivial link states, where we simply replace each maximally entangled state by some fixed other state, it is perhaps not very surprising to see that if there is a single minimal cut for a subsystem A then the resulting density matrix ρ_A will have an entanglement spectrum that converges to that of $|\gamma_A|$ copies of the link state along the minimal cut as $D \rightarrow \infty$; indeed this was already suggested in [HNQ⁺16].

A more interesting question, and the main focus of this work, is the case where there are two minimal cuts, as in Fig. 10.5b. This situation is motivated by questions in holography: it can be used to study the phase transition at the point where there are two competing minimal surfaces [MWW20]. Moreover, as we saw in Section 10.2.1 this situation is relevant to understand the Page transition in the black hole evaporation process. For a single random tensor we sketched how the replica trick can be used to compute $\mathcal{O}(1)$ corrections to the entanglement entropy. We would like to be able to understand such corrections for random tensor network states with non-trivial link states.

For our first result, we consider the situation where the ratios λ_i/λ_j of different eigenvalues remain bounded in the $D \rightarrow \infty$ limit; we refer to this as the *bounded spectral variation* limit. Formally, we consider a family of link states with increasing bond dimension D as in Eq. (10.8). For each D, the link state has an associated distribution

$$\mu_e^{(D)} = \frac{1}{D} \sum_{i=1}^{D} \delta_{D\lambda_i}\,.$$
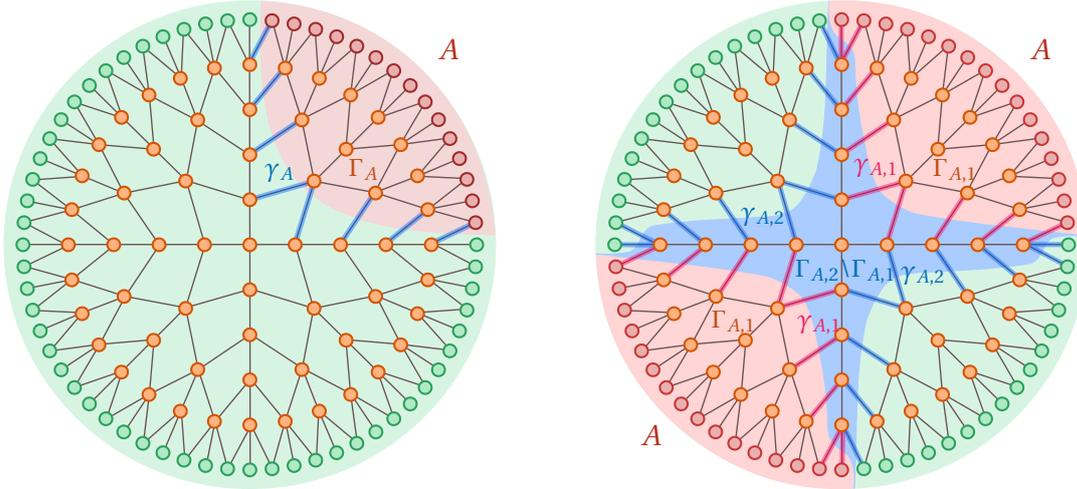
(a) The boundary domain A has a unique minimal cut Γ_A. The set of edges crossing this minimal cut is denoted by γ_A.

(b) The boundary domain A has two minimal cuts $\Gamma_{A,i}$ with cut-sets $\gamma_{A,i}$. In the notation of Section 12.1.2, the red region is $\Gamma_{A,1} = V_1$, the blue region is $\Gamma_{A,2} \setminus \Gamma_{A,1} = V_2$ and the green region is $V \setminus \Gamma_{A,2} = V_3$. Note that $\Gamma_{A,1}$ has two connected components, while $\Gamma_{A,2}$ is connected.

Figure 10.5: Tensor networks with one and two minimal cuts, compare with the competing surfaces in Fig. 10.4.

We then require that the moments

$$m_k^{(D)} = D^{k-1} \sum_{i=1}^{D} \lambda_i^k$$

of the distributions μ_e converge to a finite limit as $D \to \infty$ for all positive integer k. This means, in particular, that we must have $\lambda_i = O(1/D)$ for almost all eigenvalues λ_i.

If we let γ_A denote a minimal cut (more precisely, the set of edges crossing a minimal cut) for a boundary domain A, then we may similarly define the associated distribution

$$\mu_{\gamma_A}^{(D)} = \frac{1}{D^{|\gamma_A|}} \sum_{i_1,\dots,i_{|\gamma_A|}} \delta_{D^{|\gamma_A|}\lambda_{i_1}\cdots\lambda_{i_{|\gamma_A|}}},$$

and assume that it converges weakly to a some distribution μ_{γ_A}. Now consider the spectrum of the reduced state ρ_A, and denote its eigenvalues in non-increasing order by $\lambda_{A,i}$. Consider the empirical distribution of eigenvalues

$$\mu_A^{(D)} = \frac{1}{D^{|\gamma_A|}} \sum_{i=1}^{D^{|\gamma_A|}} \delta_{D^{|\gamma_A|}\lambda_{A,i}}.$$

In the case where there are two non-intersecting minimal cuts $\gamma_{A,1}$ and $\gamma_{A,2}$, we have weak convergence $\mu^{(D)} \Rightarrow \mu_A$, where the limiting distribution μ_A will be given by a *free product* of distributions $\mathrm{MP}(1) \boxtimes \mu_{\gamma_{A,1}} \boxtimes \mu_{\gamma_{A,2}}$, a notion from the theory of free probability. Here, $\mathrm{MP}(1)$ is the Marchenko-Pastur distribution of parameter 1. The situation is summarized by our first main result, which we state more precisely as Theorem 13.4:

Theorem (Informal)**.** *Consider a family of link states in the bounded spectral variation limit. If the tensor network has a unique minimal cut γ_A for a boundary subsystem A, then $\mu_A^{(D)}$ converges weakly, in probability, to μ_{γ_A}. If there are exactly two minimal cuts $\gamma_{A,1}$ and $\gamma_{A,2}$, then $\mu_A^{(D)}$ converges weakly, in probability, to $\mathrm{MP}(1) \boxtimes \mu_{\gamma_{A,1}} \boxtimes \mu_{\gamma_{A,2}}$.*

We also briefly discuss the closely related problem of computing the entanglement negativity spectrum in the same regime.

In Chapter 14, we investigate a different regime, in which link states are allowed to have *unbounded spectral variation* in the large D limit. This is the more relevant regime for holography, where fluctuations in the area of a surface (in Planck units) grow sublinearly but without bound in the semiclassical limit.

When the spectral variation is unbounded, one has to be careful about how to define the notion of a "minimal cut." A sensible way to formalize this is by using *one-shot entropies*: we say that a cut is minimal if the rank of the state along the cut is smaller than the inverse of the largest element in the entanglement spectrum along any other cut. This condition, while intuitive, is a little too restrictive, and one can use smooth entropies to get a weaker, but still meaningful, condition. In Theorem 14.6, we give a general condition for what it means for a cut to be "minimal," and show that the spectrum of a reduced density matrix ρ_A will be close to the spectrum along the minimal cut.

Finally, one can also consider a further generalization of the random tensor network model, where the link state $\bigotimes_e \phi_e$ is replaced by a general state. This can be used to model holographic states with bulk entropy. We show a version of subregion-subregion duality, and relate this to the quantum information processing task of *split transfer*.

The replica trick in quantum gravity

In this chapter, we give a heuristic description of certain Euclidean gravity path integrals in holography. This section serves as a motivation for the random tensor network models we study, and shows how the holographic computations closely mirror random tensor network computations, but is not needed to understand the random tensor network results. In this chapter we will depart with our convention that logarithms are to base 2 and use natural logarithms (as this is standard in the physics literature).

## 11.1 The replica trick and Euclidean path integrals

How can one compute entropies in quantum field theories? If one would like to compute the von Neumann entropy, one would have to 'compute the logarithm of the density matrix', but it is not clear how to translate this into a field theory computation. The standard way to avoid this is by combining two basic observations. The first observation is that if ρ is a density matrix, we can compute its *Rényi entropies*

$$H_k(\rho) = \frac{1}{1-k} \log \operatorname{tr}\left[\rho^k\right].$$

In principle, this formula is a legitimate definition for all $k \in (0,1) \cup (1,\infty)$, and its limit for $k \to 1$ corresponds to the von Neumann entropy $H(\rho)$. Moreover, for a finite-dimensional Hilbert space $H_k(\rho)$ is analytic in k. If we just know the integer values, and the analytic continuation is unique (for instance because the function does not grow too fast and using Carlson's theorem), then we may use this to deduce $H(\rho)$. The second basic observation is that for integer k,

$$\operatorname{tr}\left[\rho^k\right] = \operatorname{tr}\left[R(\tau)\rho^{\otimes k}\right] \tag{11.1}$$

where $R(\tau)$ is the operator which sends $|i_1\rangle \otimes \ldots \otimes |i_{k-1}\rangle \otimes |i_k\rangle \mapsto |i_k\rangle \otimes |i_1\rangle \ldots \otimes |i_{k-1}\rangle$, that is, it cyclically permutes the k copies of ρ. This is particularly useful because this allows us to compute Rényi entropies using path integrals for quantum field theories. We consider a pure quantum field theory state $|\rho\rangle$ on a space M which is prepared by a Euclidean path integral on $M \times (-\infty, 0]$. Correspondingly, $\langle\rho|$ is prepared by the time-reflected path integral on $M \times [0, \infty)$. If we let $\Phi(x, \tau)$ denote the fields of the theory, S
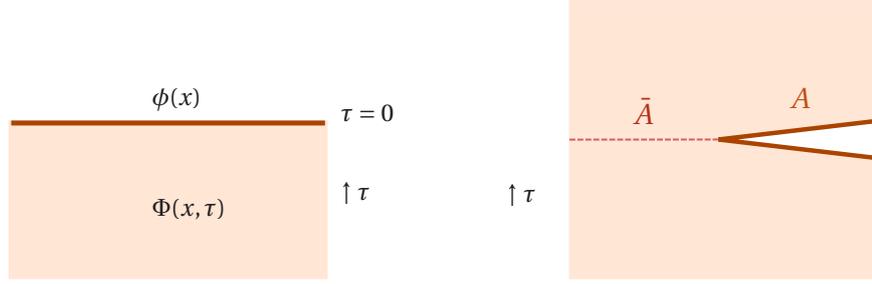
(a) A pure state prepared by a Eu-   (b) Path integral representation of the reduced density
clidean path integral.   matrix on a subsystem A.

Figure 11.1: Quantum field theory state prepared by a path integral.

the Euclidean action, then $|\rho\rangle$ is the state which computes expectation values through the path integral

$$\mathrm{tr}[\rho\phi] = \int_{M\times(-\infty,0]} D\Phi_{\Phi(x,0)=\phi(x)} e^{-S(\Phi)}$$

Let A be some subregion of M. Then the reduced density matrix on A is given by taking $|\rho\rangle$ and $\langle\rho|$, gluing together (and integrating over the fields on $\bar{A}$, the complement of A). We now use the observation in Eq. (11.1), and take k copies of this path integral, and glue the boundaries at the A system cyclically, integrating over the fields. This is illustrated in Fig. 11.2. We conclude that $\mathrm{tr}[\rho_A^k]$ is computed by a path integral $Z_{A,k}$ on a manifold $M_{A,k}$, allowing us to compute $H_k(\rho_A)$. Of course, this is formally infinite. First of all, the path integral is not normalized, so one has to 'normalize' that path integral appropriately by normalizing by Z_1 which can be thought of as $\mathrm{tr}[\rho]$. Then

$$H_k(\rho_A) = \log\frac{Z_{A,k}}{Z_1^k}. \tag{11.2}$$

Moreover, to get a finite result one has to impose a UV cut-off of size ε. There will be divergences in ε, and the 'interesting' part (that is, the part which is independent of the regularization scheme) will be the divergences of order $\log(\varepsilon^{-1})$ and potentially a constant part.

   If the quantum field theory is a CFT, this is especially pleasant. By the operator-state correspondence all states may be prepared as Euclidean path integrals. If the theory is 1+1-dimensional, one may use the conformal symmetry to show that in general (after regularization with an UV cut-off of size ε) the Rényi entropy of an interval A of length $|A|$ is given by

$$H_k(\rho_A) = \left(1+\frac{1}{k}\right)\frac{c}{6}\log\frac{|A|}{\varepsilon}$$

where c is the central charge of the CFT, leading to the *Cardy-Calabrese* formula for the entanglement entropy as $k\to 1$

$$H(\rho_A) = \frac{c}{3}\log\frac{|A|}{\varepsilon}.$$

This result matches the RT formula in $\mathrm{AdS}_3\,/\,\mathrm{CFT}_2$, as this corresponds (upon correct choice of units and regularization) to the length of a geodesic which is ankered to the boundary at the endpoints ∂A of the interval A.
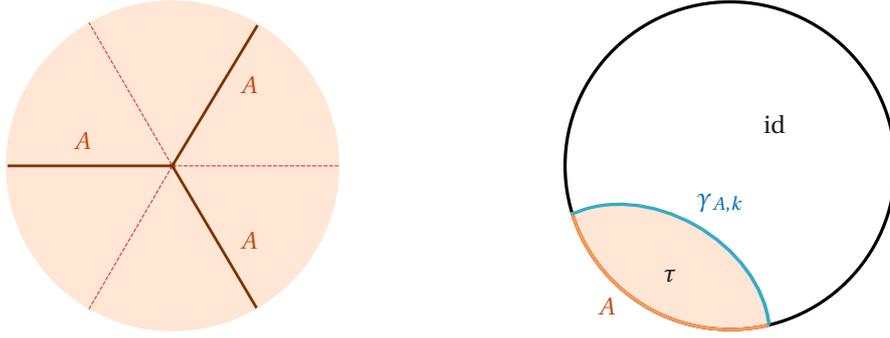
(a) For the replica trick the path integral is glued cyclically along A to obtain $M_{A,k}$. The space $M_{A,k}$ is a k-fold cover of M, branching at the boundary ∂A. In this case $k = 3$.

(b) The holographic version of the replica trick. The boundary manifold is glued as $M_{A,k}$, and the bulk manifold is glued along a cyclic permutation τ adjacent to A, and along the identity permutation along a region adjacent to the complement $\bar{A}$. These two regions are separated by a brane $\gamma_{A,k}$.

Figure 11.2: The path integral replica trick to compute $\text{tr}[\rho_A^k]$.

11.1.1 The holographic replica trick

What happens if the quantum field theory is a *holographic CFT*? Recall that we have a correspondence between the path integral of the CFT on the one hand, and the bulk quantum gravity path integral on the other hand, which we may approximate by its semiclassical saddle. We will use this to sketch a derivation of the RT formula. The original argument is due to [LM13], we follow the closely related approach of [Don16, DHM19]. We consider a state $|\rho\rangle$ which is dual to a semiclassical geometry, and we fix a boundary subsystem A. Then, the k-th Rényi entropy is given by Eq. (11.2), and we use holographic duality to compute the value of these path integrals. For sufficiently large effective central charge we may approximate by a saddle-point approximation, using the semiclassical gravitational action I to obtain

$$Z_{A,k} \approx e^{-I(B_{A,k})}$$

where $B_{A,k}$ is a (Riemannian) manifold which has $M_{A,k}$ as its conformal boundary and is a solution to the gravitational equations of motion. We note that $M_{A,k}$ has a $\mathbb{Z}_k$ symmetry, by shifting the k replicas. Let us assume that the saddle point $B_{A,k}$ also satisfies this *replica symmetry*. Then we may consider the orbifold $\hat{B}_{A,k} = B_{A,k}/\mathbb{Z}_k$, which then satisfies

$$I(B_{A,k}) = kI(\hat{B}_{A,k}).$$

This yields

$$H_k(\rho_A) \approx \frac{k}{k-1}\left(I(\hat{B}_{A,k}) - I(B_1)\right).$$

The bulk solution $\hat{B}_{A,k}$ will have two domains, one where the replicas in $B_{A,k}$ are glued cyclically (adjacent to A) and a complementary region (adjacent to the complement of A) as shown in Fig. 11.2b. Let $\gamma_{A,k}$ be the surface where these two regions meet. We may write the action as $I(\hat{B}_{A,k}) = I_{\text{away}}(\hat{B}_{A,k}) + I_{\text{brane}}(\hat{B}_{A,k})$, where $I_{\text{away}}(\hat{B}_{A,k})$ is the

action away from $\gamma_{A,k}$, and where $I_{\text{brane}}(\hat{B}_{A,k})$ corresponds to a *conical deficit* at $\gamma_{A,k}$ of size $\frac{2\pi(k-1)}{k}$. This conical deficit arises from the orbifold construction, and can be interpreted as a 'cosmic brane' in the action with tension $\frac{k-1}{4G_N k}$. The term $I_{\text{away}}(\hat{B}_{A,k})$ cancels against $I(B_1)$ and we are left with

$$I(\hat{B}_{A,k}) - I(B_1) = \frac{k-1}{4G_N k}|\gamma_{A,k}|$$

where $|\gamma_{A,k}|$ is the area of the brane. This may be analytically continued to non-integer values k, and

$$H_k(\rho_A) \approx \frac{|\gamma_{A,k}|}{4G_N} \tag{11.3}$$

where $\gamma_{A,k}$ is the brane in the saddle-point solution. In particular, continuation to $k = 1$ yields the RT formula

$$H_k(\rho_A) \approx \frac{|\gamma_A|}{4G_N}. \tag{11.4}$$

where one can show that the equations of motion for the brane impose that γ_A is now the *minimal area* surface homologous to A.

What can we say about the spectrum of ρ_A? This spectrum can be recovered from the Rényi entropies. In Eq. (11.3), we see that $|\gamma_{A,k}|$ depends on k but is otherwise fixed as we let G_N go to zero in the classical limit. It was argued in [BPSW19] that this behavior implies that we can approximate the spectrum to be such that for λ in the spectrum, the value of $-\log\lambda$ lies in an interval

$$[H(\rho_A) - \mathcal{O}(G_N^{-\frac{1}{2}}), H(\rho_A) + \mathcal{O}(G_N^{-\frac{1}{2}})],$$

which can be rephrased as the interval

$$[H(\rho_A) - \mathcal{O}(\sqrt{H(\rho_A)}), H(\rho_A) + \mathcal{O}(\sqrt{H(\rho_A)})].$$

This situation can be made more precise in the language of one-shot information theory and smooth entropies, and we will do so in the random tensor network setting in Chapter 14. This shows that the large c limit in a holographic CFT is similar to considering the many-copy limit in quantum information theory, where one also finds that for a fixed state ρ_0, one can approximate the spectrum of $\rho_0^{\otimes n}$ such that for λ in the spectrum, $-\log\lambda$ lies in the interval

$$[nH(\rho_0) - \mathcal{O}(\sqrt{n}), nH(\rho_0) + \mathcal{O}(\sqrt{n})].$$

11.1.2 Fixed area states

A useful variation on the derivation of the RT formula is to consider *fixed-area states* [DHM19]. We consider the 'area operator' $\hat{\gamma}_A$ for a subsystem A, which measures the area of a minimal surface. The operator $\hat{\gamma}_A$ actually has fluctuations, and we may write

$$|\rho\rangle = \int \mathrm{d}\alpha \, |\psi_\alpha\rangle$$

where ψ_α is an eigenvector of $\hat{A}$ with eigenvalue α, so $|\rho\rangle$ is a superposition of fixed area states. The state $|\psi_\alpha\rangle$ is a fixed area state, and can be thought of as prepared by a bulk path integral where we have restricted to bulk geometries for which $|\gamma_A| = \alpha$. The same derivation as above now leads to

$$H_k(\rho_A) \approx \frac{\alpha}{4G_N}.$$

since the area of the minimal surface is fixed to be α. Thus, in this case, all Rényi entropies are (to good approximation) equal which implies that the state has *flat entanglement spectrum* (i.e. all nonzero eigenvalues of ρ_A are approximately equal).

11.2 States at a minimal surface phase transition

Consider now a CFT state ρ on the boundary, which is prepared by a Euclidean path integral. We would like to investigate what happens if the minimal surface for A is not unique, but there are two surfaces of the same order of magnitude, as in Fig. 10.4. In order to study the entanglement spectrum of a reduced state ρ_A, we will again use the replica trick. In the case where there was a unique minimal surface, we saw how the replica trick lead to the derivation of the RT formula. Let us denote the two competing minimal surfaces in the bulk by $\gamma_{A,1}$ and $\gamma_{A,2}$. In [MWW20], it was shown how the entanglement entropy should behave at this phase transition between the two minimal surfaces. A similar computation was performed in [AP20] for the setting with two competing minimal surfaces and bulk matter. We will briefly sketch their argument, referring the interested reader to [AP20] for more details.

11.2.1 Fixed area states with two minimal surfaces

We begin by considering bipartite fixed-area states ρ_{AB}, which are states prepared by a gravitational Euclidean path integral, and in which we have fixed the size of the two competing surfaces $\gamma_{A,1}$ and $\gamma_{A,2}$. In this case, saddle points of the path integral have to satisfy the equations of motion everywhere except at the surfaces $\gamma_{A,i}$, where there could be conical singularities. The two surfaces divide the bulk into 3 regions: a_1, a_2 and a_3.

The saddle points are states with smooth geometries in the regions where the copies of regions a_i are glued to each other – the k copies of region a_1 are glued cyclically, while the copies of region a_3 are glued without permutation. On the middle region a_2, we are free to glue along an arbitrary permutation π. This breaks the replica symmetry.

Let us label such a saddle solution by $B_{A,\pi}$, and let us write ϕ_i for the conical singularity angle at $\gamma_{A,i}$. It turns out that these saddle points lead to an action of the form

$$I(B_{A,\pi}) = kI_{\text{away}}(B_{A,\pi}) + (k\phi_1 - 2\pi|C(\pi)|)\frac{|\gamma_{A,1}|}{8\pi G_N} + (k\phi_2 - 2\pi|C(\tau^{-1}\pi)|)\frac{|\gamma_{A,2}|}{8\pi G_N},$$

where g is the saddle corresponding to a single copy of the state, $I_{\text{away}}[g]$ is the action away from the surfaces, ϕ_i is the angle of the conical singularity at $\gamma_{A,i}$, $|\gamma_{A,i}|$ is the area of the surface $\gamma_{A,i}$, $|C(\pi)|$ is the number of cycles of π, and τ is the full cycle $(12\ldots k)$. In

particular, for $k = 1$ we have

$$I(B_{A,\pi}) = I_{\text{away}}(B_{A,\pi}) + (\phi_2 - 2\pi)\frac{|\gamma_{A,1}|}{8\pi G_N} + (\phi_2 - 2\pi)\frac{|\gamma_{A,2}|}{8\pi G_N},$$

so when we look at the normalized path integral, and sum over all permutations

$$\frac{Z_{A,k}}{(Z_{A,1})^k} \approx \sum_{\pi \in S_k} e^{(|C(\pi)|-k)\frac{|\gamma_{A,1}|}{4G_N} + (|C(\tau^{-1}\pi)|-k)\frac{|\gamma_{A,2}|}{4G_N}}$$

$$= \sum_{\pi \in S_k} e^{-d(\text{id},\pi)\frac{|\gamma_{A,1}|}{4G_N} - d(\pi,\tau)\frac{|\gamma_{A,2}|}{4G_N}}.$$

This is very similar to the situation in Section 10.2.4. As in that case, in this expression not all permutations will be relevant. The areas $\gamma_{A,i}$ are of the same order of magnitude, and are divergent. As a result, only the permutations for which $d(\text{id},\pi) + d(\pi,\tau)$ are minimal will contribute, as all other permutations are suppressed by at least a factor of the area of $\gamma_{A,i}$ in the action. The relevant permutations are again the non-crossing permutations $NC(k)$. We conclude that

$$\frac{Z_{A,k}}{(Z_{A,1})^k} \approx \sum_{\pi \in NC(k)} e^{(|C(\pi)|-k)\frac{|\gamma_{A,1}|}{4G_N} + (|C(\tau^{-1}\pi)|-k)\frac{|\gamma_{A,2}|}{4G_N}}. \tag{11.5}$$

This computation is in one-to-one correspondence with the computation in for a single random tensor we discussed in Section 10.2.4, as also observed in [PSSY19]. It also corresponds more generally to a random tensor network computation with two minimal cuts, as will be clear from the computations in Chapter 12. One can also add bulk matter in this path integral computation, which will again be in correspondence to a similar computation in a random tensor network [AP20]. From the moment computation in Eq. (11.5) and applying the results for the entanglement of a single random tensor, we observe that for two fixed surfaces of exactly equal size, the (appropriately scaled) entanglement spectrum is a Marchenko-Pastur distribution, giving an $\mathcal{O}(1)$ correction to the entanglement entropy, agreeing with the gravitational replica trick computation in [MWW20].

11.2.2 General states at the minimal surface phase transition

We now relax the fixed-area restriction, and study similar calculations performed in [DHM19], [MWW20], and [AP20]. Denote by $Z_{A,k}(\alpha_1, \alpha_2)$ the path integral where we have fixed the areas of $\gamma_{A,i}$ to be α_i. Then, following section 2.3 in [DHM19], the full path integral is given by

$$Z_{A,k} = \int d\alpha_1 d\alpha_2 \, Z_k(\alpha_1, \alpha_2).$$

Again, we consider the semiclassical limit, so we take our saddle-point approximation of $Z_k(\alpha_1, \alpha_2)$ in Eq. (11.5), and we also take a saddle-point approximation for the integral over α_1 and α_2. This saddle point will be at the values for α_i where the deficit angles are given by $\phi_i = \frac{2\pi}{n}$ (since then the saddle point geometry is smooth), which leads to

$$\frac{Z_{A,k}}{(Z_{A,1})^k} \approx \sum_{\pi \in NC(k)} e^{(|C(\pi)|-k)\frac{|\gamma_{A,1}^{(k)}|}{4G_N} + (|C(\tau^{-1}\pi)|-k)\frac{|\gamma_{A,2}^{(k)}|}{4G_N}} \tag{11.6}$$

where $\gamma_{A,i}^{(k)}$ are now minimal surfaces, with a dependence on k. Analytic continuation to $k = 1$ yields the usual surface prescription. In particular, if there are two surfaces that are of almost equal area, the contribution of the larger term is exponentially suppressed for any $\mathcal{O}(1)$ or larger difference in areas.

To zoom in on the region where the two surfaces are nearly equal, we write the state as a superposition of fixed area states. We may discretize the area size α_1 and α_2 over $\text{poly}(\frac{1}{G_N})$ values and approximate the state as a (finite) sum

$$|\psi\rangle_{AB} = \sum_{\alpha_1,\alpha_2} \sqrt{p(\alpha_1,\alpha_2)}\, |\psi_{\alpha_1,\alpha_2}\rangle$$

where $|\psi_{\alpha_1,\alpha_2}\rangle$ is the state where the areas are fixed as $|\gamma_{A,i}| = \alpha_i$ and p is a probability distribution over the possible areas. Then a straightforward calculation of the reduced density matrix ρ_A yields a state of the form:

$$
\begin{aligned}
\rho_A &= \sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)\rho_{A,\alpha_1,\alpha_2} + \sum_{\alpha_1\neq\alpha_1',\alpha_2\neq\alpha_2'} \sqrt{p(\alpha_1,\alpha_2)p(\alpha_1',\alpha_2')}\,\mathrm{tr}_B\!\left[|\psi_{\alpha_1,\alpha_2}\rangle\langle\psi_{\alpha_1',\alpha_2'}|\right], \\
&= \sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)\rho_{A,\alpha_1,\alpha_2} + OD_A,
\end{aligned}
$$

where OD_A are the off-diagonal elements of ρ_A. One can argue that the states $\rho_{A,\alpha_1,\alpha_2}$ are all mutually orthogonal by entanglement wedge reconstruction – the area operator can be reconstructed on A, and hence, each $\rho_{A,\alpha_1,\alpha_2}$ is perfectly distinguishable from each other. Then the entropy of the diagonal part of the state is easily computed as

$$H\!\left(\sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)\rho_{A,\alpha_1,\alpha_2}\right) = \sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)H(\rho_{A,\alpha_1,\alpha_2}) - \sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)\log p(\alpha_1,\alpha_2). \tag{11.7}$$
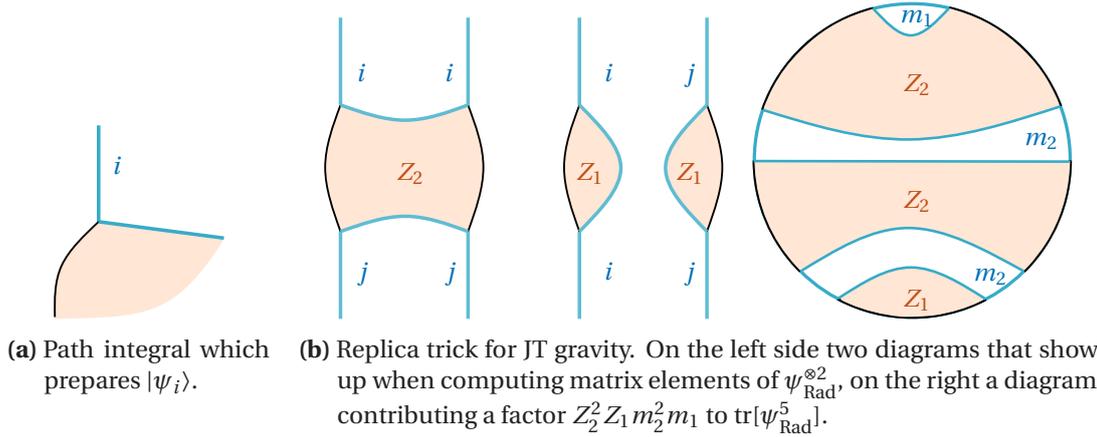
The second term is the so-called entropy of mixing, and it is a standard argument that this term is suppressed relative to the first term [MWW20] as $\mathcal{O}(\ln G_N)$ or smaller. The entropies appearing in the first term can be computed using the methods in the previous subsection.

Returning to the off-diagonal terms OD_A, [MWW20] argued that such terms should be subleading in the analytic continuation due to the relevant surfaces breaking replica symmetry. At the same time, [AP20] argued that such terms should be subleading due to reasons similar to those for the orthogonality of the diagonal elements: complementary entanglement wedge reconstruction implies one may reconstruct the bulk area operator on B, and hence, such states are perfectly distinguishable on B. Therefore, the partial trace over B vanishes for $\alpha_1 \neq \alpha_1'$, $\alpha_2 \neq \alpha_2'$.

At any rate, one reaches the conclusion:

$$H(\rho_A) = \sum_{\alpha_1,\alpha_2} p(\alpha_1,\alpha_2)\frac{\min\{\alpha_1,\alpha_2\}}{4G_N} + \mathcal{O}(\ln G_N), \tag{11.8}$$

In this computation the $\mathcal{O}(1)$ corrections due to the Marchenko-Pastur distribution along each pair of minimal cuts of equal size (or equivalently, the degeneracy in the contributions to the saddle point approximation) is irrelevant, as the entropy of mixing already leads to $\mathcal{O}(\ln G_N)$ deviations.

(a) Path integral which prepares $|\psi_i\rangle$.

(b) Replica trick for JT gravity. On the left side two diagrams that show up when computing matrix elements of $\psi_{\text{Rad}}^{\otimes 2}$, on the right a diagram contributing a factor $Z_2^2 Z_1 m_2^2 m_1$ to $\text{tr}[\psi_{\text{Rad}}^5]$.

Figure 11.3: Path integrals and the replica trick for JT gravity. See [PSSY19] for a detailed explanation of the diagrammatic notation.

11.2.3 Replica wormholes and JT gravity

One of the most basic models of quantum gravity is *JT gravity*, a $1+1$-dimensional model of gravity; see [Sár17] for a review. JT gravity is also a useful model for the near-horizon dynamics of extremal black holes in any dimension. In this case, the dual theory should be $0+1$-dimensional. In other words, it should be regular quantum mechanics rather than a quantum field theory. Indeed, in [SSS19], it was shown that JT gravity theory is dual to a random matrix model, where the Hamiltonian is a random self-adjoint matrix according to some distribution, providing another strong connection between quantum gravity and random matrix theory. It also appears that such gravitational systems may be dual to an ensemble of boundary theories [BW20], rather than a single one. Whether this is fundamental, a special feature of $1+1$-dimensional models, or due to averaging over microscopic features of the gravity theory, is a line of active research [SSSY21].

We now sketch a variation on a calculation in [PSSY19], providing proof-of-principle that the free probability techniques we introduce in Chapter 13 provide an elegant framework to understand such results. We refer the interested reader to [PSSY19] for more in-depth motivation and detailed computations.

We consider JT gravity with an end of the world (EOW) brane containing a large number n of internal states. This model has action

$$I = I_{\text{JT}} + \mu \int_{\text{brane}} \text{d}s,$$

where the action of a manifold M with metric g, induced boundary metric h, extrinsic curvature K, and dilaton ϕ is given by

$$I_{\text{JT}}[M, g] = -\frac{S_0}{2\pi} \left[\frac{1}{2} \int_M \sqrt{g} R + \int_{\partial M} \sqrt{h} K \right] - \left[\frac{1}{2} \int_M \sqrt{g} \phi(R+2) + \int_{\partial M} \sqrt{h} \phi K \right].$$

The details of this action are not very important for us; we just note that we will take the S_0 parameter to be large, and that this suppresses contributions where the manifold M has genus $\gamma > 0$ in the Euclidean path integral.

Such systems are of interest when studying a simple version of an evaporating black

hole. Let

$$|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |\psi_{B,i}\rangle |i_{\mathrm{Rad}}\rangle$$

where $|\psi_{B,i}\rangle$ is the state of the black hole with the EOW brane in state i, and $|i_{\mathrm{Rad}}\rangle$ is a reference state, which can be thought of as the radiation system. Notice that the entanglement spectrum of this state is flat. We generalize this to

$$|\psi\rangle = \sum_{i=1}^{n} \sqrt{p_i} |\psi_{B,i}\rangle |i_{\mathrm{Rad}}\rangle,$$

where the entanglement between the black hole and the radiation has some nontrivial spectrum, which we will assume to be close to uniform, so that $\frac{p_i}{n}$ is bounded by a constant for all i. In Chapter 13, we will formalize this assumption as having *bounded spectral variation.* We let

$$m_k = \sum_{i=1}^{n} n^{k-1} p_i^k$$

be the (appropriately scaled) moments of the spectrum of the entanglement spectrum of the EOW brane. Moreover, we write the path integral on a disc geometry with k boundary components and k EOW branes as $e^{-S_0} Z_k$. Then following the arguments of [PSSY19], one can compute the k-th moment of the radiation system for large n and large e^{S_0} (large n enforces a planar limit with only non-crossing partitions, while large e^{S_0} ensures that only genus $\gamma = 0$ geometries contribute), as illustrated in Fig. 11.3b. The contributions of path integral configurations connecting different replicas are called *replica wormholes.* This diagrammatic computation shows that

$$\mathrm{tr}[\psi_{\mathrm{Rad}}^k] = \sum_{\pi \in NC(k)} m_\pi \frac{Z_{\pi^{-1}\tau}}{Z_1^k} n^{-d(\pi,\mathrm{id})} e^{-S_0 d(\pi,\tau)}. \tag{11.9}$$

In this expression, we use the notation $Z_\sigma = \prod_{l \in C(\sigma)} Z_l$, where $C(\sigma)$ is the cycle type of σ, and $l \in C(\sigma)$ are the lengths of the cycles of σ, and similarly for m_σ. This expression implies that if $n \gg e^{S_0}$, the dominant contribution in Eq. (11.9) has $\pi = \tau$. On the other hand, if $n \ll e^{S_0}$, the dominant contribution in Eq. (11.9) is given by $\pi = \mathrm{id}$. This corresponds to the situation where there is a unique minimal surface (more precisely, a unique *quantum extremal surface*). We are interested in the regime at the phase transition, which is analogous to the Page time of an evaporating black hole, so we assume $n e^{-S_0} \to 1$. The coefficients m_σ correspond to the weight of the σ configuration, as determined by the number and length of the cycles in σ, and the probability distribution of eigenstates p_i. In the case of the flat entanglement spectrum, this number equals the number of closed loops between the connected components. This will also be the case for the non-trivial entanglement spectrum, but each loop will have a different weight that depend on the p_i's.

The m_k are the (scaled) moments of a probability distribution. While the explicit expression itself is not important for our purposes, the Z_l can be written as the l-th moments of a probability distribution [PSSY19]. Hence, Eq. (11.9) is a product of moments, summed over all non-crossing partitions of length k. As a result, we can express $\mathrm{tr}[\psi_{\mathrm{Rad}}^k]$ in the planar limit by way of free probability theory.

More precisely, we may define moment-generating functions for EOW, JT, and Rad:

$$M_{\mathrm{JT}}(z) = \sum_{k=1}^{\infty} \frac{Z_k}{Z_1^k} z^k,$$

$$M_{\mathrm{EOW}}(z) = \sum_{k=1}^{\infty} m_k z^k,$$

$$M_{\mathrm{Rad}} = \sum_{k=1}^{\infty} e^{S_0(k-1)} \mathrm{tr}[\psi_{\mathrm{Rad}}^k] z^k.$$

Given a moment generating function $M(z)$, which is a formal power series, denote by $M^{-1}(z)$ the power series which is its formal inverse and let

$$S(z) = \frac{1+z}{z} M^{-1}(z)$$

be the *S-transform*, and we use this to define the S-transforms S_{JT}, S_{EOW} and S_{Rad}. We will see in Theorem 13.2 that the relation between the moments in Eq. (11.9) implies that these are related as

$$S_{\mathrm{Rad}}(z) = \frac{1}{1+z} S_{\mathrm{JT}}(z) S_{\mathrm{EOW}}(z). \tag{11.10}$$

In Section 13.1 we will explain how this means that the spectrum of ψ_{Rad} can be described using notions from free probability theory.

A recursion relation for the resolvent

Given a moment generating function $M(z)$, we may also define the *resolvent function* $R(z)$ by

$$R(z) = \frac{1}{z}\left(1 + M\left(\frac{1}{z}\right)\right).$$

To relate to previous results, we consider the case where the entanglement with the radiation is maximally entangled. In this case, $S_{\mathrm{EOW}}(z) = 1$ and $S_{\mathrm{Rad}}(z) = \frac{1}{1+z} S_{\mathrm{JT}}(z)$. By definition of the S-transform and setting $z \to M_{\mathrm{Rad}}(z)$, this implies

$$\frac{1}{1 + M_{\mathrm{Rad}}(z)} S_{\mathrm{JT}}(M_{\mathrm{Rad}}(z)) = \frac{1 + M_{\mathrm{Rad}}(z)}{M_{\mathrm{Rad}}(z)} z,$$

which we may rewrite as (again using the definition of the S-transform):

$$M_{\mathrm{Rad}}(z) = M_{\mathrm{JT}}[z(1 + M_{\mathrm{Rad}}(z))].$$

In terms of the resolvent this becomes

$$\begin{aligned}
R(z) &= \frac{1}{z} + \frac{1}{z} M_R\left(\frac{1}{z}\right) \\
&= \frac{1}{z} + \frac{1}{z} M_{\mathrm{JT}}(R(z)) \\
&= \frac{1}{z} + \sum_{k=1}^{\infty} \frac{Z_k}{Z_1^k} \frac{R(z)^k}{z},
\end{aligned}$$

which is a recursion relation previously derived in [PSSY19] by a diagrammatic argument. More generally, we can interpret Eq. (11.10) as a (complicated) recursion relation that directly generalizes the above recursion relation.

Random tensor network states

In this chapter we set the stage for our results and introduce the random tensor network model in detail. We start by introducing conventions and notation. Then in Section 12.1 we introduce random tensor network models with nontrivial link states and discuss the replica trick for such states. In Section 12.2 we give a further generalization, of random tensor networks with general background states.

Notation and conventions

Recall that we denote by $\|a\|_p$ the ℓ_p-norm of a vector a, defined by $\|a\|_p^p = \sum_i |a_i|^p$. If a and b are vectors of different dimension, we extend the shorter vector by zeros and still write $\|p - q\|_p$ for their distance. For example, if $a \in \mathbb{C}^{d_1}$ and $b \in \mathbb{C}^{d_2}$ with $d_2 > d_1$, we write $\|a - b\|_1 = \sum_{i=1}^{d_1} |a_i - b_i| + \sum_{i=d_1+1}^{d_2} |b_i|$. If $\mathcal{H}$ is a Hilbert space, we introduce the notation $\mathscr{P}(\mathcal{H})$ for the set of positive semidefinite operators on $\mathcal{H}$. In this chapter often refer to positive semidefinite operators as 'density operators' or 'states', *without* requiring them to be normalized to unit trace. We write $\mathscr{P}_=(\mathcal{H})$ for the set of $\rho \in \mathscr{P}(\mathcal{H})$ with unit trace, $\text{tr}[\rho] = 1$, and we denote by $\mathscr{P}_\le(\mathcal{H})$ the set of subnormalized states, that is, $\rho \in \mathscr{P}(\mathcal{H})$ with $\text{tr}[\rho] \le 1$. Also, we recall that we use the convention that for a vector $|\phi\rangle$ we denote the corresponding pure state by ϕ, so $\phi = |\phi\rangle\langle\phi|$. Given a positive semidefinite operator ρ, we denote by $\text{spec}(\rho)$ the vector containing its spectrum in non-increasing order, and we write $\text{spec}_+(\rho)$ for the nonzero part of the spectrum. It is a well-known fact that

$$\|\text{spec}(\rho) - \text{spec}(\sigma)\|_1 = \|\text{spec}_+(\rho) - \text{spec}_+(\sigma)\|_1 \le \|\rho - \sigma\|_1 \tag{12.1}$$

(in the second expression we use the convention for the distance of vectors of possibly different dimension introduced above). If A is a quantum system with Hilbert space $\mathcal{H}_A$, we write $\mathscr{P}(A) = \mathscr{P}(\mathcal{H}_A)$, $\mathscr{P}_=(A) = \mathscr{P}_=(\mathcal{H}_A)$, and $\mathscr{P}_\le(A) = \mathscr{P}_\le(\mathcal{H}_A)$, and we use subscripts, e.g. $\rho_A \in \mathscr{P}(A)$, to indicate which system and Hilbert space a quantum state is associated with. Finally, we adopt the standard notation that if μ_n is some sequence of finite measures, we write $\mu_n \Rightarrow \mu$ if μ_n converges weakly (or in distribution) to a finite measure μ, meaning that for any bounded continuous function $f \in C_b(\mathbb{R})$,

$$\int f(x)\mathrm{d}\mu_n(x) \to \int f(x)\mathrm{d}\mu(x). \tag{12.2}$$

If μ_n is a sequence of *random* finite measures on $\mathbb{R}$, we say that the sequence μ_n *converges weakly, in probability,* to a finite measure μ, if, for any bounded continuous function $f \in C_b(\mathbb{R})$, it holds that for every $\varepsilon > 0$

$$\lim_{n \to \infty} P\left(|\int f(x) \mathrm{d}\mu_n(x) - \int f(x) \mathrm{d}\mu(x)| \geq \varepsilon \right) = 0,$$

in other words, Eq. (12.2) converges in probability. In this and later chapter, all logarithms are again to base 2.

12.1 Random tensor network states

We first review the random tensor network model, closely following [HNQ$^+$16, DQW21]. Let $G = (V, E)$ be a connected undirected graph, and let $V = V_\partial \sqcup V_b$ be a partition of the vertices into a set of boundary vertices V_∂ and bulk vertices V_b. If $A \subseteq V_\partial$, we write $\bar{A} = V_\partial \setminus A$. We assign a bond dimension D_e to each edge, and we will consider families of states with increasing bond dimensions; for example, we may take equal bond dimension $D_e = D$ for all edges and let D increase. For each vertex $x \in V$, let $\partial\{x\}$ denote the set of edges $e = (xy) \in E$ connecting x to some $y \in V$. We define Hilbert spaces $\mathcal{H}_{e,x} = \mathbb{C}^{D_e}$ for each $e \in \partial\{x\}$, and we let $\mathcal{H}_x := \bigotimes_{e \in \partial\{x\}} \mathcal{H}_{e,x}$. We call the pair (e, x) a *half-edge*. Moreover, for an edge $e = (xy) \in E$ we write $\mathcal{H}_e = \mathcal{H}_{e,x} \otimes \mathcal{H}_{e,y}$. Let $D_x = \dim(\mathcal{H}_x)$. For a subset $A \subseteq V$, we write $\mathcal{H}_A = \bigotimes_{x \in A} \mathcal{H}_x$, and similarly, for a subset $S \subseteq E$ we write $\mathcal{H}_S = \bigotimes_{e \in S} \mathcal{H}_e$. Similarly, for a set T of half-edges we write $\mathcal{H}_T = \bigotimes_{(e,x) \in T} \mathcal{H}_{e,x}$. At each edge $e = (xy) \in E$ we place a link state $\phi_e \in \mathcal{P}_=(\mathcal{H}_e)$

$$|\phi_e\rangle = \sum_{i=1}^{D_e} \sqrt{\lambda_{e,i}} |ii\rangle \in \mathcal{H}_e = \mathcal{H}_{e,x} \otimes \mathcal{H}_{e,y}. \tag{12.3}$$

Then, $\phi_{e,x} = \phi_{e,y} = \sum_{i=1}^{D} \lambda_{e,i} |i\rangle\langle i|$ is the reduced density matrix of the link state on either of the two subsystems. We refer to the vector $\mathrm{spec}(\phi_{e,x}) = \mathrm{spec}(\phi_{e,y})$, which is ordered in non-increasing fashion, as the *entanglement spectrum* of ϕ_e. Let $\phi \in \mathcal{P}_=(V)$ be the *link state* given by

$$|\phi\rangle = \bigotimes_{e \in E} |\phi_e\rangle. \tag{12.4}$$

At every bulk vertex $x \in V_b$, we place a random vector $|\psi_x\rangle \in \mathcal{H}_x$, where the entries of $|\psi_x\rangle$ are independent standard (circularly-symmetric) complex Gaussian random variables: each entry of the tensor can be written as $\frac{1}{\sqrt{2}}(x + iy)$ where x and y are independent real Gaussian random variables of mean 0 and unit variance. We note that, in the model of [HNQ$^+$16], the tensors $|\psi_x\rangle$ were not chosen as random Gaussian vectors, but as uniformly random vectors on the unit sphere. However, for our choice of Gaussian $|\psi_x\rangle$, the norm $\||\psi_x\rangle\|$ is independent of the normalized vector $|\psi_x\rangle / \||\psi_x\rangle\|$, and $|\psi_x\rangle / \||\psi_x\rangle\|$ will be a uniformly random vectors on the unit sphere. Therefore, these two models only differ by their normalization. We write $|\psi\rangle = \bigotimes_{x \in V_b} |\psi_x\rangle$. The resulting *random tensor network state* $\rho \in \mathcal{P}(V_\partial)$ is defined by

$$|\rho\rangle = (I_{V_\partial} \otimes \langle\psi|) |\phi\rangle. \tag{12.5}$$

The random tensor network state is obtained by projecting the link states onto random vectors, so that the final state lives in the boundary Hilbert space. We can make this manifest by using the cyclicity of the trace to write the density matrix:

$$\rho = \left(I_{V_\partial} \otimes \langle \psi | \right) \phi \left(I_{V_\partial} \otimes | \psi \rangle \right) = \mathrm{tr}_{V_b}\left[\left(I_{V_\partial} \otimes \psi\right)\phi\right]. \tag{12.6}$$

Note that this state need not be normalized, but we chose the standard deviation of the $|\psi_x\rangle$ such that ρ is normalized on average, given that the link state ϕ is normalized:

$$\mathbb{E}\,\mathrm{tr}[\rho] = \mathrm{tr}[\phi]. \tag{12.7}$$

In Section 12.2.1, we prove the stronger statement that ρ is normalized with high probability for appropriately connected tensor networks and large bond dimension. Note also, that in Eq. (12.3), we have chosen states which have a Schmidt decomposition in a fixed basis (the standard basis). Since we project onto uniformly random tensors, we can choose to do so without loss of generality.

12.1.1 The replica trick for random tensor networks

We now consider a boundary subset $A \subseteq V_\partial$ and use the *replica trick* to study the Rényi entropies of the reduced density matrix ρ_A. The replica trick for random tensor network models was first studied in [HNQ$^+$16], and it is the key tool we apply throughout this work. Let $\mathscr{H}$ be a Hilbert space. The Rényi entropies of a (normalized) density matrix $\rho \in \mathscr{P}_=(\mathscr{H})$ are defined by

$$H_k(\rho) = \frac{1}{1-k}\log(\mathrm{tr}[\rho^k])$$

for $k \in (0,1) \cup (1,\infty)$. For $k = 0, 1, \infty$, there are well-defined limits, given by

$$\begin{aligned} H_0(\rho) &:= \log(\mathrm{rank}(\rho)) \\ H_1(\rho) &:= -\mathrm{tr}[\rho\log(\rho)] \\ H_\infty &:= -\log(\|\rho\|_\infty). \end{aligned} \tag{12.8}$$

In particular, we see that $H(\rho) = H_1(\rho)$ is the von Neumann entropy. For reduced density matrices we also write $H(A)_\rho := H(\rho_A)$ and $H_k(A)_\rho := H_k(\rho_A)$. If $\rho \in \mathscr{P}_\le(\mathscr{H})$ is subnormalized, we let

$$H_k(\rho) = \frac{1}{1-k}\log\frac{\mathrm{tr}\left[\rho^k\right]}{\mathrm{tr}[\rho]}. \tag{12.9}$$

Denote by R the representation of S_k on $\mathscr{H}^{\otimes k}$ which permutes the k copies of $\mathscr{H}$ according to the action of S_k. We will write $R_x(\pi)$ when $\mathscr{H} = \mathscr{H}_x$ and $R_A(\pi)$ if $\mathscr{H} = \mathscr{H}_A$ for $A \subseteq V$. We let τ denote the standard k-cycle in S_k, i.e.,

$$\tau = (12\ldots k).$$

The key idea of the replica trick is the observation that the k-th moment of $\rho \in \mathscr{P}(\mathscr{H})$ can be written as

$$\mathrm{tr}\left[\rho^k\right] = \mathrm{tr}\left[R(\tau)\rho^{\otimes k}\right]. \tag{12.10}$$

Recall the notion of the cycle type of a permutation π: if π can be written as as a product of m disjoint cycles of lengths $l_1, \ldots, l_m$, then π has cycle type $C(\pi) = \{l_1, \ldots, l_m\}$. Then, for an arbitrary $\pi \in S_k$,

$$\mathrm{tr}\Big[R(\pi)\rho^{\otimes k}\Big] = \prod_{l \in C(\pi)} \mathrm{tr}\Big[\rho^l\Big].$$

Note that this is the generalization of the well-known *swap trick* for two copies of a state ρ. The other crucial ingredient is a property of the Gaussian random vectors:

$$\mathbb{E}\Big[\psi_x^{\otimes k}\Big] = \sum_{\pi \in S_k} R_x(\pi). \tag{12.11}$$

Using Eq. (12.6), we may then compute

$$
\begin{aligned}
\mathbb{E}\,\mathrm{tr}\Big[\rho_A^k\Big] &= \mathbb{E}\,\mathrm{tr}\Big[R_A(\tau)\rho_A^{\otimes k}\Big] \\
&= \mathbb{E}\,\mathrm{tr}\Big[\big(R_A(\tau) \otimes R_{V_\partial \setminus A}(\mathrm{id})\big)\rho^{\otimes k}\Big] \\
&= \mathbb{E}\,\mathrm{tr}\Big[\big(R_A(\tau) \otimes R_{V_\partial \setminus A}(\mathrm{id})\big)\big(I_{V_\partial} \otimes \psi\big)^{\otimes k}\phi^{\otimes k}\Big] \\
&= \mathrm{tr}\Big[(R_A(\tau) \otimes R_{V \setminus A}(\mathrm{id}))\,\mathbb{E}\Big[\big(I_{V_\partial} \otimes \psi\big)^{\otimes k}\Big]\phi^{\otimes k}\Big].
\end{aligned}
\tag{12.12}
$$

To further simplify this expression, we define the following set:

$$\mathcal{S}_{A,\sigma} = \big\{\{\pi_x\}_{x \in V} : \pi_x \in S_k, \text{ where } \pi_x = \sigma \text{ for } x \in A \text{ and } \pi_x = \mathrm{id} \text{ for } x \in \bar{A}\big\}, \tag{12.13}$$

for any $\sigma \in S_k$ and $A \subseteq V_\partial$. An element of $\mathcal{S}_{A,\sigma}$ assigns a permutation to each vertex in V subject to a 'boundary condition.' Now, using Eq. (12.11), we find that

$$\mathbb{E}\,\mathrm{tr}\Big[\rho_A^k\Big] = \sum_{\{\pi_x\} \in \mathcal{S}_{A,\tau}} \mathrm{tr}\Bigg[\bigotimes_{x \in V} R_x(\pi_x)\phi^{\otimes k}\Bigg].$$

Finally, we observe that for $e = (xy)$

$$\mathrm{tr}\Big[R(\pi_x) \otimes R(\pi_y)\phi^{\otimes k}\Big] = \prod_{l \in C(\pi_x^{-1}\pi_y)} \mathrm{tr}\Big[\phi_{e,x}^l\Big],$$

where we recall that $\phi_{e,x}$ is the reduced density matrix of the link state on edge $e = (xy)$. Thus, we conclude that

$$\mathbb{E}\,\mathrm{tr}\Big[\rho_A^k\Big] = \sum_{\{\pi_x\} \in \mathcal{S}_{A,\tau}} \prod_{e=(xy) \in E} \prod_{l \in C(\pi_x^{-1}\pi_y)} \mathrm{tr}\Big[\phi_{e,x}^l\Big]. \tag{12.14}$$

We can interpret the expectation as the partition function of a classical spin model

$$\mathbb{E}\,\mathrm{tr}\Big[\rho_A^k\Big] = \sum_{\{\pi_x\} \in \mathcal{S}_{A,\tau}} 2^{-\sum_{e=(xy) \in E} J_e(\pi_x, \pi_y)}, \tag{12.15}$$

where the site variables in the spin model are permutations $\pi_x \in S_k$, the interaction at the edges between sites is given by

$$J_e(\pi_x, \pi_y) = -\sum_{l \in C(\pi_x^{-1}\pi_y)} \log(\mathrm{tr}\Big[\phi_{e,x}^l\Big]) = \sum_{l \in C(\pi_x^{-1}\pi_y)} (l-1)H_l(\phi_{e,x}),$$

with H_l the l-th Rényi entropy, and the model as boundary conditions such that the permutation must be τ on A and id on $\bar{A}$. Similarly, we may place an arbitrary permutation π on A instead of τ, which yields (by exactly the same reasoning)

$$\mathbb{E}\,\mathrm{tr}\Big[R(\pi)\rho^{\otimes k}\Big] = \sum_{\{\pi_x\} \in \mathcal{S}_{A,\pi}} \prod_{e=(xy) \in E} \prod_{l \in C(\pi_x^{-1}\pi_y)} \mathrm{tr}\Big[\phi_{e,x}^l\Big]. \tag{12.16}$$

12.1.2 Maximally entangled link states and minimal cuts

We will now discuss the special case where all the link states are maximally entangled states of dimension D, which has been studied extensively in [HNQ$^+$16]. We will generalize the results we discuss here to a wider class of link states in Chapter 13. In this case, the entanglement spectra of the link states are flat: for $e \in E$, we have $\lambda_{e,i} = \frac{1}{D}$ for $i = 1, \dots, D$. In particular, for all $l \in \mathbb{N}$ we have $H_l(\phi_e) = \log(D)$ and hence

$$J_e(\pi_x, \pi_y) = \log(D) \sum_{l \in C(\pi_x^{-1}\pi_y)} (l-1).$$

This leads to the so-called *Cayley distance* on S_k:

$$d(\pi_x, \pi_y) = \sum_{l \in C(\pi_x^{-1}\pi_y)} (l-1) = k - |C(\pi_x^{-1}\pi_y)|,$$

where $|C(\pi)|$ is the number of cycles in π. Moreover, $d(\pi_x, \pi_y)$ is a metric and equals the minimal number of transpositions needed to transform π_x into π_y. We say that $\pi \in S_k$ is on a geodesic between π_1 and π_2 if $d(\pi_1, \pi) + d(\pi, \pi_2) = d(\pi_1, \pi_2)$ (recall that d is a metric). We can rewrite the spin model in terms of this distance:

$$\mathbb{E}\,\mathrm{tr}\left[\rho_A^k\right] = \sum_{\{\pi_x\} \in \mathscr{S}_{A,\sigma}} 2^{-\log(D)\sum_{e=(xy)\in E} d(\pi_x, \pi_y)}. \tag{12.17}$$

The physically inclined reader may observe that the logarithm of the bond dimension has the role of an inverse temperature, and for large D, the dominant contribution to the partition function will be the ground state of the spin model, subject to the relevant boundary conditions.

To describe the dominant contribution to the sum in Eq. (12.17) for large D, we need the *minimal cuts* for A in G. A cut for A is a subset of the vertices $\Gamma_A \subset V$ such that $\Gamma_A \cap V_\partial = A$. Throughout this work, we will denote the set of all cuts for A by $C(A)$. We will use the convention of denoting cuts (i.e. subsets of vertices) by capital Greek letters. Given a cut $\Gamma_A \in C(A)$, we will denote the set of edges crossing the cut, that is, edges connecting a vertex in Γ_A with a vertex in $V \setminus \Gamma_A$, by lowercase Greek letters γ_A (and by an abuse of language, also refer to this set as a 'cut'). A minimal cut for A is a cut such that the number of edges $|\gamma_A|$ is minimal. We write $m(A) = |\gamma_A|$ for a minimal cut $\gamma_A \in C(A)$. If $\Gamma_A \in C(A)$, we write $\Gamma_A^c = V \setminus \Gamma_A$. Note that Γ_A^c is a cut for $\bar{A} = V_\partial \setminus A$.

In the simplest case, there is a *unique* minimal cut γ_A. For this case, one can show that the dominant configuration is the one in which $\pi_x = \tau$ for $x \in \Gamma_A$ and $\pi_x = \mathrm{id}$ for $x \in V \setminus \Gamma_A$, see [HLW06], or Proposition 13.3. That is, there are two domains in the spin model corresponding to τ and id, and the minimization of the domain wall corresponds to the minimal cut in the graph.

We will also be interested in the case of exactly two non-intersecting minimal cuts $\Gamma_{A,1}$ and $\Gamma_{A,2}$. In this case, we have that $\Gamma_{A,1} \subset \Gamma_{A,2}$, or $\Gamma_{A,1} \subset \Gamma_{A,2}$. After relabeling, we may assume that the first is the case, and define the following three domains in the graph: $V = V_1 \sqcup V_2 \sqcup V_3$ given by $V_1 = \Gamma_{A,1}$, $V_2 = \Gamma_{A,2} \setminus \Gamma_{A,1}$ and $V \setminus \Gamma_{A,2}$. If there are exactly two minimal cuts, then multiple dominant configurations contribute equally to the partition function Eq. (12.17). These dominant configurations can be constructed as follows: for each π on a geodesic between τ and id, set $\pi_x = \tau$ for $x \in V_1$, $\pi_x = \pi$ for $x \in V_2$ and $\pi_x = \mathrm{id}$ for $x \in V_3$. That these are the dominant configurations follows

immediately from the fact that $d(\tau,\pi) + d(\pi,\mathrm{id}) \geq d(\tau,\mathrm{id})$, with equality if and only if π is on a geodesic between τ and id.

To understand this degeneracy, we use the following fact [NS06]: *the set of permutations π on a geodesic between τ and* id *is in a one-to-one correspondence with the set of non-crossing partitions $NC(k)$ of* $[k]$. See Section 13.1 for a definition and properties of the non-crossing partitions. Thus, the degeneracy for the the k-th moment is $|NC(k)| = C_k$ where

$$C_k = \frac{1}{k+1}\binom{k}{2k}$$

is the k-th Catalan number. These are the moments of the *Marchenko-Pastur distribution* $\mathrm{MP}(t)$

$$
\begin{aligned}
MP(t) &= \max(1-t,0)\delta_0 + \nu_t \\
\mathrm{d}\nu_t(x) &= \frac{\sqrt{4t-(x-1-t)^2}}{2\pi x}\mathbb{1}_{(x-1-t)^2 \leq 4t}\mathrm{d}x.
\end{aligned}
\tag{12.18}
$$

This allows one to show the folklore result (which we prove and extend to more general link states in Theorem 13.4) that upon an appropriate rescaling, the empirical distribution of the spectrum of ρ_A converges to a Marchenko-Pastur distribution. This is in line with the case of a single random tensor, which precisely yields a Wishart matrix (see Section 13.1.1 for a brief introduction to these objects). In the first case, where there is a unique minimal cut, the entanglement spectrum of ρ_A is flat, while, as we have seen, in the second case, the degeneracy gives rise to a nontrivial spectrum in the right scaling limit.
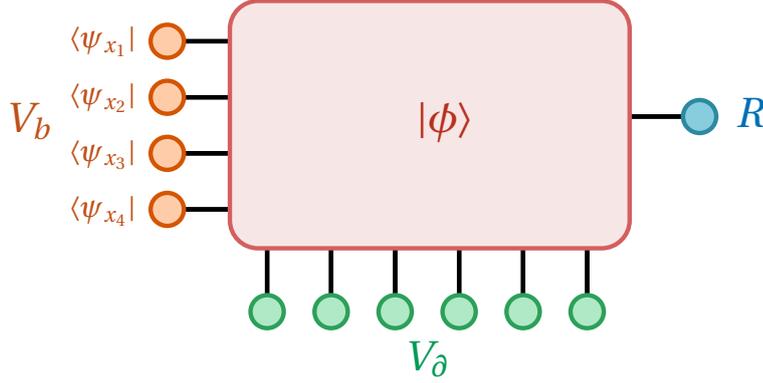
12.2 Random tensor networks with general background states

In Eq. (12.15), we computed the result of the replica trick for the k-th moment for a random tensor network state. We will also consider the more general setting where the link state is replaced by some arbitrary state ϕ_V. In this case there need not be a graph structure, and the Hilbert space at each vertex $x \in V$ can be some arbitrary Hilbert space, rather than a tensor product of Hilbert spaces labelled by half-edges. In this case, we will refer to ϕ_V as a "background state" instead of a "link state" (as the interpretation of links along the edges does not necessarily make sense in this situation). That is, where before we had a link state

$$|\phi\rangle = \bigotimes_{e\in E}|\phi_e\rangle,$$

we will now consider some *arbitrary* possibly mixed and subnormalized $\phi_V \in \mathscr{P}_{\leq}(V)$ in the tensor network construction. We can generalize Eq. (12.6) to also apply for general background states to obtain a state $\rho \in \mathscr{P}(V_\partial)$ given by

$$\rho = \mathrm{tr}_{V_b}\big[\big(I_{V_\partial}\otimes\psi\big)\phi_V\big] \tag{12.19}$$

Figure 12.1: The structure of a (purified) random tensor network with a general background state.

where $|\psi\rangle$ is a tensor product of random states at the bulk vertices. If ϕ is pure, then so is ρ. If ϕ is not pure, we can consider a purification $\phi_{VR} \in \mathscr{P}_{\leq}(VR)$ and consider R as an additional boundary system; this leads to a random tensor network state $\rho_{V_\partial R}$ which is a purification of ρ_{V_∂}. This set-up is illustrated in Fig. 12.1. While formally very similar, the resulting state is no longer a PEPS tensor network state in general.

There are multiple reasons to also allow general background states. The first reason is of a technical nature: they are useful for estimates based on *smooth entropies*, which we discuss in Chapter 14. In this application, the link state is still pure, but no longer a tensor product along the edges. A second motivation for considering general background states is that they can be used as a toy model for holographic systems where there is "bulk entropy" present. Finally, these states are closely related to protocols for the quantum information processing task of split transfer [DH10]. We comment on this connection in Chapter 14.

Even for a general background state, a version of the replica trick still applies. Consider a boundary subsystem $A \subseteq V_\partial$ with corresponding boundary state ρ_A. Then, the computation in Eq. (12.12) is still valid, and we find

$$\mathbb{E}\,\mathrm{tr}[\rho_A^k] = \sum_{\{\pi_x\} \in \mathscr{S}_{A,\tau}} \mathrm{tr}_V \left[\bigotimes_{x \in V} R_x(\pi_x) \phi_V^{\otimes k} \right] \tag{12.20}$$

where $\tau = (12\ldots k)$. However, Eq. (12.20) no longer has the interpretation of a spin model with local interactions.

For general background states, we will only need the replica trick for $k = 2$. Since S_2 has only two elements, each configuration of permutations is completely characterized by the domain $\Delta_A = \{x \in V \text{ such that } \pi_x = \tau\}$. Because of the boundary conditions in $\mathscr{S}_{A,\tau}$, the collection of these sets coincides with $C(A)$, and hence

$$\mathbb{E}\,\mathrm{tr}[\rho_A^2] = \sum_{\Delta_A \in C(A)} \mathrm{tr}[\phi_{\Delta_A}^2] = \sum_{\Delta_A \in C(A)} \mathrm{tr}[\phi] 2^{-H_2(\Delta_A)_\phi}. \tag{12.21}$$

Another useful fact is that by Eq. (12.11),

$$\mathbb{E}\rho_{V_\partial} = \phi_{V_\partial}. \tag{12.22}$$

We remark that if one only uses the $k = 2$ replica trick, one could also use tensors which are drawn from a *projective 2-design*, a distribution which produces tensors

with the same first and second moments as uniformly random tensors of unit norm [KR05, GAE07]. An example of a projective 2-design is the set of uniformly random stabilizer states. For tensors $|\psi_x\rangle$ drawn from a projective 2-design of dimension D_x, it holds that

$$\mathbb{E}\psi_x^{\otimes 2} = \frac{1}{D_x(D_x - 1)}I + \frac{1}{D_x(D_x - 1)}R_x(\tau),$$

and hence

$$\mathbb{E}\operatorname{tr}[\rho_A^2] = \frac{D_x}{D_x + 1}\sum_{\Delta_A \in C(A)}\operatorname{tr}[\phi_{\Delta_A}^2],$$

which is close to Eq. (12.21) for large D_x. Thus, it is not hard to see that all random tensor network results which only use the $k = 2$ replica trick are also valid for states with tensors drawn from projective 2-designs. This was already observed in [HNQ⁺16] and random tensor networks with random stabilizer tensors were further studied in [NW20]. The results of Chapter 14 only use the $k = 2$ replica trick, and thus will extend to states with tensors drawn from projective 2-designs. This will not be true for the results in Chapter 13, which requires usage of the replica trick for all $k \in \mathbb{N}$.

12.2.1   Normalization of random tensor network states

One immediate consequence of the replica trick for $k = 2$ is that the random tensor network state ρ will be approximately normalized with high probability, so long as a mild condition on the background state is satisfied: the bulk needs to be connected, with sufficiently entangled edges. Let

$$\eta = \max_{\Delta \subseteq V_b, \Delta \neq \emptyset}\operatorname{tr}[\phi_\Delta^2] = \max_{\Delta \subseteq V_b, \Delta \neq \emptyset}\operatorname{tr}[\phi]2^{-H_2(\Delta)_\phi}. \tag{12.23}$$

If the the state has enough correlations along each cut (or more precisely, if $H_2(\Delta)_\phi$ is large for each Δ), then η is small. Concretely, if we consider a random tensor network state with maximally entangled link states of bond dimension D, we will have $\eta \leq \frac{1}{D}$. We then have

Lemma 12.1. *For any background state $\phi \in \mathscr{P}_{\leq}(V)$, with associated $\rho \in \P(V_\partial)$ as in Eq. (12.19), it holds that for any $\varepsilon > 0$*

$$P\big(|\operatorname{tr}[\rho] - \operatorname{tr}[\phi]| \geq \varepsilon\big) \leq 2^{|V_b|}\frac{\eta}{\varepsilon^2}$$

where η is defined in Eq. (12.23).

Proof. This follows from a special case of Eq. (12.21). In this case, the empty cut contributes $\operatorname{tr}[\phi]^2$, so we find

$$\operatorname{Var}(\operatorname{tr}[\rho]) = \mathbb{E}|\operatorname{tr}[\rho] - \operatorname{tr}[\phi]|^2 = \mathbb{E}|\operatorname{tr}[\rho]^2 - \operatorname{tr}[\phi]^2| \leq 2^{V_b}\max_{\Delta \subseteq V_b, \Delta \neq \emptyset}\operatorname{tr}[\phi_\Delta^2],$$

where we have used the normalization of ρ in expectation $\mathbb{E}\operatorname{tr}[\rho] = \operatorname{tr}[\phi]$, as in Eq. (12.7). The result follows by an application of Chebyshev's inequality. ∎

We can improve this result by taking advantage of the fact that our random projectors are random Gaussian vectors, allowing us to use Gaussian concentration of measure rather than the Chebyshev's inequality. For instance, using a concentration bound for Gaussian polynomials (see for instance [AS17], Corollary 5.49) one can show that for any $\varepsilon \geq (\sqrt{2}e)^{2V_b}\eta$:

$$P\big(|\mathrm{tr}[\rho] - \mathrm{tr}[\phi]| \geq \varepsilon\big) \leq \exp\left(-\frac{|V_b|}{2e}\varepsilon^{\frac{1}{|V_b|}}\eta^{-\frac{1}{|V_b|}}\right),$$

where η is defined as in Eq. (12.23). We will not need this refinement.

Link states with bounded spectral variation

In this chapter, we study random tensor network states with link states that have *bounded spectral variation*, meaning that there is an effective bond dimension D such that the Schmidt coefficients of the link state are of the order $\frac{1}{D}$.

We start by providing background material on random matrix theory and *free probability*, which is a key tool in the study of products of random matrices. In Section 13.2, we will precisely define the notion of bounded spectral variation and generalize the results in Section 12.1.2 for random tensor network states with maximally entangled link states to this wider class of link states. This leads to the main result of this chapter, Theorem 13.4, which shows that the asymptotic entanglement spectrum can be expressed in terms of a free product of distributions. We will see that the results are similar to the quantum gravity set-up described in Section 11.2.3. Finally, in Section 13.3, we investigate the entanglement negativity for random tensor network states with link states of bounded spectral variation.

13.1   Random matrices and free probability

13.1.1   Random matrix theory and Wishart matrices

We start by reviewing relevant concepts from probability and random matrix theory that are relevant for our analyses. This material can be found in any introduction to random matrix theory, e.g. [AGZ10, BS10, PB20].

A fundamental question in random matrix theory is as follows: given a family of $n \times n$ matrices with entries selected according to some distribution, what is the asymptotic distribution of the eigenvalues as $n \to \infty$? This question has been extensively studied, and in many cases has an elegant and concise answer. We discuss a basic example which is closely related to our purposes: *Wishart matrices*. Consider a $n \times m$ matrix X whose entries are drawn i.i.d. from a Gaussian distribution with mean zero and unit variance. The sample covariance matrix of X is the $n \times n$ matrix defined as

$$Y_{n,m} = \frac{1}{m} X X^T. \tag{13.1}$$

Such random matrices are called *(real) Wishart matrices*, and can be thought of as a sample second moment matrix (where one has m realizations of an n-dimensional

random variable). One can also consider *complex* Wishart matrices: in this case the entries of the $n \times m$ matrix X are complex i.i.d. standard (circularly symmetric) complex Gaussian random variables. We then let $Y_{n,m} = \frac{1}{m} X X^\dagger$. We would like to understand the spectrum of $Y_{n,m}$, and to that end, we consider the empirical distribution of the eigenvalues. This empirical distribution is itself random, depending on the particular realization of $Y_{n,m}$. To characterize the convergence, we recall that if $\{\mu_n\}_{n \in \mathbb{N}}$ is a sequence of random finite measures on $\mathbb{R}$, we say that the sequence μ_n *converges weakly, in probability,* to a finite measure μ, if, for any bounded continuous function $f \in C_b(\mathbb{R})$, it holds that for every $\varepsilon > 0$

$$\lim_{n \to \infty} P\left(|\int f(x) \mathrm{d}\mu_n(x) - \int f(x) \mathrm{d}\mu(x)| \geq \varepsilon \right) = 0.$$

The asymptotic distribution of the eigenvalues of Wishart matrices is known to obey the *Marchenko-Pastur law* (see, for instance, Theorem 3.6 and Theorem 3.7 in [BS10]):

Theorem 13.1. *Consider (real or complex) Wishart matrices $Y_{n,m}$ and let*

$$\mu_{n,m} = \frac{1}{n} \sum_{\lambda \in \mathrm{spec}(Y_{n,m})} \delta_\lambda$$

be the empirical distribution of its eigenvalue spectrum. Suppose that the ratio of dimensions n/m converges to a constant $t > 0$ as $n \to \infty$. Then $\mu_{n,m}$ converges weakly, in probability, to the Marchenko-Pastur distribution $\mathrm{MP}(t)$ *with parameter $t > 0$, as defined in Eq.* (12.18).

Generalizations to this result are possible. For example, one still has convergence if the entries of X are chosen according to non-Gaussian distributions with mean zero and unit variance. Also, one can prove weak convergence, almost surely (rather than just in probability); see [BS10].

If $Y_{n,m} = \frac{1}{m} X X^\dagger$ is a complex Wishart matrix, X can also be interpreted as as a uniformly random pure quantum state on $\mathbb{C}^n \otimes \mathbb{C}^m$, and $Y_{n,m}$, up to normalization, as the reduced density matrix on $\mathbb{C}^n$ [HLW06]. Note that $\frac{1}{n} Y_{n,m}$ is normalized in expectation in the sense that $\mathbb{E} \frac{1}{n} Y_{n,m} = 1$. So, complex Wishart matrices can be used as a model for the reduced state of a random bipartite quantum state, and this allows one to quantify the 'typical entanglement' of a random state. Equivalently, in the tensor network setting, the matrix $\frac{1}{\sqrt{n}} X$ can be thought of as a random tensor network state with a single bulk vertex, two boundary vertices, and maximally entangled link states. We can then can interpret $\frac{1}{n} Y_{n,m}$ as the reduced density matrix on one of the boundary vertices. We will provide a generalization of Theorem 13.1 for the entanglement spectrum of random tensor network states in Theorem 13.4.

13.1.2 Free probability

The topic of probability distributions in random matrix theory is closely related to *free probability* and, in particular, to the notion of the *free product*. We provide a brief introduction here; the material in this section is very standard, and we only review a few relevant aspects. For an extensive treatment, see, for instance, Chapter 5 in [AGZ10] or the books [NS06, MS17, PB20]. As we will see later, the free product will allow us to concisely formulate replica trick results involving multiple minimal cuts.

A *non-commutative probability space* is a pair $(\mathscr{A}, \omega)$, where $\mathscr{A}$ is a C^*-algebra and ω is a state on $\mathscr{A}$. An element $a \in \mathscr{A}$ is called a non-commutative random variable. The key example to have in mind is the space of $n \times n$ random matrices, where the matrix entries are distributed according to some probability distribution, and $\omega(a) = \mathbb{E}\frac{1}{n}\text{tr}[a]$ defines a tracial state. If $a \in \mathscr{A}$, the *distribution* (or *law*) μ_a of a is defined as a map on polynomials, which evaluates on a polynomial p as $\mu_a(p) = \omega(p(a))$. If a is self-adjoint, it has real spectrum and we can extend the domain of μ_a to all bounded continuous functions $f \in C_b(\mathbb{R})$, by using the functional calculus to define $f(a)$ and by letting $\mu_a(f) = \omega(f(a))$. In this case we can identify μ_a with a distribution such that, for $f \in C_b(\mathbb{R})$, we have $\mu_a(f) = \int f(x)\mathrm{d}\mu_a(x)$. In particular, if a is an $n \times n$ self-adjoint random matrix, then $\mu_a(f) = \frac{1}{n}\mathbb{E}\sum_{\lambda \in \text{spec}(a)} f(\lambda)$, and we may identify μ_a with the empirical measure of the eigenvalues of a. If $\mathscr{A}$ is a commutative algebra, these notions reduce to the usual notions of probability theory, where ω is the expectation.

We call a set of n non-commutative random variables $\{a_i\}$ on a non-commutative probability space $(\mathscr{A}, \omega)$ *freely independent* or just *free* if, for any set of $k \geq 2$ polynomials $\{p_j\}$, the variables satisfy

$$\omega(p_1(a_{i_1}) \dots p_k(a_{i_k})) = 0$$

whenever $\omega(p_m(a_{i_m})) = 0$ for all $1 \leq m \leq k$ and no two adjacent indices i_m and i_{m+1} for $1 \leq m \leq k-1$ are equal. One can see that two freely independent variables a_1, a_2 satisfy:

$$0 = \omega((a_1 - \omega(a_1))(a_2 - \omega(a_2))) = \omega(a_1 a_2) - \omega(a_1)\omega(a_2), \tag{13.2}$$

which, in the commutative case with random variables x_1, x_2, is the classical bivariate independence condition $\mathbb{E}[x_1 x_2] = \mathbb{E}[x_1]\mathbb{E}[x_2]$. While the definition of free independence is stronger than independence (commuting independent random variables are generically not free), the role of free independence is analogous to the role of classical independence for commuting random variables: it allows one to, in principle, compute the joint mixed moments of the variables.

We will be interested in the *multiplicative free convolution* or *free product* (there also exists an additive convolution or just *free convolution*) of distributions. Suppose a, b are non-commutative self-adjoint free random variables on $(\mathscr{A}, \omega)$ with distributions μ_a and μ_b. Then we denote the distribution of ab by $\mu_{ab} = \mu_a \boxtimes \mu_b$. Note that, generally, ab need not be self-adjoint. However, if ω is tracial (as in the random matrix case) and a is positive, the distribution of ab coincides with that of $\sqrt{a}b\sqrt{a}$ which is self-adjoint, and we can identify μ_{ab} with a distribution on $\mathbb{R}$. If μ_a and μ_b are compactly supported distributions, then so is $\mu_a \boxtimes \mu_b$.

As a concrete example of the freeness and the free product, let X_n and Y_n be two families of random $n \times n$ positive diagonal matrices with uniformly bounded norm, such that their spectrum converges weakly to probability distributions μ and ν respectively. Let U_n be a family of Haar random unitary $n \times n$ matrices. Then as n goes to infinity, X_n and $Y_n' = U_n Y_n U_n^\dagger$ will be freely independent (so they are *asymptotically free*), and we would like to study their product. The product of positive matrices need not be self-adjoint, so we consider $Z_n = \sqrt{X_n} Y_n' \sqrt{X_n}$ which is a positive matrix. One may then show that the distribution of the spectrum of Z_n weakly converges in probability to $\mu \boxtimes \nu$. See Corollary 5.4.11 in [AGZ10] for a precise statement and proof.

The free product may be analyzed using generating functions. If we are given a

(non-commutative) random variable a with distribution μ_a, let

$$m_{a,k} = \int x^k \mathrm{d}\mu_a(x)$$

be the k-th moment of μ_a. Then the *moment generating function* is the formal power series

$$M_{\mu_a}(z) = \sum_{k=1}^{\infty} m_{a,k} z^k. \tag{13.3}$$

We define the *S-transform* to be the formal power series

$$S_{\mu_a}(z) = \frac{1+z}{z} M_a^{-1}(z),$$

where $M_a^{-1}(z)$ is the power series corresponding to the formal inverse of $M_\mu(z)$ under composition, which is well-defined as long as $m_{a,1} \neq 0$. For compactly supported distributions, the moment generating function, and hence the S-transform, uniquely determines the distribution.

If a and b are non-commutative self-adjoint free random variables, then

$$S_{\mu_a \boxtimes \mu_b}(z) = S_{\mu_{ab}}(z) = S_{\mu_a}(z) S_{\mu_b}(z). \tag{13.4}$$

This also provides a completely combinatorial interpretation of the free product, without reference to the associated non-commutative probability spaces. That is, given compactly supported distributions μ and ν, we can define $\mu \boxtimes \nu$ by Eq. (13.4): it is the compactly supported distribution with moments prescribed by $S_{\mu_a \boxtimes \mu_b}(z)$, which determines $M_{\mu_a \boxtimes \mu_b}(z)$. The free product is commutative and associative.

Example. As an example, we compute the S-transform of the Marchenko-Pastur distribution $\mu \sim MP(1)$. The distribution function is given by

$$\mathrm{d}\mu(x) = \frac{1}{2\pi} \sqrt{4x^{-1} - 1} \mathrm{d}x.$$

The moments can be computed directly:

$$m_k = \sum_{i=0}^{k-1} \frac{1}{i+1} \binom{k}{i} \binom{k-1}{i} \tag{13.5}$$

After some work, one can show that the moments above lead to a closed-form moment generating function

$$M(z) = \frac{2z - 1 - \sqrt{1-4z}}{2z}.$$

One may then invert the expression and obtain the S-transform

$$S(z) = \frac{1}{1+z}.$$

Similarly, for the Marchenko-Pastur distribution $MP(t)$ with parameter t, which has distribution as given in Eq. (12.18), we find that

$$S(z) = \frac{1}{t+z}.$$

See, for instance, [BBCC11].

13.1.3 Non-crossing partitions

Given the set $k \in \mathbb{N}$, let $NC(k)$ denote the set of *non-crossing partitions* of $[k]$. A non-crossing partition of $[k]$ is a partition $[k] = X_1 \sqcup \ldots \sqcup X_m$ which is such that, if for some α we have $i < j \in X_\alpha$, then there are no $k, l \in X_\beta$ for $\beta \neq \alpha$ with $k < i < l < j$ or $i < k < j < l$. To any non-crossing partition, we associate a permutation $\pi \in S_k$ by mapping each subset $\{i_1, \ldots, i_l\}$ to the cycle $(i_1, \ldots, i_l)$ with $i_1 < \ldots < i_l$. In a slight abuse of notation, we will write $\pi \in NC(k)$. For any $\pi \in S_k$, and for a sequence of numbers f_k for $k = 1, 2, \ldots$, we write

$$f_\pi = \prod_{l \in C(\pi)} f_l \tag{13.6}$$

where $C(\pi)$ is the cycle type of π. We will need the following result, which is a straightforward consequence of the combinatorics of the S-transform.

Theorem 13.2. *Consider compactly supported probability distributions μ, ν, ρ. Suppose that the moments of ρ are given by*

$$m_k^\rho = \sum_{\pi \in NC(k)} m_\pi^\mu m_{\pi^{-1}\tau_k}^\nu$$

where $\tau_k = (12 \ldots k)$ is the full cycle. Then

$$\rho = \mathrm{MP}(1) \boxtimes \mu \boxtimes \nu.$$

Proof. We let $\mathscr{F}$ be the transformation that sends a formal power series $f(z)$ to the power series $\frac{1}{z} f^{-1}(z)$. This is such that for some distribution μ, the S-transform is given by $S_\mu(z) = (1 + z)\mathscr{F}(M_\mu)(z)$. Moreover, given two power series $f(z) = \sum_k f_k z^k$ and $g(z) = \sum_k g_k z^k$, define a convolution operation $\circledast$ by

$$(f \circledast g)(z) = \sum_k \left(\sum_{\pi \in NC(k)} f_\pi g_{\pi^{-1}\tau_k} \right) z^k$$

where τ_k is the full cycle in S_k. Then Theorem 18.14 in [NS06] states that for any two f and g with $f_1 \neq 0 \neq g_1$, it holds that

$$\mathscr{F}(f \circledast g)(z) = \mathscr{F}(f)(z)\mathscr{F}(g)(z).$$

Then the S-transform of ρ can be written:

$$S_\rho(z) = (1 + z)\mathscr{F}(M_\rho)(z) = (1 + z)\mathscr{F}(M_\mu)(z)\mathscr{F}(M_\nu)(z)$$

$$= \frac{1}{1 + z} S_\mu(z) S_\nu(z).$$

This implies the desired result, as the S-transform of $\mathrm{MP}(1)$ is given by $\frac{1}{1+z}$, and the S-transform uniquely determines a compactly supported distribution. $\blacksquare$

We remark briefly that free independence can equivalently be formulated in terms of the vanishing of *free cumulants*, which are themselves defined in terms of sums over non-crossing partitions. We refer the interested reader to any of the previously cited references for a more in-depth discussion on the role of non-crossing partitions in free probability. For our purposes, the fact that non-crossing partitions are intimately related to free independence will allows us to later phrase random tensor network results in terms of free probability.

13.2 Entanglement spectrum of random tensor network states

We now return to studying random tensor network states. Consider a family of link states in $\mathscr{P}_=(V)$ with states ϕ_e along the edges $e \in E$ as in Eq. (12.3), and assume that along each edge, the bond dimensions scale with a parameter D, so $D_e = \Theta(D)$. Our key assumption is that the link states have *bounded spectral variation* – we formalize this condition by demanding that the empirical distribution of the rescaled entanglement spectrum of the link states

$$\mu_e^{(D)} := \sum_{i=1}^{D_e} \frac{1}{D_e} \delta_{D_e \lambda_{e,i}} \tag{13.7}$$

has all moments converging to the moments $m_{e,k}$ of a compactly supported probability distribution μ_e, as D goes to infinity. We assume that the link states are normalized, so $m_{e,1} = 1$. This condition implies that all elements of the entanglement spectrum of the link state are of order D^{-1}.

For a minimal cut γ_A, let $\mu_{\gamma_A}^{(D)}$ be the distribution for the spectrum of the tensor product of the link states in γ_A:

$$\mu_{\gamma_A}^{(D)} = \bigotimes_{e \in \gamma_A} \mu_e^{(D)} = \frac{1}{D_{\gamma_A}} \sum_{\{i_e\}} \delta_{D_{\gamma_A} \prod_{e \in \gamma_A} \lambda_{e,i_e}}$$

where $i_e = 1, \ldots, D_e$ and $D_{\gamma_A} = \prod_{e \in \gamma_A} D_e$. We define the tensor product of distributions as follows: if X_1 and X_2 are independent real valued random variables with distributions μ_{X_i}, then $\mu_{X_1} \otimes \mu_{X_2}$ is defined as the distribution of the product $X_1 X_2$. The distribution $\mu_{\gamma_A}^{(D)}$ has k-th moment given by $m_{\gamma_A,k}^{(D)} = \prod_{e \in \gamma_A} m_{e,k}^{(D)}$, and we can see that $m_{\gamma_A,k}^{(D)}$ converges to $m_{\gamma_A,k}$, the moments of the distribution

$$\mu_{\gamma_A} := \bigotimes_{e \in \gamma_A} \mu_e.$$

Let $\mathrm{spec}(\rho_A) = \{\lambda_{A,i}\}$ (recall that $\mathrm{spec}(\rho_A)$ is ordered in non-increasing order). Let γ_A be a cut for A. By a standard argument, the number of nonzero eigenvalues of ρ_A (that is, $\mathrm{rank}(\rho_A)$) is upper bounded by D_{γ_A}. If γ_A is the unique minimal cut, then we define

$$\mu_A^{(D)} := \frac{1}{D_{\gamma_A}} \sum_{i=1}^{D_{\gamma_A}} \delta_{D_{\gamma_A} \lambda_{A,i}}. \tag{13.8}$$

If there are multiple minimal cuts, it is ambiguous which γ_A, and hence, which D_{γ_A}, we should pick; we choose the cut for which D_{γ_A} is minimal in Eq. (13.8), and we will denote this minimal cut by $\gamma_{A,1}$. The moments of $\mu_A^{(D)}$ are given by

$$m_{A,k}^{(D)} := \int z^k \mathrm{d}\mu_A(z) = D_{\gamma_A}^{k-1} \sum_{i=1}^{D_{\gamma_A}} (\lambda_{A,i})^k.$$

Note that the distribution $\mu_A^{(D)}$ is random, and correspondingly, the moments $m_{A,k}^{(D)}$ are random variables. In contrast, the moments $m_{e,k}^{(D)}$ and $m_{\gamma_A,k}^{(D)}$ are numbers depending only on the bond dimension.

The theorem we want to prove will follow straightforwardly from a key intermediate result: as D goes to infinity, all the moments of the boundary distribution $\mu_A^{(D)}$ converge to the moments of μ_{γ_A}. We use the notation in Eq. (13.6) to write expressions like

$$m_{\gamma_A,\pi} = \prod_{l \in C(\pi)} m_{\gamma_A,l}$$

for a permutation $\pi \in S_k$. We will then apply the *method of moments* to show that convergence of moments implies convergence in distribution. As a remark on notation, in the error bounds in both the current section and Chapter 14, when we use $\mathcal{O}$-notation, the constants may depend on the graph underlying the tensor network (typically our bounds scale as $2^{|V_b|}$, where V_b is the set of bulk vertices).

Proposition 13.3. *If there exists a unique minimal cut γ_A for A, then*

$$\lim_{D \to \infty} \mathbb{E}m_{A,k}^{(D)} = m_{\gamma_A,k}. \tag{13.9}$$

If there exist exactly two minimal cuts $\gamma_{A,1}$ and $\gamma_{A,2}$, which do not intersect (meaning that $\gamma_{A,1} \cap \gamma_{A,2} = \emptyset$) and for which $\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}$ converges to a constant $t \leq 1$, then

$$\lim_{D \to \infty} \mathbb{E}m_{A,k}^{(D)} = \sum_{\pi \in NC(k)} t^{d(\pi,\mathrm{id})} m_{\gamma_{A,1},\tau^{-1}\pi} m_{\gamma_{A,2},\pi}. \tag{13.10}$$

Moreover, in both cases the variance goes to zero as D goes to infinity: for every k

$$\mathbb{E}\left[\left(m_{A,k}^{(D)} - \mathbb{E}\left[m_{A,k}^{(D)} \right] \right)^2 \right] = \mathcal{O}\left(\frac{1}{D} \right).$$

Proof. We first provide a sketch of the proof. It proceeds via the following steps:

(i) Write the expectation of the moments of $\mu_A^{(D)}$ as the partition function for a classical spin model, as in Section 12.1.1.

(ii) Show that the contributions from terms of the form given in the statement of the proposition dominate the partition function by carefully tracking the powers of D, and showing that all other contributions are suppressed polynomially in D.

(iii) Show that the variance of the moments vanishes in the limit $D \to \infty$ by direct computation.

We begin with Step 1. First, we observe that the k-th moment of $\mu_A^{(D)}$ is given by $m_{A,k}^{(D)} = D_{\gamma_A}^{k-1} \mathrm{tr}[\rho_A^k]$. Consider the expression in Eq. (12.16) for the replica trick with permutation π on A:

$$Z_{k,\pi} := \mathbb{E}\,\mathrm{tr}\left[R_A(\pi)\rho_A^{\otimes k} \right] = \sum_{\{\pi_x\} \in \mathscr{S}_{A,\pi}} \prod_{e=(xy) \in E} \prod_{l \in C(\pi_x^{-1}\pi_y)} \mathrm{tr}\left[\phi_{e,x}^l \right]. \tag{13.11}$$

Recall that the set $\mathscr{S}_{A,\pi}$, as defined in Eq. (12.13), consists of assignments of permutations to each $x \in V$, subject to $\pi_x = \pi$ for $x \in A$ and $\pi_x = \mathrm{id}$ for $x \in \bar{A}$. As in Eq. (12.14), if $\pi = \tau$, then we indeed have $Z_{k,\tau} = \mathbb{E}\,\mathrm{tr}[\rho_A^k]$, so

$$\mathbb{E}m_{A,k}^{(D)} = D_{\gamma_A}^{k-1} Z_{k,\tau}. \tag{13.12}$$

On the other hand, if we let $k = 2n$ and $\pi = \tilde{\tau} = (12\ldots n)(n+1\ n+2\ldots 2n)$, then we see that $Z_{k,\pi} = \mathbb{E}\left[\operatorname{tr}[\rho_A^n]^2\right]$, and hence

$$\mathbb{E}\left[\left(m_{A,n}^{(D)}\right)^2\right] = D_{\gamma_A}^{2n-2} Z_{k,\tilde{\tau}}. \tag{13.13}$$

Recall that $m_{e,l}^{(D)} = D_e^{l-1}\operatorname{tr}[\phi_{e,x}^l]$, and write

$$Z_{k,\pi} = \sum_{\{\pi_x\}\in\mathscr{S}_{A,\pi}} Z_k(\{\pi_x\})$$

where

$$
\begin{aligned}
Z_k(\{\pi_x\}) &:= \prod_{e=(xy)\in E} \prod_{l\in C(\pi_x^{-1}\pi_y)} \operatorname{tr}\left[\phi_{e,x}^l\right] \\
&= \prod_{e=(xy)\in E} D_e^{|C(\pi_x^{-1}\pi_y)|-k} \prod_{l\in C(\pi_x^{-1}\pi_y)} m_{e,l}^{(D)} \\
&= \prod_{e=(xy)\in E} D_e^{-d(\pi_x,\pi_y)} m_{e,\pi_x^{-1}\pi_y}^{(D)}.
\end{aligned}
\tag{13.14}
$$

This accomplishes Step 1: we have recast the problem of computing moments into a question of computing a partition function for a classical spin model with fixed boundary conditions.

For Step 2, we want to show that the dominant contribution(s) to $Z_{k,\pi}$ as D goes to infinity are those given in the statement of the proposition. This will simply be a matter of checking powers of D, and using the triangle inequality property of the Cayley distance. If Γ_A is the unique minimal cut, then we let $\pi_x^{\min} = \pi$ for $x \in \Gamma_A$ and $\pi_x^{\min} = \mathrm{id}$ for $x \in V \setminus \Gamma_A$, and we have

$$Z_k(\{\pi_x^{\min}\}) = D_{\gamma_A}^{-d(\pi,\mathrm{id})} \prod_{e\in\gamma_A} m_{e,\pi}^{(D)} = D_{\gamma_A}^{-d(\pi,\mathrm{id})} m_{\gamma_A,\pi}^{(D)}. \tag{13.15}$$

If there are exactly two minimal cuts $\Gamma_{A,1} \subset \Gamma_{A,2}$, we partition $V = V_1 \sqcup V_2 \sqcup V_3$ into three sets of vertices, with $V_1 = \Gamma_{A,1}$, $V_2 = \Gamma_{A,2} \cap \Gamma_{A,1}^c$ and $V_3 = \Gamma_{A,2}^c$. Now consider the permutations $\sigma \in S_k$ that are on a geodesic between π and id (recall that this implies that $d(\pi,\sigma) + d(\sigma,\mathrm{id}) = d(\pi,\mathrm{id})$), and consider the configuration given by

$$
\begin{aligned}
\pi_x^\sigma &= \pi \text{ for } x \in V_1, \\
\pi_x^\sigma &= \sigma \text{ for } x \in V_2, \\
\pi_x^\sigma &= \mathrm{id} \text{ for } x \in V_3.
\end{aligned}
$$

By hypothesis, $\gamma_{A,1}$ and $\gamma_{A,2}$ do not intersect, and hence, the edges in each cut are distinct. Then this configuration has weight

$$
\begin{aligned}
Z_k(\{\pi_x^\sigma\}) &= \prod_{e_1\in\gamma_{A,1}} D_{e_1}^{-d(\pi,\sigma)} \prod_{l_1\in C(\pi^{-1}\sigma)} m_{e_1,l_1}^{(D)} \prod_{e_2\in\gamma_{A,2}} D_{e_2}^{-d(\sigma,\mathrm{id})} \prod_{l_2\in C(\sigma)} m_{e,l}^{(D)} \\
&= D_{\gamma_{A,1}}^{-d(\pi,\mathrm{id})} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} \prod_{e_1\in\gamma_{A,1}} m_{e_1,\pi^{-1}\sigma}^{(D)} \prod_{e_2\in\gamma_{A,2}} m_{e_2,\sigma}^{(D)} \\
&= D_{\gamma_{A,1}}^{-d(\pi,\mathrm{id})} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1},\pi^{-1}\sigma}^{(D)} m_{\gamma_{A,2},\sigma}^{(D)},
\end{aligned}
\tag{13.16}
$$

where $D_{\gamma_A,1}/D_{\gamma_A,2}$ converges to t, by assumption. Now, to show that these configurations yield the dominant contributions, we will need to use that $D_e = \Theta(D)$, so let us write $\frac{D}{D_e} = C_e^{(D)} = \Theta(1)$. Then for general configurations labeled by π_x, we may rewrite Eq. (13.14) as

$$Z_k(\{\pi_x\}) = \prod_{e=(xy)\in E} D^{-d(\pi_x,\pi_y)}(C_e^{(D)})^{d(\pi_x,\pi_y)} m_{e,\pi_x^{-1}\pi_y}^{(D)}.$$

The configurations we claimed to be dominant satisfy $Z_k(\{\pi\}) = \Theta(D^{-m(A)d(\pi,\mathrm{id})})$, where we recall that $m(A)$ is the size of a minimal cut for A. Now we will show that all other configurations satisfy $Z_k(\{\pi\}) = \mathcal{O}(D^{-m(A)d(\pi,\mathrm{id})-1})$. To this end, consider some arbitrary configuration $\{\pi_x\} \in \mathscr{S}_{A,\pi}$. Let P be a maximal set of edge-disjoint paths in G from A to $\bar{A}$. It is a well-known fact that such a set has size $m(A)$, by the max-flow min-cut theorem. Let

$$C_k^{(D)} := \left(\max_{e\in E, l=1,\dots,k}(C_e^{(D)})^{l-1}\right)\left(\max_{e\in E, \pi\in S_k} m_{e,\pi}^{(D)}\right).$$

Then we may bound

$$Z_k(\{\pi_x\}) \le (C_k^{(D)})^{|E|} \prod_{e=(xy)\in E} D^{-d(\pi_x,\pi_y)} \le (C_k^{(D)})^{|E|} \prod_{p\in P} D^{-\sum_{e=(xy)\in p} d(\pi_x,\pi_y)}. \tag{13.17}$$

The first inequality is clear from the definition of $C_k^{(D)}$, and in the second inequality, we simply restrict to a subset of the edges we sum over. Note that $C_k^{(D)} = \mathcal{O}(1)$. Then, by the triangle inequality for the Cayley distance d, it holds that

$$\sum_{e=(xy)\in p} d(\pi_x,\pi_y) \ge d(\pi,\mathrm{id})$$

with equality if and only if the only edges (xy) for which $\pi_x \ne \pi_y$ are on a path in P, and each of the paths is a geodesic. Then we conclude

$$Z_k(\{\pi_x\}) \le C_k^{(D)} \prod_{p\in P} D^{-d(\pi,\mathrm{id})} = C_k^{(D)} D^{-m(A)d(\pi,\mathrm{id})},$$

and we see that the weight of every configuration can be bounded by the product of a $\mathcal{O}(1)$ number and a polynomial in D.

Now, as promised, we show that if $\{\pi_x\}$ is not one of the minimal configurations described above, we actually have

$$Z_k(\{\pi_x\}) = \mathcal{O}(D^{-m(A)d(\pi,\mathrm{id})-1}). \tag{13.18}$$

To see this, we rewrite the triangle inequality for the Cayley distance as:

$$\prod_{e=(xy)\in E} D^{-d(\pi_x,\pi_y)} \le \prod_{p\in P} D^{-\sum_{e=(xy)\in p} d(\pi_x,\pi_y)} \le D^{-m(A)d(\pi,\mathrm{id})} \tag{13.19}$$

with equality if and only if the π_x are on a geodesic path in P. We now show that this is satisfied only for the configurations we claimed to be minimal. Assume that $\{\pi_x\} \in \mathscr{S}_{A,\pi}$ is such that the inequalities in Eq. (13.19) are equalities and let

$$\Delta_n = \{x \in V \text{ such that } d(\pi_x,\pi) \le n\}.$$

Then $\Delta_n \in C(A)$ for $0 \le n < d(\pi, \mathrm{id})$, and we denote by δ_n the associated set of edges crossing the cut. Each edge $(xy) \in \delta_n$ must be such that $\pi_x \ne \pi_y$, so it must be on a path in P, and because the permutations are geodesics along the paths, they must be on different paths. Hence $|\delta_n| \le |P| = m(A)$, implying each Δ_n is a minimal cut. This immediately implies the claim if there is a unique minimal cut, since we must have $\Delta_{d(\pi,\mathrm{id})-1} = \Delta_0 = \Gamma_A$. If there are exactly two minimal cuts, then we must have that $\pi_x = \pi$ for $x \in V_1$, $\pi_x = \mathrm{id}$ for $x \in V_3$, and there must be some l such that for all $x \in V_2$ we have $d(\pi, \pi_x) = l$ and $d(\pi_x, \mathrm{id}) = d(\pi, \mathrm{id}) - l$. Then in order to have equality in Eq. (13.19), we must have that for all $x \in V_2$, π_x equals some fixed permutation σ, because the assumption of having exactly two cuts implies that V_2 is connected, and we must have $d(\pi_x, \pi_y) = 0$ for all $(xy) \in E$ with $x, y \in V_2$. This proves Eq. (13.18).

In conclusion, if there is a unique minimal cut, then by Eq. (13.15) and Eq. (13.18), we find

$$Z_{k,\pi} = D_{\gamma_A}^{-d(\pi,\mathrm{id})} m_{\gamma_A,\pi}^{(D)} + \mathcal{O}(D^{-m(A)d(\pi,\mathrm{id})-1}), \tag{13.20}$$

and if there are exactly two (non-intersecting) cuts, then by Eq. (13.16) and Eq. (13.18), we find

$$Z_{k,\pi} = \sum_{\sigma, d(\pi,\sigma)+d(\sigma,\mathrm{id})=d(\pi,\mathrm{id})} D_{\gamma_{A,1}}^{-d(\pi,\mathrm{id})} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1},\pi^{-1}\sigma}^{(D)} m_{\gamma_{A,2},\sigma}^{(D)} \tag{13.21}$$

$$+ \mathcal{O}(D^{-m(A)d(\pi,\mathrm{id})-1}).$$

Finally, we set $\pi = \tau$ for the full cycle τ and we use Eq. (13.12). For a unique minimal cut γ_A, by Eq. (13.20)

$$\begin{aligned} \mathbb{E} m_{\gamma_A,k} &= D_{\gamma_A}^{k-1} Z_{k,\tau} \\ &= m_{\gamma_A,\pi}^{(D)} + \mathcal{O}(D_{\gamma_A}^{k-1} D^{-m(A)(k-1)-1}) \\ &= m_{\gamma_A,k}^{(D)} + \mathcal{O}\left(\frac{1}{D}\right). \end{aligned} \tag{13.22}$$

using $d(\tau, \mathrm{id}) = k - 1$ and $D_{\gamma_A} = \Theta(D^{m(A)})$. This proves Eq. (13.9) as $m_{\gamma_A,k}^{(D)}$ converges to $m_{\gamma_A,k}$.

For two non-intersecting minimal cuts, we saw that the dominant contribution is due to configurations $\{\pi_x^\sigma\}$ for σ on a geodesic between τ and id. Then applying the observation that σ is on such a geodesic if and only if σ is a non-crossing partition similarly yields that by Eq. (13.21)

$$\begin{aligned} \mathbb{E} m_{\gamma_A,k} &= D_{\gamma_A}^{k-1} Z_{k,\tau} \\ &= D_{\gamma_A}^{k-1} \sum_{\sigma \in NC(k)} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1},\pi^{-1}\sigma}^{(D)} m_{\gamma_{A,2},\sigma}^{(D)} + \mathcal{O}(D_{\gamma_A}^{k-1} D^{-m(A)(k-1)-1}) \\ &= \sum_{\sigma \in NC(k)} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1},\tau^{-1}\sigma}^{(D)} m_{\gamma_{A,2},\sigma}^{(D)} + \mathcal{O}\left(\frac{1}{D}\right). \end{aligned} \tag{13.23}$$

Since $m_{\gamma_{A,2},\sigma}^{(D)} \to m_{\gamma_{A,2},\sigma}$ and $D_{\gamma_{A,1}}/D_{\gamma_{A,2}} \to t$, this proves Eq. (13.10).

This accomplishes Step 2: we have shown that the configurations $\{\pi_x^{\min}\}$ (in case of a unique minimal cut for A) and $\{\pi_x^\sigma\}$ (in case there are exactly minimal cuts for A)

dominate in the computation of the expectation of the k-th moment of μ_A in terms of powers of D.

We complete the proof by showing that the variance of $m_{A,k}$ vanishes as $D \to \infty$. We use the observation in Eq. (13.13), applying the analysis of $Z_{2k,\pi}$ to the case where we take $\pi = \tilde{\tau} = (12\ldots k)(k+1\,k+2\ldots 2k)$. If there is a unique minimal cut, then using Eq. (13.20) and the fact that $d(\tilde{\tau}, \mathrm{id}) = 2k - 2$, we find

$$\mathbb{E}\left[\left(m_{A,k}^{(D)}\right)^2\right] = D_{\gamma_A}^{2k-2} Z_{2k,\tilde{\tau}} = \prod_{l \in C(\tilde{\tau})} m_{\gamma_A,l}^{(D)} + \mathcal{O}(D_{\gamma_A}^{2k-2} D^{-m(A)(2k-2)-1})$$

$$= (m_{\gamma_A,k}^{(D)})^2 + \mathcal{O}\left(\frac{1}{D}\right).$$

By Eq. (13.22) we know that $(\mathbb{E}m_{A,k}^{(D)})^2 = (m_{\gamma_A,k}^{(D)} + \mathcal{O}(\frac{1}{D}))^2$, and we conclude that the variance obeys

$$\mathbb{E}\left[\left(m_{A,k}^{(D)} - \mathbb{E}\left[m_{A,k}^{(D)}\right]\right)^2\right] = \mathbb{E}\left[(m_{A,k}^{(D)})^2\right] - \left(\mathbb{E}m_{A,k}^{(D)}\right)^2 = \mathcal{O}\left(\frac{1}{D}\right).$$

For the case with exactly two minimal cuts, a similar argument holds. Here, the key observation is that σ is on a geodesic between $\tilde{\tau}$ and id if and only if $\sigma = \sigma_1 \sigma_2$ where σ_1 is on a geodesic between $(12\ldots k)$ and id and σ_2 is on a geodesic between $(k+1\,k+2\ldots 2k)$. Using Eq. (13.21), this implies that we can bound $\mathbb{E}(m_{A,k}^{(D)})^2 = D_{\gamma_A}^{2k-2} Z_{2k,\tilde{\tau}}$ as

$$\sum_{\sigma_1,\sigma_2 \in NC(k)} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma_1,\mathrm{id})+d(\sigma_2,\mathrm{id})} m_{\gamma_{A,1},\tau^{-1}\sigma_1}^{(D)} m_{\gamma_{A,1},\tau^{-1}\sigma_2}^{(D)} m_{\gamma_{A,2},\sigma_1}^{(D)} m_{\gamma_{A,2},\sigma_2}^{(D)} + \mathcal{O}\left(\frac{1}{D}\right)$$

$$= \left(\sum_{\sigma \in NC(k)} \left(\frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}}\right)^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1},\tau^{-1}\sigma}^{(D)} m_{\gamma_{A,2},\sigma}^{(D)}\right)^2 + \mathcal{O}\left(\frac{1}{D}\right).$$

By Eq. (13.23), we see that this coincides with $(\mathbb{E}m_{A,k}^{(D)})^2$, up to $\mathcal{O}(\frac{1}{D})$, and hence Eq. (13.11) holds. ∎

We now have the ingredients to prove that the entanglement spectrum of random tensor networks with link states with bounded spectral variation can be written in a simple fashion. We will use the *method of moments* to translate the above result on convergence of moments to convergence in distribution. The basic statement is that, given certain conditions on the distributions in question, if the moments of a sequence of distribution μ_n converge to those of μ, then $\mu_n \Rightarrow \mu$ – see for instance Theorem 30.8 in [Bil08].

The method of moments is valid, so long as a distribution μ is completely determined by its moments. This occurs if the k-th moment $m_{\mu,k}$ is bounded as

$$m_{\mu,k} \le AB^k k! \tag{13.24}$$

for constants A, B for all k. If the distributions have compact support, as in Proposition 13.3, then this condition is satisfied.[1]

---

[1] A basic example of a distribution which does not have compact support, but is nevertheless uniquely determined by its moments is a standard Gaussian distribution. On the other hand, a standard example of distributions that are not determined by their moments are the densities on $\mathbb{R}_{\ge 0}$ with $d\mu_1(x) = \sqrt{2\pi} x^{-1} e^{-(\log x)^2/2} dx$ and $d\mu_2(x) = (1 + \sin(2\pi \log x)) d\mu_1(x)$, for which it can be verified that the n-th moments of both distributions are equal to $e^{-n^2/2}$, while the distributions are clearly not identical.

Now that we have established the convergence of moments in Proposition 13.3, we have our main result of the (conditional) convergence in distribution. As in Proposition 13.3 we consider a family of random tensor network state with link states with bounded spectral variation with increasing D, as defined in the beginning of this section.

Theorem 13.4. *If there exists a unique minimal cut γ_A for A, then $\mu_A^{(D)} \Rightarrow \mu_{\gamma_A}$, in probability, as $D \to \infty$. If there exist exactly two minimal cuts $\gamma_{A,1}$ and $\gamma_{A,2}$, which do not intersect and for which $\lim_{D \to \infty} \frac{D_{\gamma_{A,1}}}{D_{\gamma_{A,2}}} = t \leq 1$, then $\mu_A^{(D)} \Rightarrow MP(1) \boxtimes \mu_{\gamma_{A,1}} \boxtimes \mu_{\gamma_{A,2}}(t)$, in probability, where $\mu_{\gamma_{A,2}}(t) = (1-t)\delta_0 + t\mu_{\gamma_{A,2}}$*

Proof. It is straightforward to see that the k-th moment of $\mu_{\gamma_{A,2}}(t)$ is given by $t^{k-1}m_{\gamma_{A,2},k}$, and then the result follows immediately from Proposition 13.3, Theorem 13.2, and the method of moments. Because we assumed that for any minimal cut γ_A for A, the limiting distributions μ_{γ_A} are compactly supported, they are uniquely determined by their moments. Hence, the method of moments is valid, and the convergence of moments implies convergence in distribution. ∎

Remark 13.5. In Theorem 13.4 we assumed that the two cuts where non-intersecting. What happens if there are still only exactly two minimal cuts, but $\gamma_{A,1} \cap \gamma_{A,2}$ is nonempty? This extension is straightforward. Let $\gamma_A^{(a)} := \gamma_{A,1} \cap \gamma_{A,2}$ and let $\gamma_{A,i}^{(b)} = \gamma_{A,i} \setminus \gamma_A^{(a)}$ for $i = 1,2$. In line with previous notation, let $\mu_{\gamma_A^{(a)}}$ and $\mu_{\gamma_{A,i}^{(b)}}$ denote the corresponding limiting distributions of the entanglement spectra along these sets, with moments $m_{\gamma_A^{(a)},k}$ and $m_{\gamma_{A,i}^{(b)},k}$. The only step in the proof of Proposition 13.3 where we used that the cuts were non-intersecting is when we computed the value of $Z_k(\{\pi_x\})$ for the optimal configuration. If the cuts do intersect, and we consider the configuration with $\pi_x = \tau$ for $x \in V_1$ with τ the complete cycle, $\pi_x = \sigma$ for $x \in V_2$ and $\sigma \in NC(k)$, and $\pi_x = \mathrm{id}$ for $x \in V_3$, then a quick calculation shows

$$Z_k(\{\pi_x\}) \to D_{\gamma_{A,1}}^{-d(\pi,\mathrm{id})}(D_{\gamma_A,1}/D_{\gamma_A,2})^{d(\sigma,\mathrm{id})} \prod_{e \in \gamma_A^{(a)}} m_{e,k} \prod_{e_1 \in \gamma_{A,1}^{(b)}} m_{e_1,\pi^{-1}\sigma} \prod_{e_2 \in \gamma_{A,2}^{(b)}} m_{e_2,\sigma}.$$

Apart from this modification, the proof of Proposition 13.3 is still valid, leading to

$$Z_{k,\tau} = m_{\gamma_A^{(a)},k} \sum_{\sigma \in NC(k)} t^{d(\sigma,\mathrm{id})} m_{\gamma_{A,1}^{(b)},\tau^{-1}\sigma} m_{\gamma_{A,2}^{(b)},\sigma}.$$

If, in Theorem 13.4, we do not assume that the cuts are nonintersecting, then the partition function above leads to a limiting distribution given by

$$\mu_{\gamma_A^{(a)}} \otimes \left(MP(1) \boxtimes \mu_{\gamma_{A,1}^{(b)}} \boxtimes \mu_{\gamma_{A,2}^{(b)}}(t) \right).$$

13.3 Nontrivial link states and entanglement negativity

As another application of the theory of free probability, we will compute the entanglement negativity spectrum for random tensor network states with link states with bounded spectra. In [DQW21], it was shown how to compute the entanglement negativity spectrum for a random tensor network state with maximally entangled link states using a replica trick. Using the methods from the previous section, we can analyze

the negativity for entangled link states with bounded spectral variation. We remark that similar computations have recently been performed in [DMW21] in the context of replica wormholes, and our assumption on the link states is a generalization of the "pairwise connected regime" in [DMW21]. Another work investigating nontrivial entanglement negativity spectra in random tensor networks is [KFNR21], where they focus on the effect of having multiple minimal cuts in the network. As our analysis will be a straightforward combination of the arguments in [DQW21] and Section 13.2, we will be rather concise; the main message of this section is to show that the language of free probability applies to other random tensor network computations as well.

We first recall how negativity functions as an entanglement measure for mixed states. Let $\mathcal{T}$ be the superoperator which maps an operator $X \mapsto X^{\mathsf{T}}$, and $\mathcal{I}$ the identity superoperator. For $\rho_{AB} \in \mathcal{P}(AB)$,

$$\rho_{AB}^{T_B} := (\mathcal{I}_A \otimes \mathcal{T}_B)(\rho_{AB})$$

is the partial transpose of ρ_{AB} on the B system. The *logarithmic* or *entanglement negativity* is given by

$$E_N(\rho_{AB}) = \log \frac{\|\rho_{AB}^{T_B}\|_1}{\mathrm{tr}[\rho]}.$$

It is a measure for the entanglement of the mixed state ρ_{AB}: if $E_N(\rho_{AB}) > 0$ the state must be entangled. We call $\mathrm{spec}(|\rho_{AB}^{T_B}|)$ the *entanglement negativity spectrum*. In analogy to the Rényi entropies, we can generalize the logarithmic negativity to a one-parameter family of negativities. The k-th Rényi negativity is given by

$$N_k(\rho_{AB}) = \mathrm{tr}\left[(\rho_{AB}^{T_B})^k\right].$$

If we let $N_m^{(\mathrm{even})}(\rho_{AB}) = N_{2m}(\rho_{AB})$, then the logarithmic negativity is obtained an analytic continuation in the Rényi index $m \to \frac{1}{2}$ of $\log(N_m^{(\mathrm{even})}(\rho_{AB}))$. More precisely, in the expression

$$\log \sum_{\lambda \in \mathrm{spec}(\rho_{AB}^{T_B})} |\lambda|^\alpha,$$

we may take $\alpha \to \frac{1}{2}$ to obtain $E_N(\rho_{AB}) + \log \mathrm{tr}[\rho]$.

In the context of random tensor networks, we partition the boundary in three regions: $V_\partial = A \sqcup B \sqcup C$, and we would like to compute the Rényi negativities of the reduced state ρ_{AB}. We will then use this to determine the entanglement negativity spectrum, and compute the entanglement negativity. The idea is that the k-th Rényi negativity can be computed using a replica trick, by placing the full cycle $\tau_k = (12\ldots k) \in S_k$ on A and $\tau_k^{-1} = (k\, k+1\ldots 1)$ on B:

$$N_k(\rho_{AB}) = \mathrm{tr}\left[\rho_{AB}^{\otimes k}\left(R_A(\tau) \otimes R_B(\tau^{-1})\right)\right]$$
$$= \mathrm{tr}\left[\rho_{ABC}^{\otimes k}\left(R_A(\tau) \otimes R_B(\tau^{-1}) \otimes R_C(\mathrm{id})\right)\right].$$

Let us first discuss the case with maximally entangled link states, following [DQW21]. The same arguments as in Section 12.1.2 show that one can compute the expectation

(a) Illustration of the case where $\gamma_C = \gamma_A \cup \gamma_B$, where the dominant configuration is given by τ_k on Γ_A, τ_k^{-1} on Γ_B and id on Γ_C.

(b) Illustration of the case where $\gamma_C \neq \gamma_A \cup \gamma_B$, where the dominant configuration is given by τ_k on Γ_A, τ_k^{-1} on Γ_B and id on Γ_C and some non-crossing pairing π_1 on the domain V_1.
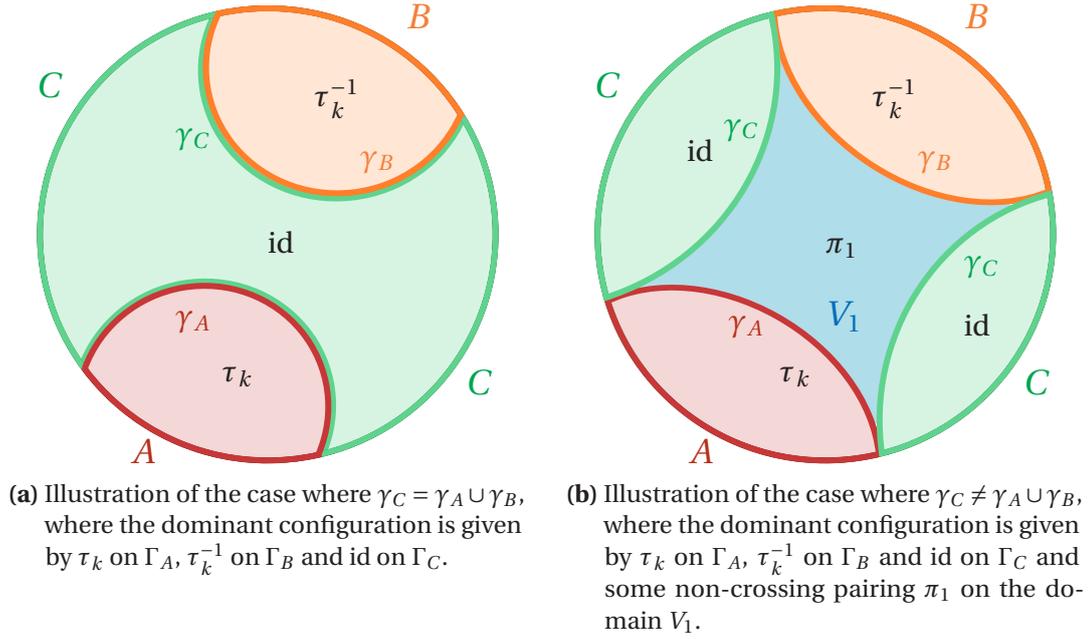
Figure 13.1: Tensor networks with one and two minimal cuts. The relevant ground state configuration domains are denoted by Γ_A.

of $N_k(\rho_{AB})$ for a random tensor network state using a spin model, now with boundary conditions of τ_k on A, τ_k^{-1} on B and id on C. We will assume that the minimal cuts Γ_A, Γ_B and Γ_C are unique. Note that the minimal cut for AB is given by $\Gamma_{AB} = \Gamma_C^c$. From the theory of multi-commodity flows, it is known that there exist sets of edge-disjoint paths $P = P_{AB} \cup P_{AC} \cup P_{BC}$, where P_{AB} consists of paths from A to B, and similarly for sets of paths P_{AC} and P_{BC}, and which are such that

$$|P_{AB}| + |P_{AC}| = |\gamma_A|, \qquad |P_{AB}| + |P_{BC}| = |\gamma_B|, \qquad |P_{AC}| + |P_{BC}| = |\gamma_C|.$$

This can be used to show (in analogous fashion to the proof of Proposition 13.3) that, if $k = 2n$ is even, any spin model configuration contributing to $\mathbb{E}N_k(\rho_{AB})$ is of order $\mathcal{O}(D^{-(n-1)(|\gamma_A|+|\gamma_B|)-n|\gamma_C|})$. If $k = 2n+1$ is odd, any spin model configuration contributing to $\mathbb{E}N_k(\rho_{AB})$ is of order of magnitude $\mathcal{O}(D^{-n(|\gamma_A|+|\gamma_B|+|\gamma_C|)})$.

In order to determine what happens as $D \to \infty$, we need to determine the dominant configurations. Let r be the number of connected components of $V \setminus (\Gamma_A \cup \Gamma_B \cup \Gamma_C)$. There are two distinct cases. The first is when the the minimal cut for AB (which is the complement of the minimal cut for C) is the union of the minimal cuts for A and B, so $\Gamma_{AB} = \Gamma_A \cup \Gamma_B$ and hence $\gamma_{AB} = \gamma_A \cup \gamma_B$. Then the minimal cuts naturally partition the bulk vertices into three cuts Γ_A, Γ_B and Γ_C, and we have $r = 0$. In this case, the dominant configurations in the spin model are those where the vertices in Γ_A are assigned τ_k, those in Γ_B are assigned τ_k^{-1} and those in Γ_C are assigned id. This is illustrated in Fig. 13.1a.

The second case is when $\Gamma_A \cup \Gamma_B \subsetneq \Gamma_{AB}$ and hence $\gamma_{AB} \neq \gamma_A \cup \gamma_B$. Now, we have again the domains Γ_A, Γ_B and Γ_C, but upon removing these vertices, there may also be connected components $V_1, \ldots, V_r$ which are not connected to A, B or C. Here, the minimal configurations are those for which, again, the vertices in Γ_A are assigned τ_k, those in Γ_B are assigned τ_k^{-1} and those in Γ_C are assigned id, and where in each component V_i the vertices are assigned a permutation π_i which is such that it satisfies three

conditions: it must be on a geodesic between τ_k and τ_k^{-1}, on a geodesic between τ_k and id and on a geodesic between τ_k^{-1} and id. If $k = 2n$ is even, such permutations are given by *non-crossing pairings*: permutations corresponding to non-crossing partitions in which each cycle has length 2. The set of non-crossing pairings on $2n$ elements is in bijection with the set of non-crossing partitions on n elements, so the number of non-crossing pairings on $2n$ elements is given by $|NC(n)| = C_n$. One way to obtain this correspondence is as follows. If π is a non-crossing pairing, $\tau_{2n}\pi$ will map even numbers to even numbers, and restricting to the even numbers and relabeling $2i \mapsto i$ yields a non-crossing partition $\sigma \in NC(n)$. Moreover, restricting to the odd numbers and relabeling $2i + 1 \mapsto i$ yields the non-crossing partition $\sigma^{-1}\tau \in NC(n)$. This leads to C_n^r dominant contributions to $\mathbb{E}N_{2n}(\rho_{AB})$ which are of size $D^{-(n-1)(|\gamma_A|+|\gamma_B|)-n|\gamma_C|}$ since we can choose a non-crossing pairing π_i for each component. Such a configuration is illustrated in Fig. 13.1b.

For odd $k = 2n + 1$, we similarly have permutations which correspond to a non-crossing partition. These permutations have a single fixed point and all other cycles with length 2. This leads to $((2n + 1)C_n)^r$ dominant contributions to $\mathbb{E}N_{2n}(\rho_{AB})$, of size $D^{-(n-1)(|\gamma_A|+|\gamma_B|)-n|\gamma_C|}$. We also note that $\mathrm{rank}(\rho_{AB}^{T_B}) \leq D^{|\gamma_A|+|\gamma_A|}$. If $\mathrm{spec}(\rho_{AB}^{T_B}) = \{s_i\}$, then we define the measure

$$\mu_{AB}^{(D)} = \frac{1}{D^{|\gamma_A|+|\gamma_B|}} \sum_{i=1}^{D^{|\gamma_A|+|\gamma_B|}} \delta_{D^{\frac{1}{2}(|\gamma_A|+|\gamma_B|+|\gamma_C|)}s_i}. \tag{13.25}$$

This has moments given by

$$m_{AB,k}^{(D)} = \int x^k \mathrm{d}\mu_{AB}^{(D)}(x) = D^{(\frac{k}{2}-1)(|\gamma_A|+|\gamma_B|)+\frac{k}{2}|\gamma_C|}N_k(\rho_{AB}).$$

If we take the expectation of the moments, we again need to distinguish the two cases. If $|\gamma_A| + |\gamma_B| = |\gamma_C|$, we see that the powers of D cancel for the dominant configurations, so $m_{AB,k}^{(D)} \to 1$ for all k. On the other hand, for $|\gamma_A| + |\gamma_B| > |\gamma_C|$, we see that for $D \to \infty$ with odd k, we have $\mathbb{E}m_{AB,k}^{(D)} \to 0$. For even k, we recover the degeneracy of the dominant configurations, leading to

$$\lim_{D\to\infty} m_{AB,k}^{(D)} = \begin{cases} 1 & \text{if } |\gamma_A| + |\gamma_B| = |\gamma_C|, \\ 0 & \text{if } k \text{ odd and } |\gamma_A| + |\gamma_B| > |\gamma_C|, \\ C_{k/2}^r & \text{if } k \text{ even and } |\gamma_A| + |\gamma_B| > |\gamma_C|, \end{cases} \tag{13.26}$$

where r is the number of connected components of $V \setminus (\Gamma_A \cup \Gamma_B \cup \Gamma_C)$. In fact, one can show that, as in Proposition 13.3, the variance goes to zero as well, and hence the method of moments allows one to conclude that $\mu_{AB}^{(D)}$ converges weakly, in probability, to

$$\mu_{AB} = \begin{cases} \sigma^{\otimes r} & \text{if } r > 0, \\ \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1} & \text{if } r = 0 \text{ and } |\gamma_A| + |\gamma_B| > |\gamma_C|, \\ \delta_1 & \text{if } r = 0 \text{ and } |\gamma_A| + |\gamma_B| = |\gamma_C|, \end{cases} \tag{13.27}$$

where σ is the semi-circle distribution with density

$$\mathrm{d}\sigma(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbb{1}_{|x|\leq 2}\,dx.$$

Alternatively, one may study the empirical distribution of the *squared* entanglement negativity spectrum

$$v_{AB}^{(D)} = \frac{1}{D^{|\gamma_A|+|\gamma_B|}} \sum_i \delta_{D^{|\gamma_A|+|\gamma_B|+|\gamma_C| s_i^2}}. \tag{13.28}$$

This distribution has k-th moment given by $m_{AB,2k}^{(D)}$, and on comparison with the limiting moments in Eq. (13.26), one can conclude that $v_{AB}^{(D)} \Rightarrow v_{AB}$, in probability, where

$$v_{AB} = MP(1)^{\otimes r}. \tag{13.29}$$

The logarithmic negativity can be computed using the distribution $\mu_{AB}^{(D)}$ or $v_{AB}^{(D)}$ as

$$E_N(\rho_{AB}) = \log \int |\lambda| d\mu_{AB}^{(D)}(\lambda) + \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|) - \log \text{tr}[\rho] \tag{13.30}$$

$$= \log \int \sqrt{\lambda} dv_{AB}^{(D)}(\lambda) + \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|) - \log \text{tr}[\rho]. \tag{13.31}$$

The convergence of $v_{AB}^{(D)}$ to v_{AB} implies[2] that $E_N(\rho_{AB}) - \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|)$ converges in probability to

$$\log \int \sqrt{\lambda} dv_{AB}(\lambda) = r \log \frac{8}{3\pi}.$$

See Appendix D of [DQW21] for details and proofs.

A straightforward combination of the arguments in Section 13.2 and [DQW21] shows that the same configurations are the dominant contributions for link states with bounded spectral variation as in Section 13.2. To determine the limiting distribution in this case, we can generalize Eq. (13.27) in the same fashion as in Section 13.2. We assume the minimal cuts Γ_A, Γ_B and Γ_C are unique. We also assume that $\gamma_A \cap \gamma_B = \emptyset$, and in the case where $\gamma_C = \gamma_{AB} \neq \gamma_A \cup \gamma_B$ (so $|\gamma_A| + |\gamma_B| > |\gamma_C|$), all pairwise intersections between γ_A, γ_B and γ_C are empty. This excludes the case where $|\gamma_A| + |\gamma_B| > |\gamma_C|$, but $r = 0$. We let $\gamma_{A,i}$ and $\gamma_{B,i}$ denote the components of γ_A and γ_B which are connected to V_i, and we let $\mu_{\gamma_{A,i}}^{(D)}$ and $\mu_{\gamma_{B,i}}^{(D)}$ denote the distribution of the link state spectrum along these sets, with associated k-th moments $m_{\gamma_{A,i},k}^{(D)}$, $m_{\gamma_{B,i},k}^{(D)}$, which we assume to converge to the moments $m_{\gamma_{A,i},k}$, $m_{\gamma_{B,i},k}$ of compactly supported distributions $\mu_{\gamma_{A,i}}$ and $\mu_{\gamma_{B,i}}$. For convenience, we assume $D_e = D$ for all edges $e \in E$.

We can now compute the dominant contributions to $\mathbb{E}N_k(\rho_{AB})$. If $\gamma_C = \gamma_A \cup \gamma_B$, then there is a unique dominant configuration, which contributes $D^{-(k-1)|\gamma_C|} m_{\gamma_C,k}$. If $|\gamma_A| + |\gamma_B| > |\gamma_C|$ and $k = 2n$ is even, consider the configuration which assigns π_i to V_i, where each π_i is a non-crossing pairing. For each edge $e \in \gamma_C$, we have $m_{e,\pi_i}^{(D)} = (m_{e,2}^{(D)})^n$, so this configuration contributes

$$D^{-(n-1)(|\gamma_A|+|\gamma_B|)-n|\gamma_C|} \left(m_{\gamma_C,2}^{(D)}\right)^n \prod_{i=1}^r \left(m_{\gamma_{A,i},\tau_{2n}^{-1}\pi_i}^{(D)} m_{\gamma_{B,i},\tau_{2n}\pi_i}^{(D)}\right).$$

---

[2]The function $f(\lambda) = \sqrt{\lambda}$ is not in $C_b(\mathbb{R})$, but the method of moments actually shows a stronger convergence, allowing test functions to have polynomial growth.

Recalling the construction of the equivalence between $NC(n)$ and non-crossing pairings on $2n$ elements, we see that

$$m^{(D)}_{\gamma_{B,i},\tau_{2n}\pi_i} = m^{(D)}_{\gamma_{B,i},\sigma_i} m^{(D)}_{\gamma_{B,i},\sigma_i^{-1}\tau_n}$$

for some unique $\sigma_i \in NC(n)$. Similarly, one may verify

$$m^{(D)}_{\gamma_{A,i},\tau_{2n}^{-1}\pi_i} = m^{(D)}_{\gamma_{A,i},\sigma_i} m^{(D)}_{\gamma_{A,i},\sigma_i^{-1}\tau_n}.$$

This implies that the contribution of all dominant configurations is given by

$$\left(m^{(D)}_{\gamma_C,2}\right)^n \prod_{i=1}^{r} \left(\sum_{\sigma \in NC(n)} m^{(D)}_{\gamma_{A,i},\sigma} m^{(D)}_{\gamma_{B,i},\sigma} m^{(D)}_{\gamma_{A,i},\sigma^{-1}\tau_n} m^{(D)}_{\gamma_{B,i},\sigma^{-1}\tau_n} \right)$$

As in the maximally entangled case, upon rescaling, the odd moments vanish as $D \to \infty$. In conclusion, the resulting asymptotic moments are given by

$$\lim_{D \to \infty} \mathbb{E} m^{(D)}_{AB,k} = \begin{cases} m_{\gamma_C,k} & \text{if } |\gamma_A| + |\gamma_B| = |\gamma_C|, \\ 0 & \text{if } k \text{ odd and } |\gamma_A| + |\gamma_B| > |\gamma_C|, \\ m_k & \text{if } k \text{ even and } |\gamma_A| + |\gamma_B| > |\gamma_C| \end{cases} \tag{13.32}$$

with

$$m_{2n} = m^n_{\gamma_C,2} \prod_{i=1}^{r} \left(\sum_{\sigma \in NC(n)} m_{\gamma_{A,i},\sigma} m_{\gamma_{B,i},\sigma} m_{\gamma_{A,i},\sigma^{-1}\tau_n} m_{\gamma_{B,i},\sigma^{-1}\tau_n} \right).$$

As before, one can also show, in similar fashion to the proof of Proposition 13.3, that the variance of the moments goes to zero as $D \to \infty$. For the case $|\gamma_A| + |\gamma_B| > |\gamma_C|$, we consider $\nu^{(D)}_{AB}$ similar to Eq. (13.28), but with an additional rescaling by $m_{\gamma_C,2}$:

$$\nu^{(D)}_{AB} = \frac{1}{D^{|\gamma_A|+|\gamma_B|}} \sum_i \delta_{D^{|\gamma_A|+|\gamma_B|+|\gamma_C|} m^{-1}_{\gamma_C,2} s_i^2}.$$

This has moments, which compute $N^{(\text{even})}_k(\rho_{AB})$, converging to

$$\lim_{D \to \infty} \mathbb{E} \int x^k \mathrm{d}\nu^{(D)}_{AB}(x) = \prod_{i=1}^{r} \left(\sum_{\sigma \in NC(n)} m_{\gamma_{A,i},\sigma} m_{\gamma_{B,i},\sigma} m_{\gamma_{A,i},\sigma^{-1}\tau_n} m_{\gamma_{B,i},\sigma^{-1}\tau_n} \right).$$

Thus, by the method of moments and Theorem 13.2, it holds that $\nu^{(D)}_{AB}$ converges weakly, in probability, to

$$\nu_{AB} = \begin{cases} \bigotimes_{i=1}^{r} \nu_i & \text{if } r > 0, \\ \mu_{\gamma_C} & \text{if } r = 0, \end{cases} \tag{13.33}$$

where ν_i is given by

$$\nu_i = (\mu_{\gamma_{A,i}} \otimes \mu_{\gamma_{B,i}})^{\boxtimes 2} \boxtimes \mathrm{MP}(1).$$

This reduces to Eq. (13.29) if the link states are maximally entangled. We can use this to compute the logarithmic negativity, as we did previously. For $r > 0$,

$$E_N(\rho_{AB}) = \log \int \sqrt{\lambda} \mathrm{d}\nu^{(D)}_{AB}(\lambda) + \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|) + \frac{1}{2}\log m_{\gamma_C,2} - \log \mathrm{tr}[\rho],$$

from which we find that $E_N(\rho_{AB}) - \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|)$ converges in probability to

$$\log \int \sqrt{\lambda} \mathrm{d}\nu_{AB}(\lambda) + \frac{1}{2}\log m_{\gamma_{C,2}}.$$

For the case $|\gamma_A| + |\gamma_B| = |\gamma_C|$, it is more elegant to use the limiting distribution of $\mu_{AB}^{(D)}$, as defined in Eq. (13.25). By the method of moments and Eq. (13.32), $\mu_{AB}^{(D)} \Rightarrow \mu_{\gamma_C}$, in probability. We may then compute the entanglement negativity as

$$E_N(\rho_{AB}) = \log \int |\lambda| \mathrm{d}\mu_{AB}^{(D)}(\lambda) + \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|) - \log \mathrm{tr}[\rho],$$

and hence $E_N(\rho_{AB}) - \frac{\log D}{2}(|\gamma_A| + |\gamma_B| - |\gamma_C|)$ converges in probability to $\log \int |\lambda| \mathrm{d}\mu_{AB}(\lambda)$.

Discussion and open questions

In this chapter we restricted to link states with bounded spectral variation. In the next chapter we will investigate more general link states. Another restriction we made is that we only considered the case with at most two minimal cuts. It would be interesting, already in the maximally entangled case, to characterize the convergence of appropriately scaled empirical distributions for any possible collection of minimal cuts. In this case, the same arguments would allow one to show that one can again use the replica trick to compute the moments of the reduced density matrix. The dominant configurations of permutations will be given by configurations which are constants on the connected components of the graph upon removing the minimal cuts. However, in this case the combinatorial problem of counting the relevant permutations will be more complicated.

Another avenue of investigation is to study more fine-grained properties of the spectrum. For instance, in the context of random matrix theory one typically also studies the behaviour of the largest and smallest eigenvalues of the random matrix.

Link states with unbounded spectral variation

We will now consider a different regime, where the link states have unbounded spectral variation. Our methods in this chapter are distinct from the previous one, and the two chapters can be considered separately.

14.1 One-shot entropies

We begin by introducing one of our main tools for studying entanglement spectra in random tensor network states: *one-shot entropies*. In quantum information theory, the rates of certain important protocols, such as compression or state merging can be expressed as entropic quantities. One-shot entropies are the appropriate analogs for settings where one would like to analyze a task for a single or finite number of copies of the relevant state. Asymptotic rates in terms of ordinary von Neumann entropies are then recovered in the limit of infinitely many independent copies. For an extensive introduction to this point of view, see [Tom15]; here we provide the basic definitions and introduce the relevant concepts.

A random tensor network built from link states that are maximally entangled (or more generally have bounded spectral variation) effectively can be analyzed using asymptotic tools. Indeed, if we have a maximally entangled state of exponentially large dimension $D = 2^n$, then this is equal to the n-th tensor power of a qubit maximally entangled state, so we are effectively in a asymptotic situation. However, if we allow for link states with unbounded spectral variation or even completely general background states, as in Section 12.2, then it is more natural to use tools from one-shot quantum information theory.

We take the Rényi entropies as a starting point, which we defined in Eq. (12.9) for subnormalized states. Let $\mathcal{H}$ be some Hilbert space and for $\rho \in \mathcal{P}_\leq(\mathcal{H})$ we define the (unconditional) *min-entropy* and the *max-entropy* by

$$H_{\min}(\rho) = -\log \|\rho\|_\infty$$
$$H_{\max}(\rho) = \log\left(\mathrm{tr}[\sqrt{\rho}]^2\right)$$

which coincide with the Rényi entropies $H_\infty(\rho)$ and $H_{\frac{1}{2}}(\rho)$ for $\rho \in \mathcal{P}_=(\mathcal{H})$. As usual, if ρ_A is the reduced density matrix on a system A, we may write $H_{\min}(A)_\rho = H_{\min}(\rho_A)$ and $H_{\max}(A)_\rho = H_{\max}(\rho_A)$.

Often, when applied to study quantum information processing tasks, it is useful to allow a small error. This leads to the introduction of smooth entropies. To this end we define the trace distance between $\rho, \sigma \in \mathscr{P}_{\leq}(\mathscr{H})$ to be

$$T(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1 + \frac{1}{2}|\mathrm{tr}[\rho - \sigma]|,$$

where the last term, which is absent in usual definitions of the trace distance, accounts for subnormalized states. It is easy to see that $T(\rho, \sigma) \leq \|\rho - \sigma\|_1 \leq 2T(\rho, \sigma)$. We define the *smooth min- and max-entropies* of $\rho \in \mathscr{P}_{\leq}(\mathscr{H})$ as

$$H_{\min}^{\varepsilon}(\rho) = \sup_{\rho^{\varepsilon} \in \mathscr{P}_{\leq}(\mathscr{H}), T(\rho^{\varepsilon}, \rho) \leq \varepsilon} H_{\min}(\rho^{\varepsilon})$$

$$H_{\max}^{\varepsilon}(\rho) = \sup_{\rho^{\varepsilon} \in \mathscr{P}_{\leq}(\mathscr{H}), T(\rho^{\varepsilon}, \rho) \leq \varepsilon} H_{\max}(\rho^{\varepsilon}).$$

The smooth entropies are such that one recovers the usual von Neumann entropies in the limit of many independent copies. Indeed, the following *asymptotic equipartition property* holds:

$$\lim_{n \to \infty} \frac{1}{n} H_{\min}^{\varepsilon}(\rho^{\otimes n}) = H(\rho) = \lim_{n \to \infty} \frac{1}{n} H_{\max}^{\varepsilon}(\rho^{\otimes n})$$

for any $0 < \varepsilon < 1$. Variations on this definition are possible. For instance, one can choose a different distance measure, which will yield different entropies. However, for the usual choices, the differences go to zero as ε goes to zero, so the particular choice is often immaterial. One useful distance measure is the purified distance, which is given for $\rho, \sigma \in \mathscr{P}_{\leq}(\mathscr{H})$ by

$$P(\rho, \sigma) = \sqrt{1 - F_*(\rho, \sigma)^2}$$

where $F_*(\rho, \sigma)$ is the *generalized fidelity* between ρ and σ, which is defined by

$$F_*(\rho, \sigma) = F(\rho, \sigma) + \sqrt{(1 - \mathrm{tr}[\rho])(1 - \mathrm{tr}[\sigma])}$$

in terms of the ordinary *fidelity* $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$. The Fuchs-van de Graaff inequalities relate the trace distance and purified distance:

$$T(\rho, \sigma) \leq P(\rho, \sigma) \leq \sqrt{2T(\rho, \sigma)} \tag{14.1}$$

for $\rho, \sigma \in \mathscr{P}_{\leq}(\mathscr{H})$.

There are also conditional versions of the Rényi entropies. Consider a bipartite quantum state $\rho_{AB} \in \mathscr{P}_{=}(AB)$. For the von Neumann entropy, the conditional entropy can simply be defined as $H(A|B)_{\rho} = H(AB)_{\rho} - H(B)_{\rho}$. However, it turns out that this is not a good definition in the Rényi case. There are various ways to define a Rényi conditional entropy $H_k(A|B)$; we use a version based on the so-called sandwiched Rényi relative entropy. For $k = 2$, this gives a *quantum conditional collision entropy*, which will be useful for defining minimal cuts and which is defined as follows. For $\rho_{AB} \in \mathscr{P}_{\leq}(AB)$, let

$$H_2(A|B)_{\rho|\rho} := -\log \mathrm{tr}\left[\left((I \otimes \rho_B)^{-\frac{1}{4}} \rho_{AB} (I \otimes \rho_B)^{-\frac{1}{4}}\right)^2\right] + \log \mathrm{tr}[\rho]. \tag{14.2}$$

Finally, there are also conditional versions of the min- and max-entropy. If $\rho_{AB} \in \mathscr{P}_{\leq}(AB)$ and $\sigma_B \in \mathscr{P}_{\leq}(B)$, we define

$$H_{\min}(A|B)_{\rho|\sigma} = -\inf\{\lambda : \rho_{AB} \leq 2^\lambda I_A \otimes \sigma_B\}$$
$$H_{\max}(A|B)_{\rho|\sigma} = \log\|\rho_{AB}(I \otimes \sigma_B)\|_1^2$$

and we let

$$H_{\min}(A|B)_\rho = \sup_{\sigma \in \mathscr{P}_{\leq}(B)} H_{\min}(A|B)_{\rho|\sigma}$$
$$H_{\max}(A|B)_\rho = \sup_{\sigma \in \mathscr{P}_{\leq}(B)} H_{\max}(A|B)_{\rho|\sigma}.$$

There is a duality between max- and min-entropies. If $\rho \in \mathscr{P}_{\leq}(ABC)$ is a *pure* state, it holds that

$$H_{\min}(A|B)_\rho = -H_{\max}(A|C)_\rho. \tag{14.3}$$

We will use the fact that for a normalized state $\rho_{AB} \in \mathscr{P}_{=}(AB)$ (Corollary 5.10 in [Tom15])

$$H_{\min}(A|B)_{\rho|\rho} \leq H_{\min}(A|B)_\rho \leq H_2(A|B)_{\rho|\rho}. \tag{14.4}$$

14.2 Recovery isometries

Recall that we study random tensor network states with *link states*, which are pure states which are a tensor product of edge states $\phi = \bigotimes_{e \in E} \phi_e \in \mathscr{P}_{=}(V)$ for some graph $G = (V, E)$, and in Section 12.2 we considered more general *background states* $\phi_V \in \mathscr{P}_{\leq}(V)$, where we no longer have a tensor product structure along the edges of some graph, and applying the replica trick does not yield a local spin model for the moments of the tensor network state. This situation is of independent interest, but will also be useful as an intermediate step when applying bounds based on one-shot entropies to link states. In Chapter 13, we studied link states for which the entanglement spectrum of the edge states ϕ_e had bounded variation, and we used the replica trick to compute the moments of the spectrum of ρ_A for a boundary subsystem A. For general background states we saw that the replica trick for $k = 2$ extends as in Eq. (12.21). What are the minimal cuts in this setting? Based on Eq. (12.21) a first guess would be that $\Gamma_A \in C(A)$ would be a minimal cut (i.e. correspond to the dominant term in the replica trick) if for all other cuts $\Delta_A \in C(A)$ we would have $H_2(\Gamma_A)_\phi \ll H_2(\Delta_A)_\phi$. If the state is a link state, this corresponds to adding weights to the edges of the graph corresponding to the Rényi-2 entropies along the edges, and computing a weighted minimal cut. Indeed, this would yield an accurate approximation of $\text{tr}[\rho_A^2]$ and hence of $H_2(\rho_A)$. However, if the spectrum of ρ_A is not close to a flat spectrum, this does not imply that $\text{spec}_+(\rho_A)$ is close to $\text{spec}_+(\phi_{\Gamma_A})$. We would like to show that for link states with unbounded spectral variation, and an appropriate minimal cut condition for $\Gamma_A \in C(A)$, it is still true that $\text{spec}_+(\rho_A)$ is close to $\text{spec}_+(\phi_{\Gamma_A})$.

We will adapt the $k = 2$ replica trick for general background states to get a bound on the difference in trace norm between $\text{spec}_+(\rho_A)$ and $\text{spec}_+(\phi_{\Gamma_A})$ in terms of *conditional* Rényi-2 entropies,[1] as defined in Eq. (14.2). In Section 14.2.1, we will use this to formulate a condition for cut minimality in terms of smooth entropies for link states.

---

[1] Note that while $H(A|B)_\phi = H(AB)_\phi - H(B)_\phi$, in general $H_2(A|B)_{\phi|\phi} \neq H_2(AB)_\phi - H_2(B)_\phi$.

The main result of this subsection is a tensor network version of *one-shot decoupling*. Let $\phi_V \in \mathscr{P}_\le(V)$. We allow ϕ_V to be a general state, which need not be pure and also need not be a product state along the edges of some graph. Let R be a purifying system and $\phi_{VR} \in \mathscr{P}_\le(VR)$ be a purification of ϕ_V. Then we can construct the random tensor network state $\rho_{V_\partial R}$ where the boundary systems are given by $V_\partial \cup R$, which is a purification of the random tensor network state ρ_{V_∂} as in Eq. (12.19) by

$$\phi_{V_\partial R} = \mathrm{tr}_{V_b}[(I_{V_\partial R} \otimes \psi)\phi] \tag{14.5}$$

where ψ is a tensor product of random tensors. We briefly recall our notation for boundary subsystems and cuts: for a boundary subsystem $A \subseteq V_\partial$, we denote its boundary complement by $\bar{A} = V_\partial \setminus A$, and for a cut $\Gamma_A \in C(A)$, we let $\Gamma_A^c = V \setminus \Gamma_A$, which is a cut for $\bar{A}$. The purifying system R can be thought of as an additional boundary system in the tensor network construction.

In Theorem 14.4, we will assume that we have a cut $\Gamma_A \in C(A)$ which is such that for all cuts $\Delta_A \in C(A)$ for which $\Delta_A \subsetneq \Gamma_A$ we have $H_2(\Gamma_A \setminus \Delta_A | \Gamma_A^c R)_{\phi|\phi} \gg 1$, and similarly for all cuts $\Delta_A \in C(A)$ for which $\Gamma_A \subsetneq \Delta_A$ we have $H_2(\Delta_A \setminus \Gamma_A | \Gamma_A R)_{\phi|\phi} \gg 1$. We sow that this condition implies that with high probability there exist isometries $V_A : \mathscr{H}_A \to \mathscr{H}_{\Gamma_A}$ and $V_{\bar{A}} : \mathscr{H}_{\bar{A}} \to \mathscr{H}_{\Gamma_A^c}$ such that

$$(V_A \otimes V_{\bar{A}} \otimes I_R)|\rho\rangle \approx |\phi\rangle. \tag{14.6}$$

The approximation accuracy will be measured in trace norm. In particular, this implies that $\mathrm{spec}_+(\rho_A) \approx \mathrm{spec}_+(\phi_{\Gamma_A})$. If the state ϕ is a link state, $\mathrm{spec}_+(\phi_{\Gamma_A})$ is precisely the entanglement spectrum along the cut γ_A. The isometries V_A and $V_{\bar{A}}$ are *recovery isometries*, which allow us to 'recover' Γ_A from the A system, and similarly we can recover Γ_A^c from $\bar{A}$.

The result is closely related to quantum error correction. One way to interpret this is is as follows: consider a subspace $\mathscr{H}_S$ of $\mathscr{H}_V$ and let R be a reference system of dimension $\dim(\mathscr{H}_S)$, and ϕ_{VR} a maximally entangled state between S and R. Then Eq. (14.6) can be interpreted as saying that if we encode the subspace S by projecting onto random tensors, the information in Γ_A is protected, after encoding, against an erasure error on $\bar{A}$. This idea is also discussed in [PYHP15] for perfect tensor network models, and in [HNQ$^+$16] for random tensor networks with maximally entangled link states. In holography, the notion of local recovery isometries and their error correction interpretation goes under the name of *entanglement wedge reconstruction* or *subregion-subregion duality*. See [AR19, AP22] for a detailed discussion of entanglement wedge reconstruction in holographic systems with bulk entropy, relating to one-shot entropies. We provide more details in Section 14.3.

Our approach to showing Eq. (14.6) is that we start by projecting only on the random tensors in Γ_A, and not on the random tensors in Γ_A^c. This yields a random tensor network state σ on $A\Gamma_A^c R$.

We then show that, by a version of one-shot decoupling, the reduced state on $\sigma_{\Gamma_A^c R}$ has not changed much from $\phi_{\Gamma_A^c R}$. By Uhlmann's theorem, this implies that there exists an isometry V_A such that $(V_A \otimes I_{\Gamma_A^c R})|\sigma\rangle \approx |\phi\rangle$. Combining this with a similar result for Γ_A^c we obtain Eq. (14.6), as will be made precise in Theorem 14.4.

In our construction of σ, we can relabel the vertices in the graph, and think of the vertices in $\Gamma_A \setminus A$ as the bulk vertices V_b, the boundary subsystem A as the complete

boundary V_∂, and relabel all other subsystems as the reference system R. Then we prove the following result, which is closely related to the one-shot decoupling results in [DBWR14].

Proposition 14.1. *Consider a random tensor network state $\rho_{V_\partial R}$ as in Eq. (14.5) with a (purified) background state $\phi_{VR} \in \mathscr{P}_\le(VR)$. Let $A = V_\partial$ and let $\Gamma_A = V$ and suppose that for any cut $\Delta_A \in C(A)$ other than Γ_A*

$$H_2(\Gamma_A \setminus \Delta_A | R)_{\phi|\phi} \ge K$$

then

$$\mathbb{E}\|\rho_R - \phi_R\|_1 \le 2^{\frac{|V_b|}{2}} \sqrt{\mathrm{tr}[\phi]} 2^{-\frac{1}{2}K}.$$

Note that, since $\Gamma_A = V_b \cup A$, the sets $\Gamma_A \setminus \Delta_A$ for $\Delta_A \in C(A) \setminus \{\Gamma_A\}$ are exactly the non-empty subsets of V_b. The formulation in terms of $\Delta_A \in C(A)$ will be natural when we apply this result in Theorem 14.4.

Proof. We closely follow the strategy in [DBWR14, DH10]. We first note a basic fact (Lemma 3.7 in [DBWR14]): for any operator X and ω a subnormalized density matrix, it holds that

$$\|X\|_1 \le \|\omega^{-\frac{1}{4}} X \omega^{-\frac{1}{4}}\|_2. \tag{14.7}$$

The proof is an application of the Cauchy-Schwarz inequality. We use (14.7) with $\omega = \phi_R$ and Jensen's inequality to see that

$$\mathbb{E}\|\rho_R - \phi_R\|_1 \le \sqrt{\mathbb{E}\,\mathrm{tr}[(\tilde\rho_R - \tilde\phi_R)^2]}$$

where $\tilde\rho_{V_\partial R} = (I \otimes \phi_R)^{-\frac{1}{4}} \rho_{V_\partial R}(I \otimes \phi_R)^{-\frac{1}{4}}$ and $\tilde\phi_{VR} = (I \otimes \phi_R)^{-\frac{1}{4}} \phi_{VR}(I \otimes \phi_R)^{-\frac{1}{4}}$. By Eq. (12.22) we have $\mathbb{E}\tilde\rho_R = \tilde\phi_R$, and the replica trick in Eq. (12.21) yields

$$\mathbb{E}\,\mathrm{tr}[(\tilde\rho_R - \tilde\phi_R)^2] = \mathbb{E}\,\mathrm{tr}[\tilde\rho_R^2] - \mathrm{tr}[\tilde\phi_R^2]$$

$$= \sum_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} \mathrm{tr}[\tilde\phi_{(\Gamma_A \setminus \Delta_A)R}^2]$$

$$= \sum_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} \mathrm{tr}[\phi] 2^{-H_2(\Gamma_A \setminus \Delta_A | R)_{\phi|\phi}}$$

using the definition of $\tilde\phi$ and Eq. (14.2) and hence

$$\left(\mathbb{E}\|\rho_R - \phi_R\|_1\right)^2 \le 2^{|V_b|} \mathrm{tr}[\phi] 2^{-K}.$$

$\blacksquare$

Suppose that in the set-up of Proposition 14.1, we would have equality $\rho_R = \phi_R$. Then, by Uhlmann's theorem, their purifications ρ_{AR} and $\phi_{\Gamma_A R}$ are related by some isometry V_A from A to Γ_A. The following lemma is useful to extend to the case where the reduced states are close in trace distance.

Lemma 14.2. *Suppose $\rho_{AB} \in \mathscr{P}(AB)$ and $\sigma_{AC} \in \mathscr{P}_\le(AC)$ are pure states on bipartite Hilbert spaces $\mathscr{H}_A \otimes \mathscr{H}_B$ and $\mathscr{H}_A \otimes \mathscr{H}_C$ respectively. Then*

$$\min_V \|(I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger) - \sigma_{AC}\|_1 \le 2\sqrt{2\|\rho_A - \sigma_A\|_1 + 2\|\rho_A - \sigma_A\|_1^2}.$$

where the minimum is over all isometries $V : \mathscr{H}_B \to \mathscr{H}_C$.

Proof. Uhlmann's theorem states that if $\rho_{AB} \in \mathscr{P}(AB)$ and $\sigma_{AC} \in \mathscr{P}(AC)$ are pure quantum states with $\dim(\mathscr{H}_B) \leq \dim(\mathscr{H}_C)$, then there exists an isometry $V : \mathscr{H}_B \to \mathscr{H}_C$ such that

$$P(\rho_A, \sigma_A) = P((I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger), \sigma_{AC}) = P(\rho_A, \sigma_A)$$

and, in particular, the isometry is the solution to an optimization problem:

$$P(\rho_A, \sigma_A) = \min_V P((I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger), \sigma_{AC}).$$

Moreover, if both ρ and σ are subnormalized, by Eq. (14.1), we can bound

$$\min_V \|(I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger) - \sigma_{AC}\|_1 \leq \min_V 2P((I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger), \sigma_{AC})$$
$$= 2P(\rho_A, \sigma_A) \leq 2\sqrt{2\|\rho_A - \sigma_A\|_1}.$$

From this it follows that if σ is subnormalized and ρ has $\mathrm{tr}[\rho] > 1$,

$$\min_V \|(I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger) - \sigma_{AC}\|_1 \leq 2\sqrt{2\,\mathrm{tr}[\rho]\|\rho_A - \sigma_A\|_1}.$$

Since $\mathrm{tr}[\rho] \leq \mathrm{tr}[\sigma] + \|\rho - \sigma\|_1$ and $\mathrm{tr}[\sigma] \leq 1$ we conclude that

$$\min_V \|(I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger) - \sigma_{AC}\|_1 \leq 2\sqrt{2\|\rho_A - \sigma_A\|_1 + 2\|\rho_A - \sigma_A\|_1^2}.$$

for arbitrary ρ and subnormalized σ. $\blacksquare$

Finally, we will need a basic lemma relating tensor network states with differing background states:

Lemma 14.3. *Suppose we consider random tensor network states $\rho_{V_\partial R}$ and $\tilde{\rho}_{V_\partial R}$ with (purified) background states $\phi_{VR}, \tilde{\phi}_{VR} \in \mathscr{P}_\leq(VR)$ and projecting onto the same random tensors. Then*

$$\mathbb{E}\|\rho_{V_\partial R} - \tilde{\rho}_{V_\partial}R\|_1 \leq \|\phi_{VR} - \tilde{\phi}_{VR}\|_1.$$

Proof. Let $\phi - \tilde{\phi} = \Delta_+ - \Delta_-$ where both Δ_+ and Δ_- are positive semidefinite and are such that $\|\phi - \tilde{\phi}\|_1 = \mathrm{tr}[\Delta_+ + \Delta_-]$. Then we can also consider the random tensor network states σ_+ and σ_- which take Δ_+ and Δ_- as background states, and by the linearity of Eq. (14.5) in the background state we have $\rho - \tilde{\rho} = \sigma_+ - \sigma_-$. By Eq. (12.22), $\mathbb{E}\sigma_\pm = \Delta_\pm$. We then estimate

$$\mathbb{E}\|\rho - \tilde{\rho}\|_1 = \mathbb{E}\|\sigma_+ - \sigma_-\|_1 \leq \mathbb{E}(\|\sigma_+\|_1 + \|\sigma_-\|_1)$$
$$= \mathbb{E}\,\mathrm{tr}[\sigma_+ + \sigma_-] = \mathrm{tr}[\Delta_+ + \Delta_-] = \|\phi - \tilde{\phi}\|_1.$$

where we have used that σ_+ and σ_- are positive semidefinite and hence $\|\sigma_\pm\|_1 = \mathrm{tr}[\sigma_\pm]$. $\blacksquare$

With all our tools assembled, we are ready to prove the main result of this subsection. We again let $\phi_{VR} \in \mathscr{P}_\leq(VR)$ be a background state with R be a purifying system, and we let $\rho_{V_\partial R}$ be the associated random tensor network state as constructed in Eq. (14.5). Let Γ_A be an arbitrary cut for the boundary region A. In Theorem 14.4, we provide a

criterion to determine whether Γ_A is a minimal cut in terms of conditional entropies. Informally speaking, the following result shows that if Γ_A is a minimal cut in this sense, we can recover the subsystem Γ_A from the boundary subsystem A, while we can recover Γ_A^c from the boundary subsystem $\bar{A}$. For general ϕ, Theorem 14.4 is closely related to the task of *split transfer*, see Section 14.3 for a discussion. The following result closely follows Proposition 18 of [DH10].

Theorem 14.4 (Recovery isometries)**.** *Let $\phi_{VR} \in \mathscr{P}_{\le}(VR)$ and let $\rho_{V_\partial R}$ be the associated random tensor network state as in Eq. (14.5). Let $\Gamma_A \in C(A)$ and suppose that*

$$H_2(\Gamma_A \setminus \Delta_A | \Gamma_A^c R)_{\phi|\phi} \ge K_1 \tag{14.8}$$

for all cuts $\Delta_A \in C(A)$ such that $\Delta_A \subsetneq \Gamma_A$ and

$$H_2(\Delta_A \setminus \Gamma_A | \Gamma_A R)_{\phi|\phi} \ge K_2 \tag{14.9}$$

for all cuts $\Delta_A \in C(A)$ such that $\Gamma_A \subsetneq \Delta_A$. Then

$$\mathbb{E} \min_{V_A, V_{\bar{A}}} \|(V_A \otimes V_{\bar{A}} \otimes I_R) \rho_{V_\partial R} (V_A^\dagger \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1 = \mathcal{O}(\mathrm{tr}[\phi]^{\frac{1}{4}} (2^{-\frac{1}{4}K_1} + 2^{-\frac{1}{4}K_2})). \tag{14.10}$$

where the minimum is over isometries $V_A : \mathcal{H}_A \to \mathcal{H}_{\Gamma_A}$ and $V_{\bar{A}} : \mathcal{H}_{\bar{A}} \to \mathcal{H}_{\Gamma_A^c}$.

Proof. Let $\sigma_{A\Gamma_A^c R}$ be the state where we have contracted along the tensors in Γ_A but not along those in Γ_A^c, and similarly let $\tau_{\bar{A}\Gamma_A R}$ be the state where we have contracted along the tensors in Γ_A^c but not along those in Γ_A. We first use Proposition 14.1 to show that $\sigma_{\Gamma_A^c R} \approx \phi_{\Gamma_A^c R}$ and $\tau_{\Gamma_A R} \approx \phi_{\Gamma_A R}$. Indeed, for σ we simply apply Proposition 14.1 with $V_b \cap \Gamma_A$ as the set of bulk vertices, A as the set of boundary vertices and $\Gamma_A^c R$ as the reference system. This gives

$$\mathbb{E}\|\sigma_{\Gamma_A^c R} - \phi_{\Gamma_A^c R}\|_1 = \mathcal{O}(\sqrt{\mathrm{tr}[\phi]} 2^{-\frac{1}{2}K_1}).$$

A similar application of Proposition 14.1, with $V_b \cap \Gamma_A^c$ as the set of bulk vertices, $\bar{A}$ as the set of boundary vertices and $\Gamma_A R$ as the reference system, shows

$$\mathbb{E}\|\tau_{\Gamma_A R} - \phi_{\Gamma_A R}\|_1 = \mathcal{O}(\sqrt{\mathrm{tr}[\phi]} 2^{-\frac{1}{2}K_2}).$$

We note that for any isometries $V_A : \mathcal{H}_A \to \mathcal{H}_{\Gamma_A}$ and $V_{\bar{A}} : \mathcal{H}_{\bar{A}} \to \mathcal{H}_{\Gamma_A^c}$

$$\|(V_A \otimes V_{\bar{A}} \otimes I_R)\rho(V_A^\dagger \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1 \le \|(I_{\Gamma_A} \otimes V_{\bar{A}} \otimes I_R)\tau(I_{\Gamma_A} \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1$$
$$+ \|\tau - (V_A \otimes I_{\bar{A}R})\rho(V_A^\dagger \otimes I_{\bar{A}R})\|_1,$$

where we have added and subtracted $(I_{\Gamma_A} \otimes V_{\bar{A}} \otimes I_R)\tau(I_{\Gamma_A} \otimes V_{\bar{A}}^\dagger \otimes I_R)$, applied the triangle inequality and then used the invariance of the trace norm under isometries in the second term. We use this to estimate

$$\mathbb{E} \min_{V_A, V_{\bar{A}}} \|(V_A \otimes V_{\bar{A}} \otimes I_R)\rho(V_A^\dagger \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1$$
$$\le \mathbb{E}\big(\min_{V_{\bar{A}}} \|(I_{\Gamma_A} \otimes V_{\bar{A}} \otimes I_R)\tau(I_{\Gamma_A} \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1 \tag{14.11}$$
$$+ \min_{V_A} \|\tau - (V_A \otimes I_{\bar{A}R})\rho(V_A^\dagger \otimes I_{\bar{A}R})\|_1\big),$$

where the minimum is over isometries $V_A : \mathscr{H}_A \to \mathscr{H}_{\Gamma_A}$ and $V_{\bar{A}} : \mathscr{H}_{\bar{A}} \to \mathscr{H}_{\Gamma_A^c}$. For the first term of Eq. (14.11), we apply Lemma 14.2 to get

$$\min_{V_{\bar{A}}} \|(I_{\Gamma_A} \otimes V_{\bar{A}} \otimes I_R)\tau(I_{\Gamma_A} \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1 \leq 2\sqrt{2\|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1 + \|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1^2}$$

$$\leq 2\sqrt{2}\left(\sqrt{\|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1} + \|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1\right)$$

and by Jensen's inequality

$$\mathbb{E}\min_{V_{\bar{A}}} \|(I_{\Gamma_A} \otimes V_{\bar{A}} \otimes I_R)\tau(I_{\Gamma_A} \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1$$

$$\leq 2\sqrt{2}\left(\sqrt{\mathbb{E}\|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1} + \mathbb{E}\|\tau_{\Gamma_A R} - \phi_{\Gamma_{AR}}\|_1\right) \tag{14.12}$$

$$= \mathcal{O}(\mathrm{tr}[\phi]^{\frac{1}{4}} 2^{-\frac{1}{4}K_2})$$

For the second term of Eq. (14.11), we can think of τ and $(V_A \otimes I_{\bar{A}R})\rho(V_A^\dagger \otimes I_{\bar{A}R})$ as the random tensor network states with ϕ and $(V_A \otimes I_{\Gamma_A^c R})\sigma(V_A^\dagger \otimes I_{\Gamma_A^c R})$ as link states, applying random tensors in Γ_A^c. Then, denoting by $\mathbb{E}_{\Gamma_A^c}$ the expectation value over all random tensors in Γ_A^c, by Lemma 14.3

$$\mathbb{E}_{\Gamma_A^c}\|\tau - (V_A \otimes I_{\bar{A}R})\rho(V_A^\dagger \otimes I_{\bar{A}R})\|_1 = \mathcal{O}(\|\phi_{VR} - (V_A \otimes I_{\bar{A}R})\sigma(V_A^\dagger \otimes I_{\bar{A}R})\|_1).$$

We thus estimate

$$\mathbb{E}\min_{V_A}\|\tau - (V_A \otimes I_{\bar{A}R})\rho(V_A^\dagger \otimes I_{\bar{A}R})\|_1 = \mathcal{O}(\mathbb{E}_{\Gamma_A}\min_{V_A}\|\phi_{VR} - (V_A \otimes I_{\bar{A}R})\sigma(V_A^\dagger \otimes I_{\bar{A}R})\|_1)$$

for which we may argue exactly as in Eq. (14.12) and using Lemma 14.2 that

$$\mathbb{E}_{\Gamma_A}\min_{V_A}\|\phi_{VR} - (V_A \otimes I_{\Gamma_A^c R})\sigma(V_A^\dagger \otimes I_{\Gamma_A^c R})\|_1 = \mathcal{O}(\mathrm{tr}[\phi]^{\frac{1}{4}} 2^{-\frac{1}{4}K_1}).$$

We conclude that

$$\mathbb{E}\min_{V_A, V_{\bar{A}}}\|(V_A \otimes V_{\bar{A}} \otimes I_R)\rho(V_A^\dagger \otimes V_{\bar{A}}^\dagger \otimes I_R) - \phi_{VR}\|_1 = \mathcal{O}(\mathrm{tr}[\phi]^{\frac{1}{4}}(2^{-\frac{1}{4}K_1} + 2^{-\frac{1}{4}K_2})).$$

■

We hence find that the closeness of the boundary and background state can be bounded via conditional Rényi-2 entropies of cuts. In particular, for large K_1, K_2, the recovery isometries can recover states to good accuracy, and $\mathbb{E}\|\mathrm{spec}_+(\rho_A) - \mathrm{spec}_+(\phi_{\Gamma_A})\|$ is small. However, this result is not yet completely satisfying. Indeed, the conditional Rényi-2 entropy is not a 'robust' quantity, in the sense that a small deformation of ϕ can drastically change the values of the conditional Rényi-2 entropies in Eq. (14.8) and Eq. (14.9). For this reason, we would like a condition with *smoothed* entropies. We first note that one can actually show that for the condition in Eq. (14.8), we can bound

$$H_2(\Gamma_A \setminus \Delta_A | \Gamma_A^c R)_{\phi|\phi} \geq H_{\min}(\Gamma_A \setminus \Delta_A | \Gamma_A^c R)_\phi \tag{14.13}$$

and similarly for Eq. (14.9),

$$H_2(\Delta_A \setminus \Gamma_A | \Gamma_A R)_{\phi|\phi} \geq H_{\min}(\Delta_A \setminus \Gamma_A | \Gamma_A R)_\phi \tag{14.14}$$

To make the condition 'robust', we would like to replace these by smoothed entropies and express a condition in terms of $H^\varepsilon_{\min}(\Gamma_A \setminus \Delta_A | \Gamma^c_A R)_\phi$ and $H^\varepsilon_{\min}(\Delta_A \setminus \Gamma_A | \Gamma_A R)_\phi$. This will require *simultaneous smoothing*: finding a state $\phi^\varepsilon_{VR} \in \mathscr{P}_\le(VR)$ which is close to ϕ, such that

$$H_{\min}(\Gamma_A \setminus \Delta_A | \Gamma^c_A R)_{\phi^\varepsilon} \ge H^\varepsilon_{\min}(\Gamma_A \setminus \Delta_A | \Gamma^c_A R)_\phi$$

or

$$H_{\min}(\Delta_A \setminus \Gamma_A | \Gamma_A R)_{\phi^\varepsilon} \ge H^\varepsilon_{\min}(\Delta_A \setminus \Gamma_A | \Gamma_A R)_\phi$$

for all relevant cuts Δ_A. If we have a general background state, it is not known how this can be done [DF13, Dut11]. However, if the background state is actually a link state, we can bound the smooth conditional entropies in terms of non-conditional smooth one-shot entropies, for which we can perform the simultaneous smoothing.

14.2.1 Approximation of the boundary spectrum

The primary result in this subsection is Theorem 14.6, which states that if the background state is actually a link state as in Eq. (12.4), then the spectrum of the boundary state is well-approximated by the spectrum of the minimal cut link state in expectation, where the approximation accuracy is controlled by *smooth* one-shot entropies. It is a straightforward application of Theorem 14.4. For a link state the, (unsmoothed) condition in Eq. (14.13) and Eq. (14.13) boils down to the difference between the min-entropy along δ_A and the max-entropy along γ_A. In fact, we show a slightly stronger statement that we can ignore the intersection $\gamma_A \cap \delta_A$.

We will now introduce notation in order to state Theorem 14.6. Let Δ_A and Γ_A be cuts with cut-sets δ_A and γ_A. We want to quantify the sense in which the cut along γ_A is 'smaller' than the one along δ_A, by showing that the smooth max-entropy along the edge set $\gamma_A \setminus (\delta_A \cap \gamma_A)$ is smaller than the smooth min-entropy along $\delta_A \setminus (\delta_A \cap \gamma_A)$. We ignore the intersection of γ_A and δ_A intersect. To this end, we fix a cut Γ_A, and we define for any cut $\Delta_A \in C(A)$ for which $\Delta_A \subsetneq \Gamma_A$, the sets of half-edges

$$C^{\Delta_A}_{\Gamma_A} = \{(e, x) : e = (xy), x \in \Gamma_A \setminus \Delta_A, y \in \Gamma^c_A\}$$
$$D^{\Delta_A}_{\Gamma_A} = \{(e, x) : e = (xy), x \in \Gamma_A \setminus \Delta_A, y \in \Delta_A\},$$

and similarly for any cut $\Delta_A \in C(A)$ for which $\Gamma_A \subsetneq \Delta_A$

$$C^{\Delta_A}_{\Gamma_A} = \{(e, x) : e = (xy), x \in \Delta_A \setminus \Gamma_A, y \in \Gamma_A\}$$
$$D^{\Delta_A}_{\Gamma_A} = \{(e, x) : e = (xy), x \in \Delta_A \setminus \Gamma_A, y \in \Delta^c_A\}.$$

We let

$$\mathscr{C}(\Gamma_A) = \{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A\} \cup \{\Delta_A \in C(A), \Gamma_A \subsetneq \Delta_A\}$$

The sets $C^{\Delta_A}_{\Gamma_A}$ and $D^{\Delta_A}_{\Gamma_A}$ are sets of half-edges along γ_A and δ_A respectively (leaving out the intersection $\gamma_A \cap \delta_A$). This is illustrated in Fig. 14.1. Then we let

$$h^\varepsilon_{\Gamma_A}(\Delta_A) := H^\varepsilon_{\min}(D^{\Delta_A}_{\Gamma_A})_\phi - H^\varepsilon_{\max}(C^{\Delta_A}_{\Gamma_A})_\phi.$$

We now define the condition for Γ_A to be a generalized minimal cut:
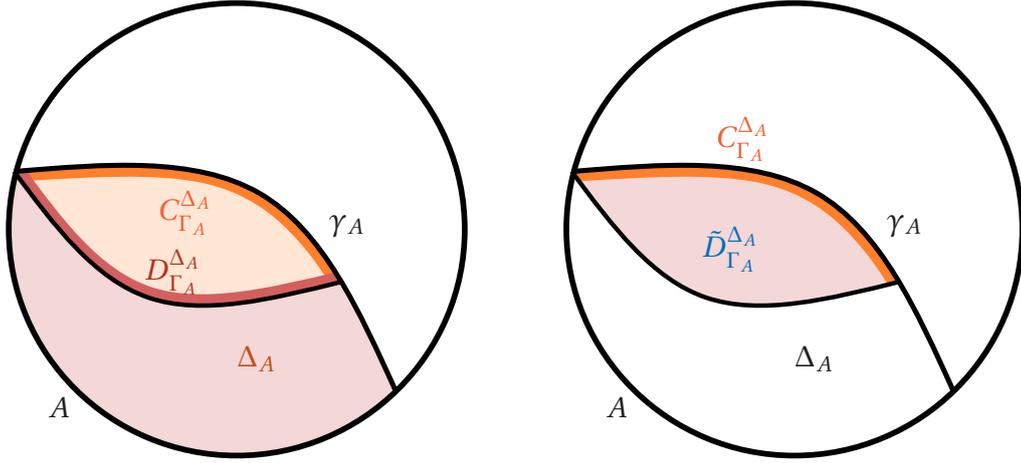
(a) Illustration of the sets of half-edges $C_{\Gamma_A}^{\Delta_A}$ and $D_{\Gamma_A}^{\Delta_A}$ given a cut $\Delta_A \subsetneq \Gamma_A$.

(b) Illustration of the sets of half-edges used in the proof of Theorem 14.6.

Figure 14.1: Tensor networks with one and two minimal cuts. The relevant ground state config-uration domains are denoted by Γ_A.

Definition 14.5 (Generalized minimal cut)**.** A cut Γ_A is a (ε, K)-*minimal cut* if

$$h_{\Gamma_A}^{\varepsilon}(\Delta_A) \geq K$$

for all $\Delta_A \in \mathscr{C}(\Gamma_A)$.

The utility of such a condition is that it controls the degree to which the spectrum of the corresponding cut link state ϕ_{Γ_A} is close to the boundary state ρ_A:

Theorem 14.6. *Consider a random tensor network state ρ constructed with a general link state $\phi \in \mathscr{P}_=(V)$ as in Eq. (12.4). Let A be a boundary region of the network, ρ_A the corresponding boundary state, and Γ_A an (ε, K)-minimal cut. Then the spectra of ρ_A and the link state ϕ_{Γ_A} on A are related as:*

$$\mathbb{E}\|\mathrm{spec}_+(\rho_A) - \mathrm{spec}_+(\phi_{\Gamma_A})\|_1 = \mathcal{O}(2^{-\frac{1}{4}K} + \sqrt{\varepsilon}).$$

In order to prove Theorem 14.6, we start by collecting some results relevant for the simultaneous smoothing of the link state. For general states, joint smoothing is an open problem [DF13, Dut11], however, if we use a link state we can straightforwardly do so, in similar fashion to joint smoothing of classical states [DF13]. A useful basic fact is that if we consider a basis in which a state ϕ is diagonal, then the optimizers for both the smooth min- and max-entropies may be taken to be diagonal in the same basis [Ren08]. To be precise, if $\phi = \sum_{i=1}^d \lambda_i |i\rangle\langle i|$ is a density matrix which is diagonal in the basis $\{|i\rangle\}$, then $H_{\min}^{\varepsilon}(\phi) = H_{\min}(\phi^\varepsilon)$ where

$$\phi^\varepsilon = \sum_i \min\{\lambda_i, 2^{-H_{\min}^{\varepsilon}(\phi)}\} |i\rangle\langle i|$$

and $T(\phi, \phi^\varepsilon) = \varepsilon$. Similarly, $H_{\max}^{\varepsilon}(\phi) = H_{\max}(\tilde{\phi}^\varepsilon)$ where

$$\tilde{\phi}^\varepsilon = \sum_i \tilde{\lambda}_i |i\rangle\langle i|$$

where $\tilde{\lambda}_i \leq \lambda_i$ and $T(\phi, \phi^\varepsilon) = \varepsilon$.

Lemma 14.7. *For a link state $\phi \in \mathscr{P}_=(V)$ as in Eq. (12.4), there exists a pure $\phi^\varepsilon \in \mathscr{P}_\le(V)$, which is such that*

$$T(\phi, \phi^\varepsilon) \le \sqrt{2^{|V|+1}\varepsilon}$$

and

$$H_{\min}(\phi_\Delta^\varepsilon) \ge H_{\min}^\varepsilon(\phi_\Delta)$$

for any vertex subset $\Delta \subseteq V$.

Proof. We may assume without loss of generality that $\phi = \bigotimes_{e \in E} \phi_e$ is such that each ϕ_e has Schmidt decomposition in the standard basis,

$$|\phi_e\rangle = \sum_{i=1}^{D_e} \sqrt{\lambda_{e,i}} |ii\rangle$$

This means we may write

$$|\phi\rangle = \sum_I \sqrt{\lambda_I} |I\rangle$$

where I runs over all possible basis elements along each edge $I = \{i_e\}_{e \in E}$ and

$$\lambda_I = \prod_{e \in E} \lambda_{e,i_e} \qquad |I\rangle = \bigotimes_{e \in E} |i_e i_e\rangle.$$

If we pick some subset $\Delta \subseteq V$ and we let the corresponding edge cut set be δ, and write E_1 and E_2 for the sets of vertices connecting only vertices with Δ or Δ^c respectively, then we may write $|\phi\rangle = |\phi_\delta\rangle \otimes |\phi_{E_1}\rangle \otimes |\phi_{E_2}\rangle$ and the reduced state on the vertices in Δ will be given by

$$\phi_\Delta = \phi_{E_1} \otimes \bigotimes_{\substack{e=(xy)\in\delta \\ x\in\Delta}} \phi_{e,x}.$$

An optimal smoothing for the min-entropy then is given by

$$\phi_\Delta^{\varepsilon,\Delta} = \phi_{E_1} \otimes \tilde\phi$$

where

$$\tilde\phi := \sum_{\{i_e\}_{e\in\delta}} \min\{\prod_{e\in\delta} \lambda_{e,i_e}, 2^{-H_{\min}^\varepsilon(\phi_\Delta)}\} \bigotimes_{e\in E} |i_e\rangle\langle i_e|.$$

In particular, this is the reduced state of the pure state $\phi^{\varepsilon,\Delta}$ defined by

$$|\phi^{\varepsilon,\Delta}\rangle = |\phi_{E_1}\rangle \otimes |\phi_{E_1}\rangle \otimes |\phi_\delta^{\varepsilon,\Delta}\rangle$$

where

$$|\phi_\delta^{\varepsilon,\Delta}\rangle = \sum_{\{i_e\}_{e\in\delta}} \min\{\prod_{e\in\delta} \sqrt{\lambda_{e,i_e}}, 2^{-\frac{1}{2}H_{\min}^\varepsilon}\} \bigotimes_{e\in E} |i_e i_e\rangle$$

and which has $P(\phi, \phi^{\varepsilon, \Delta}) \le \sqrt{2\varepsilon}$, since by the Fuchs-van de Graaff inequalities Eq. (14.1)

$$P(\phi, \phi^{\varepsilon, \Delta}) = P(\phi_\Delta, \phi_\Delta^{\varepsilon, \Delta}) \le \sqrt{2 T(\phi_\Delta, \phi_\Delta^{\varepsilon, \Delta})} \le \sqrt{2\varepsilon}.$$

In particular, we note that this state is of the form

$$|\phi^{\varepsilon, \Delta}\rangle = \sum_I \sqrt{\lambda_I^{\varepsilon, \Delta}} |I\rangle$$

for some nonnegative numbers $0 \le \lambda_I^{\varepsilon, \Delta} \le \lambda_I$. We let

$$|\phi^{\varepsilon}\rangle = \sum_I \min_{\Delta \subseteq V} \sqrt{\lambda_I^{\varepsilon, \Delta}} |I\rangle .$$

We now want to bound $P(\phi, \phi^{\varepsilon})$. To this end we observe that the fidelity can be bounded as

$$1 - F(\phi, \phi^{\varepsilon}) = \sum_I \sqrt{\lambda_I}(\sqrt{\lambda_I} - \min_{\Delta \subseteq V} \sqrt{\lambda_I^{\varepsilon, \Delta}})$$

$$\le \sum_I \sqrt{\lambda_I} \sum_{\Delta \subseteq V} (\sqrt{\lambda_I} - \sqrt{\lambda_I^{\varepsilon, \Delta}})$$

$$= \sum_{\Delta \subseteq V} 1 - F(\phi, \phi^{\varepsilon, \Delta})$$

and trivially for any Δ

$$1 + F(\phi, \phi^{\varepsilon}) \le 1 + F(\phi, \phi^{\varepsilon, \Delta}).$$

Thus, using that $F(\phi, \phi^{\varepsilon}) = F_*(\phi, \phi^{\varepsilon})$ since ϕ is normalized

$$P(\phi, \phi^{\varepsilon}) = \sqrt{1 - F(\phi, \phi^{\varepsilon})^2} = \sqrt{(1 - F(\phi, \phi^{\varepsilon}))(1 + F(\phi, \phi^{\varepsilon}))}$$

$$\le \sqrt{(\sum_{\Delta \subseteq V} 1 - F(\phi, \phi^{\varepsilon, \Delta}))((1 + F(\phi, \phi^{\varepsilon})))}$$

$$\le \sqrt{(\sum_{\Delta \subseteq V} 1 - F(\phi, \phi^{\varepsilon, \Delta}))((1 + F(\phi, \phi^{\varepsilon, \Delta})))}$$

$$\le \sqrt{\sum_{\Delta \subseteq V} (1 - F(\phi, \phi^{\varepsilon, \Delta})^2)}$$

$$\le \sqrt{2^{|V|+1}\varepsilon}$$

since $1 - F(\phi, \phi^{\varepsilon, \Delta})^2 = P(\phi, \phi^{\varepsilon, \Delta})^2 \le 2\varepsilon$. Thus, by the Fuchs-van de Graaff inequalities Eq. (14.1)

$$T(\phi, \phi^{\varepsilon}) \le P(\phi, \phi^{\varepsilon}) \le \sqrt{2^{|V|+1}\varepsilon}.$$

Moreover, in the basis $\{|I\rangle\}$ the state $|\phi^{\varepsilon}\rangle$ has elementwise smaller coefficients than $|\phi^{\varepsilon, \Delta}\rangle$, so we see that the reduced state $\phi_\Delta^{\varepsilon}$ is elementwise smaller than $\phi_\Delta^{\varepsilon, \Delta}$. This implies, by the Perron-Frobenius theorem, that the largest eigenvalue of $\phi_\Delta^{\varepsilon, \Delta}$ is larger than the largest eigenvalue of $\phi_\Delta^{\varepsilon}$, and hence $H_{\min}(\phi_\Delta^{\varepsilon}) \ge H_{\min}(\phi_\Delta^{\varepsilon, \Delta}) = H_{\min}^{\varepsilon}(\phi_\Delta)$. ∎

We will also use a similar result to obtain simultaneously smoothed max-entropies and min-entropies for classical states along a single cut.

Lemma 14.8. *Let E be a finite set, and let E_1 and E_2 label quantum systems with Hilbert spaces $\mathcal{H}_{E_i} = \otimes_{e \in E} \mathcal{H}_{e,i}$ for $i = 1, 2$. For $S \subseteq E$, let $\mathcal{H}_{S_i} = \otimes_{e \in S} \mathcal{H}_{e,i}$. Now suppose that $\phi \in \mathcal{P}_{=}(E_1 E_2)$ is a product state*

$$\phi = \bigotimes_{e \in E} \phi_e$$

where each ϕ_e is a pure state on $\mathcal{H}_{e,1} \otimes \mathcal{H}_{e,2}$. Then there exists $\phi^{\varepsilon} \in \mathcal{P}_{\leq}(E_1 E_2)$ which is such that

$$T(\phi, \phi^{\varepsilon}) \leq 2\sqrt{2^{|E|}\varepsilon}$$

and

$$H_{\max}(\phi^{\varepsilon}_{S_i}) \leq H^{\varepsilon}_{\max}(\phi_{S_i}) \quad and \quad H_{\min}(\phi^{\varepsilon}_{S_i}) \geq H^{\varepsilon}_{\min}(\phi_{S_i})$$

for all subsets $S \subseteq E$ and $i = 1, 2$.

Proof. The proof is similar to that of Lemma 14.7. We write out ϕ in the product basis, using the Schmidt decomposition of each $|\phi_e\rangle$

$$\phi = \sum_{I = \{i_e\}_{e \in E}} \sqrt{\lambda_I} |I_1 I_2\rangle$$

where $|I_i\rangle = \otimes_{e \in E} |i_{e,i}\rangle$. The reduced state ϕ_{E_1} is given by

$$\phi_{E_1} = \sum_{I = \{i_e\}_{e \in E}} \lambda_I |I_1\rangle\langle I_1|.$$

Choose a subset $S \subseteq E$ and and consider the following states in $\mathcal{P}_{\leq}(E_1)$

$$\phi^{\varepsilon, S_1, \min} = \sum_{I = \{i_e\}_{e \in E}} \lambda_I^{\varepsilon, S, \min} |I_1\rangle\langle I_1|$$

$$\phi^{\varepsilon, S_1, \max} = \sum_{I = \{i_e\}_{e \in E}} \lambda_I^{\varepsilon, S, \max} |I_1\rangle\langle I_1|$$

which can be chosen such that

$$0 \leq \lambda_I^{\varepsilon, S, \min} \leq \lambda_I, \quad T(\phi^{\varepsilon, S_1, \min}, \phi_{E_1}) \leq \varepsilon \quad and \quad H_{\min}(\phi^{\varepsilon, S_1, \min}_{S_1}) = H^{\varepsilon}_{\min}(\phi_{S_1}.$$

Similarly

$$0 \leq \lambda_I^{\varepsilon, S, \max} \leq \lambda_I, \quad T(\phi^{\varepsilon, S_1, \max}, \phi_{E_1}) \leq \varepsilon \quad and \quad H_{\max}(\phi^{\varepsilon, S_1, \max}_{S_1}) = H^{\varepsilon}_{\max}(\phi_{S_1}).$$

Then we let $\phi^{\varepsilon, S_1} \in \mathcal{P}_{\leq}(E_1)$ be given by

$$\phi^{\varepsilon, S_1} = \sum_{I = \{i_e\}_{e \in E}} \min\{\lambda_I^{\varepsilon, S, \min}, \lambda_I^{\varepsilon, S, \max}\} |I_1\rangle\langle I_1|.$$

Then it is easy to see that $T(\phi^{\varepsilon,S_1}, \phi_{E_1}) \le 2\varepsilon$ and $H_{\min}(\phi_{S_1}^{\varepsilon,S}) \ge H_{\min}^\varepsilon(\phi_{S_1})$ and $H_{\max}(\phi_{S_1}^{\varepsilon,S}) \le H_{\min}^\varepsilon(\phi_{S_1})$. Now let $\phi^\varepsilon \in \mathcal{P}_{\le}(E_1 E_2)$ be given by

$$\phi^\varepsilon = \sum_{I=\{i_e\}_{e \in E}} \min_{S \subseteq E}\{\sqrt{\lambda_I^{\varepsilon,S}}\} |I_1 I_2\rangle$$

so

$$\phi_{E_1}^\varepsilon = \sum_{I=\{i_e\}_{e \in E}} \min_{S \subseteq E}\{\lambda_I^{\varepsilon,S}\} |I_1\rangle\langle I_1|.$$

Then, since $\phi_{E_1}^\varepsilon \le \phi_{E_1}^{\varepsilon,S}$ we see that $H_{\min}(\phi_{S_1}^\varepsilon) \ge H_{\min}(\phi_{S_1}^{\varepsilon,S}) \ge H_{\min}^\varepsilon(\phi_{S_1})$ and similarly that $H_{\max}(\phi_{S_1}^\varepsilon) \le H_{\max}(\phi_{S_1}^{\varepsilon,S}) \le H_{\max}^\varepsilon(\phi_{S_1})$, and

$$T(\phi_{E_1}^\varepsilon, \phi_{E_1}) = \sum_I |\min_{S \subseteq E}\{\lambda_I^{\varepsilon,S}\} - \lambda_I| \le \sum_I \sum_{S \subseteq E} |\lambda_I^{\varepsilon,S} - \lambda_I|$$
$$\le \sum_{S \subseteq E} T(\phi^{\varepsilon,S}, \phi) \le 2^{|E|+1}\varepsilon.$$

It is easy to see this implies that $T(\phi_{E_1}^\varepsilon, \phi_{E_1}) \le 2\sqrt{2^{|E|}}\varepsilon$. By symmetry under exchanging E_1 and E_2, for any $S \subseteq E$, we must have $H_{\min/\max}^\varepsilon(\phi_{S_1}) = H_{\min/\max}^\varepsilon(\phi_{S_2})$ as well as $H_{\min/\max}(\phi_{S_1}^\varepsilon) = H_{\min/\max}(\phi_{S_2}^\varepsilon)$, so ϕ^ε satisfies all required conditions. ∎

The last ingredient we will need for the main result of this section is a relation between the Rényi-2 conditional entropy and the min and max-entropies:

Lemma 14.9. *Let $\phi = \phi_A \otimes \phi_{BC} \in \mathcal{P}_{\le}(ABC)$. Then*

$$H_2(AB|C)_{\phi|\phi} \ge H_{\min}(A)_\phi - H_{\max}(B)_\phi + \log \mathrm{tr}[\phi].$$

Proof. We prove the result for normalized $\phi \in \mathcal{P}_{=}(ABC)$, and then the result for subnormalized $\phi \in \mathcal{P}_{\le}(ABC)$ follows by applying to $\tilde{\phi} = \frac{\phi}{\mathrm{tr}[\phi]}$. By the product structure of ϕ, it is clear that

$$H_2(AB|C)_{\phi|\phi} = H_2(A)_\phi + H_2(B|C)_{\phi|\phi}.$$

Note that $H_2(A)_\phi \ge H_{\min}(A)_\phi$ by the standard monotonicity property of Rényi entropies in the Rényi index. For the second term, let R be a purifying system and ϕ_{BCR} be a purification of ϕ. Then

$$H_2(B|C)_{\phi|\phi} \ge H_{\min}(B|C)_\phi = -H_{\max}(B|R)_\phi \ge -H_{\max}(B)_\phi$$

using Eq. (14.4), Eq. (14.3) and the purity of ϕ_{BCR}, and in the last inequality, a data processing inequality (Theorem 6.19 in [Tom15]). ∎

We now prove Theorem 14.6:

Proof of Theorem 14.6. Let $\tilde{\phi}$ be a (subnormalized) pure background state. By Theorem 14.4 and the fact that $\tilde{\phi}_V$ is pure (i.e. $R = \emptyset$):

$$\mathbb{E}\|\mathrm{spec}_+(\rho_A) - \mathrm{spec}_+(\tilde{\phi}_{\Gamma_A})\|_1 \le \mathbb{E}\min_{V_A, V_{\bar{A}}} \|(V_A \otimes V_{\bar{A}} \otimes I_R)\rho(V_A^\dagger \otimes V_{\bar{A}}^\dagger \otimes I_R) - \tilde{\phi}\|_1$$
$$= \mathcal{O}(\mathrm{tr}[\tilde{\phi}]^{\frac{1}{4}}(2^{-\frac{1}{4}K_1} + 2^{-\frac{1}{4}K_2})),$$

where for any Δ_A with $\Gamma_A \subsetneq \Delta_A$, we define

$$K_1 = \min_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} H_2(\Gamma_A \setminus \Delta_A | \Gamma_A^c)_{\tilde{\phi}|\tilde{\phi}}$$

$$K_2 = \min_{\Delta_A \in C(A), \Gamma_A \subsetneq \Delta_A} H_2(\Delta_A \setminus \Gamma_A | \Gamma_A)_{\tilde{\phi}|\tilde{\phi}}.$$

So, we need to compute K_1 and K_2. Let E_1 be the set of edges (xy) for which $x, y \in \Gamma_A$, and let E_2 be the set of edges (xy) for which $x, y \in \Gamma_A^c$, so $E = E_1 \sqcup E_2 \sqcup \gamma_A$. Then the Hilbert space $\mathcal{H}_V$ factorizes as $\mathcal{H}_{E_1} \otimes \mathcal{H}_{E_2} \otimes \mathcal{H}_{\gamma_A}$. Let us assume that the state $\tilde{\phi}$ is a pure state such that it is a product state with respect to this factorization (we need this condition, rather than a product state over all edges, to accommodate for the smoothing), so

$$|\tilde{\phi}\rangle = |\tilde{\phi}_{E_1}\rangle \otimes |\tilde{\phi}_{E_2}\rangle \otimes |\tilde{\phi}_{\gamma_A}\rangle.$$

Then for $\Delta_A \subsetneq \Gamma_A$, let

$$\tilde{D}_{\Gamma_A}^{\Delta_A} = \{(e, x) : e = (xy), x \in \Gamma_A \setminus \Delta_A, y \in \Gamma_A\}.$$

Then, the collection of all half-edges (e, x) for which $x \in \Gamma_A \setminus \Delta_A$ equals the union $C_{\Gamma_A}^{\Delta_A} \sqcup \tilde{D}_{\Gamma_A}^{\Delta_A}$, and hence, by Lemma 14.9, with $A = \tilde{D}_{\Gamma_A}^{\Delta_A}$, $B = C_{\Gamma_A}^{\Delta_A}$ and $C = \Gamma_A^c$, and the product structure of $\tilde{\phi}$, we obtain

$$H_2(\Gamma_A \setminus \Delta_A | \Gamma_A^c)_{\tilde{\phi}|\tilde{\phi}} \geq H_{\min}(\tilde{D}_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{E_1}} - H_{\max}(C_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{\gamma_A}} + \log(\mathrm{tr}[\tilde{\phi}]).$$

Thus,

$$K_1 \geq \min_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} H_{\min}(\tilde{D}_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{E_1}} - H_{\max}(C_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{\gamma_A}} + \log \mathrm{tr}[\tilde{\phi}]. \qquad (14.15)$$

A similar argument bounds K_2 as

$$K_2 \geq \min_{\Delta_A \in C(A), \Gamma_A \subsetneq \Delta_A} H_{\min}(\tilde{D}_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{E_2}} - H_{\max}(C_{\Gamma_A}^{\Delta_A})_{\tilde{\phi}_{\gamma_A}} + \log \mathrm{tr}[\tilde{\phi}]. \qquad (14.16)$$

where

$$\tilde{D}_{\Gamma_A}^{\Delta_A} = \{(e, x) : e = (xy), x \in \Delta_A \setminus \Gamma_A, y \in \Gamma_A^c\}.$$

In particular, it is easy to see that this gives the desired result for $\varepsilon = 0$ if we take $\tilde{\phi} = \phi$.

Let us now discuss smoothing the link state. We start with a state which is a product along the edges E, and hence it can be written as $|\phi\rangle = |\phi_{E_1}\rangle \otimes |\phi_{E_2}\rangle \otimes |\phi_{\gamma_A}\rangle$. We now apply Lemma 14.7 to ϕ_{E_i} to obtain a pure state $\phi_{E_i}^\varepsilon$ such that $T(\phi_{E_i}, \phi_{E_i}^\varepsilon) \leq \mathcal{O}(\sqrt{\varepsilon})$ and for each cut $\Delta_A \subsetneq \Gamma_A$

$$H_{\min}(\tilde{D}_{\Gamma_A}^{\Delta_A})_{\phi_{E_1}^\varepsilon} \geq H_{\min}^\varepsilon(\tilde{D}_{\Gamma_A}^{\Delta_A})_\phi = H_{\min}^\varepsilon(D_{\Gamma_A}^{\Delta_A})_\phi$$

and a similar result for $\Gamma_A \subsetneq \Delta_A$ and $\phi_{E_2}^\varepsilon$. For ϕ_{γ_A} we use Lemma 14.8 with E_1 and E_2 the half-edges on either side of the cut to obtain a pure state $\phi_{\gamma_A}^\varepsilon \in \mathcal{P}_{\leq}(\gamma_A)$ which is

such that $T(\phi_{\gamma_A}, \phi^\varepsilon_{\gamma_A}) = \mathcal{O}(\sqrt{\varepsilon})$ and which satisfies for any cut Δ_A for which $\Delta_A \subsetneq \Gamma_A$ or $\Gamma_A \subsetneq \Delta_A$

$$H_{\max}(\phi^\varepsilon_{C^{\Delta_A}_{\Gamma_A}}) \leq H^\varepsilon_{\max}(\phi_{C^{\Delta_A}_{\Gamma_A}})$$
$$= H^\varepsilon_{\max}(C^{\Delta_A}_{\Gamma_A})_\phi.$$

We then let

$$\phi^\varepsilon = \phi^\varepsilon_{E_1} \otimes \phi^\varepsilon_{E_2} \otimes \phi^\varepsilon_{\gamma_A}$$

be the smoothed (pure) state. Let $\rho^\varepsilon_{V_\partial}$ be the random tensor network state with background state ϕ^ε. By the $\varepsilon = 0$ result for $\tilde{\phi} = \phi^\varepsilon$,

$$\mathbb{E}\|\mathrm{spec}_+(\rho^\varepsilon_A) - \mathrm{spec}_+(\phi^\varepsilon_{\Gamma_A})\|_1 = \mathcal{O}(\mathrm{tr}[\phi^\varepsilon]^{\frac{1}{4}}(2^{-\frac{1}{4}K_1} + 2^{-\frac{1}{4}K_2})) = \mathcal{O}(2^{-\frac{1}{4}K'_1} + 2^{-\frac{1}{4}K'_2})$$

where from Eq. (14.15) we get

$$K'_1 = K_1 - \log\mathrm{tr}[\phi^\varepsilon] \geq \min_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} H_{\min}(\tilde{D}^{\Delta_A}_{\Gamma_A})_{\phi^\varepsilon_{E_1}} - H_{\max}(C^{\Delta_A}_{\Gamma_A})_{\phi^\varepsilon_{\gamma_A}}$$
$$\geq \min_{\Delta_A \in C(A), \Delta_A \subsetneq \Gamma_A} H^\varepsilon_{\min}(\tilde{D}^{\Delta_A})_\phi - H^\varepsilon_{\max}(C^{\Delta_A}_{\Gamma_A})_\phi \geq K$$

and K_2 may be similarly bounded by using (14.16). We conclude, using Lemma 14.3, that

$$\mathbb{E}\|\mathrm{spec}_+(\rho_A) - \mathrm{spec}_+(\phi_{\Gamma_A})\|_1 \leq \mathbb{E}\|\mathrm{spec}_+(\rho^\varepsilon_A) - \mathrm{spec}_+(\phi^\varepsilon_{\Gamma_A})\|_1$$
$$+ \mathbb{E}\|\rho_{V_\partial} - \rho^\varepsilon_{V_\partial}\| + \|\phi_{\Gamma_A} - \phi^\varepsilon_{\Gamma_A}\|_1$$
$$= \mathcal{O}(2^{-\frac{1}{4}K}) + \mathcal{O}(\sqrt{\varepsilon}) + \mathcal{O}(\sqrt{\varepsilon})$$

as desired.   ∎

14.3   Random tensor network states and split transfer

Our results involving general background states are closely related to the quantum-information-theoretic task of *split transfer* introduced in [DH10], which can be understood as a variant of quantum state merging. The standard setup is as follows: two parties, Alice and Bob, share a state $\phi_{ABC_1...C_m}$, with Alice controlling system A, Bob controlling system B, the systems $C_1, ..., C_m$ being m "helpers". Let R be a reference system and $\phi_{ABRC_1...C_m}$ a purification of ϕ. Initially, the state is shared not only by Alice and Bob, but also with all the helper systems C_i. The goal of split transfer is to try to redistribute the state to Alice, Bob, and R, using local quantum operations and classical communication (LOCC) between Alice, Bob and the helper systems, and possibly with the assistance of additional maximally entangled states.

To be a little more detailed, a split transfer protocol consists of

(i) A partitioning of the set of the helper systems: $T_A \sqcup T_B = \{C_1, ..., C_m\}$.

(ii) For each $C_i \in T_A$ a number $K_{A,i}$ of shared maximally entangled qubits between Alice and C_i, and for each $C_i \in T_B$, a number $K_{B,i}$ of shared maximally entangled qubits between Bob and C_i.

(iii) An LOCC operation between Alice, Bob and the helper systems, which is such that after applying the protocol Alice and Bob share a state $\psi_{ABRC_1...C_m}$, which is close to $\phi_{ABRC_1...C_m}$ and such that now Alice possesses systems A and T_A, whereas Bob controls B and T_B. Moreover, after applying the protocol they may be in possession of a number $L_{A,i}$ or $L_{B,i}$ of (approximately) maximally entangled qubits between C_i, and respectively A or B.

In this case we say that the split transfer protocol has entanglement costs $K_{A,i} - L_{A,i}$ for all $C_i \in T_A$ and $K_{B,i} - L_{B,i}$ for all $C_i \in T_B$. A precise definition can be found as Definition 14 in [DH10].

Intuitively, the helper systems need to transfer their correlations with R to Alice and Bob, but without touching R. For instance, a very naive protocol could be that the helpers simply teleport their full system to either Alice or Bob, consuming EPR pairs. We can construct a potentially much more efficient protocol by way of random measurements. The precise procedure is detailed in Proposition 16 of [DH10]. Roughly speaking, such a protocol functions because random measurements have the effect of *decoupling* the helper systems from R. The helpers perform simultaneous random measurements on their systems, and send the results of their measurements to Alice and Bob. Then, Alice and Bob can use their share of the global state and their portions of the maximally-entangled states to perform a decoding operation conditioned on the results of the random measurements. The state they receive will be a purification of $\psi_{ABC_1...C_m}$, which will then be equivalent to the original $\psi_{ABRC_1...C_m}$ up to local isometries. The way we set up the split transfer protocol above was in a one-shot fashion: we get a single copy of ϕ and need to determine the optimal entanglement cost for the protocol.

One can also consider asymptotic versions (where one has many copies available and one cares about the optimal rate). An example application is the *entanglement of assistance*. Suppose that Alice, Bob and the helper systems C_i get many copies of a pure state ϕ. At what rate can they distill maximally entangled pairs between Alice and Bob, if Alice and Bob are allowed to perform LOCC operations with all the helper systems? In this case, the answer is that the rate is given by

$$\min_{T_A} S(AT_A)_\phi,$$

that is, by minimizing the entanglement entropy over all bipartitions. This rate is reminiscent of the importance of minimal cuts in a random tensor network. This connection was already explored in [HNQ$^+$16].

To see in some detail how the task of split transfer relates to random tensor networks, we consider three boundary regions, A and B, under the control of Alice and Bob, and the purifying system R. Each of the "assisting" parties represents a bulk vertex. We would like to know whether there exists a protocol in which the assisting parties are allowed to perform local operations and classical communication (LOCC) such that the state ϕ_{VR} is redistributed into a state ρ_{ABR} held by Alice and Bob that can be transformed by local isometries, acting only on A and B, to a state close to ϕ_{VR}. The protocol given in [DH10] consists of simultaneous random measurements by each of the helpers. This precisely corresponds to the random projections performed in constructing the random tensor network state with this background state!

In this light we can interpret Theorem 14.4 as a result on split transfer. Let us assume that $\phi \in \mathscr{P}_=(ABRC_1...C_n)$, and let us denote as usual by ρ_{ABR} the associated random

tensor network state and choose a partitioning $T_A \sqcup T_B$ of the assisting (bulk) parties. Since $H_2(A|B)_{\phi|\phi} \geq H_{\min}(A|B)_\phi$, Theorem 14.4 directly yields

Theorem. *Suppose that*

$$H_{\min}(S_A|BRT_B)_\phi \geq K_1 \tag{14.17}$$

for all non-empty subsets $S_A \subseteq T_A$ and

$$H_{\min}(S_B|ART_A)_\phi \geq K_2 \tag{14.18}$$

for all non-empty subsets $S_B \subseteq T_B$. Then

$$\mathbb{E} \min_{V_A, V_B} \|(V_A \otimes V_B \otimes I_R)\rho(V_A^\dagger \otimes V_B^\dagger \otimes I_R) - \phi_{ABRC_1...C_n}\|_1 = \mathcal{O}((2^{-\frac{1}{4}K_1} + 2^{-\frac{1}{4}K_2}). \tag{14.19}$$

where the minimum is over isometries $V_A : \mathcal{H}_A \to \mathcal{H}_{AT_A}$ and $V_B : \mathcal{H}_B \to \mathcal{H}_{BT_B}$.

This result shows that if K_1 and K_2 are sufficiently large, then after measurement in a random basis, the state possessed by Alice and Bob can, with high probability, be used to approximately reconstruct ϕ by acting with local isometries on the systems of Alice and Bob. If for the initial state ϕ the conditions in Eq. (14.17) and Eq. (14.17) are not satisfied, we can use another state where we have added an appropriate number of maximally entangled Bell pairs between the assisting parties, and this can be used to determine the entanglement costs.[2] An interesting open question in this context is whether one can generalize this result using *smooth* entropies in Eq. (14.17) and Eq. (14.17). The problem is that for a general state ϕ, one would need to perform simultaneous smoothing for all the relevant subsystems, which is an open problem.

There is an alternative approach, in which it is straightforwardly possible to use smooth entropies [DH10]. In this approach, one merges each party in T_A one by one, and similarly for T_B. That is, we choose some ordering $T_A = \{1, \ldots, m\} = [m]$ and we apply a sequence of state merging protocols where we merge the state in m steps, where a single step merges $A \cup T_A \setminus [i-1]$ into $A \cup T_A \setminus [i]$. In this case it is not hard to see that, if we allow some error the entanglement cost is determined by the smooth conditional entropies $H_{\min}^\varepsilon(\{i+1\}|BR[i]T_B)_\phi$. We perform a similar protocol for B and the assisting T_B systems.

14.3.1 Split transfer and recovery in holography

Split transfer is closely related to *subregion-subregion duality*, or *entanglement wedge reconstruction*, in holography. Suppose we have an AdS semiclassical stationary geometry which is dual to some boundary CFT state ρ. We fix a time-reversal invariant spatial slice and partition the boundary into A and $\bar{A}$. Let γ_A be the minimal surface for A. Then we denote by Γ_A the entanglement wedge of A, the region which is enclosed by A and γ_A. The claim of entanglement wedge reconstruction is that if we act with some low-energy local bulk operator in the entanglement wedge, we can reconstruct the action of this operator on the boundary system A.

---

[2]In fact, the protocol in [DH10] is slightly more general than what we describe; rather than measuring a random state one could also measure a random projection of rank greater than 1. This can be used to obtain EPR pairs between the helpers and and Alice and Bob to get nonzero $L_{A,i}$ and $L_{B,i}$.

One way to make this more precise is by considering a *code subspace* of bulk states S (which should be thought of as a subspace of states given by acting with low-energy operators on a fixed semiclassical space-time). We may then take a reference system R of the same dimension as S, and consider a maximally entangled state ϕ between R and S. The AdS/CFT correspondence now provides an encoding of the bulk into the boundary, yielding a boundary state $\rho_{A\bar{A}R}$. In this set-up, the claim of entanglement wedge reconstruction is that we can act with an isometry on A to recover $\phi_{\Gamma_A R}$ (and Γ_A is actually the maximal such region). If the entanglement wedge for $\bar{A}$ is the complement of Γ_A, then it is clear that this requirement is closely related to split transfer (if we only consider A and Γ_A the corresponding task is state merging). These ideas and the precise relation to quantum information theory have been developed in a large number of works, amongst which [HNQ$^+$16, Har17, AP20, AP22].

In particular, in [AP20, AP22] the idea that one-shot quantum information is relevant to entanglement wedge reconstruction has been advanced. This makes sense, as the natural setting in holography is a single-shot setting. While the limit of large effective central charge, and hence small G_N, is reminiscent of a many-copy limit, there are situations of interest (for instance in relation to the black hole information paradox) where the reference system is actually large, or in other words, we consider a situation with large bulk entropy. In this case, where we assume we have a bulk state ϕ, the RT formula in Eq. (10.4) gets adapted to be the *quantum extremal surface* formula:

$$H(\rho_A) = \min \text{ext}_{\gamma_A}\{\frac{|\gamma_A|}{4G_N} + H(\Gamma_A)_\phi\}$$

where we minimize over extremal surfaces γ_A, and we minimize the joint contribution of the area of γ_A and the bulk entropy contained in the associated entanglement wedge Γ_A. We note that this formula has a natural tensor network interpretation: consider a tensor network state where the background state is given by $\phi = \phi_{V^{(b)}R} \otimes \phi_{V^{(l)}}$, where $\phi_{V^{(b)}R}$ is a general background state (which accounts for bulk entropy) and $\phi_{V^{(l)}}$ is a link state on a graph $G = (V, E)$. Let us take maximally entangled link states with dimension D. Then, for some cut Γ_A with edge set γ_A, we have

$$H(\phi_{\Gamma_A}) = \log(D)|\gamma_A| + H(\phi_{\Gamma_A}^{(b)})$$

and we may hope that minimization over this quantity along the cuts gives a good approximation to the entropy. Whether this is indeed valid, depends on the structure of the background state ϕ. A proposal put forth in [AP20] is that the surface γ_A with entanglement wedge Γ_A gives the max-*entanglement wedge* if Γ_A is the largest region which is such that for any other surface δ_A homologous to A, with Δ_A the region enclosed by A and δ_A, and where Δ_A is contained in Γ_A, it holds that

$$H_{\min}^\varepsilon(\Gamma_A \setminus \Delta_A | \Gamma_A^c R) \gg \frac{|\gamma_A| - |\delta_A|}{4G_N}.$$

In this case, Γ_A should be the largest region which can be (approximately) reconstructed from A. Again, one can think of a random tensor network where the background state is a tensor product of a bulk state and a maximally entangled link state of dimension D with $\log(D) = \Theta(G_N^{-1})$ along the discretization of the space. Then this condition is (apart from the simultaneous smoothing problem) equivalent to Eq. (14.13). If Γ_A is the max-entanglement wedge for A, and its complement Γ_A^c is the max-entanglement wedge

for $\bar{A}$, then the holographic encoding of the bulk state into the boundary can be seen as a version of one-shot split transfer (if we restrict to just Γ_A and A, it is a version of one-shot quantum state merging, as discussed in [AP20]). See [AP20] for a detailed discussion of this proposal for holographic systems.

Discussion and open questions

In this chapter we explained how one can use one-shot entropies for random tensor networks. We did not yet discuss the case where there are *two* competing minimal cuts in this framework. This should be possible, and we hope to address this question in forthcoming work [CPWW]. This would lead to an entropy computation similar to Eq. (11.8). Another important, but substantially harder open question, is the joint smoothing problem. This is relevant to random tensor network states with general background states, but would in general be highly relevant to multiparty quantum information tasks.

Bibliography

[ABF20]   Pablo Arrighi, Cédric Bény, and Terry Farrelly. A quantum cellular automaton for one-dimensional QED. *Quantum Information Processing*, 19(3):88, 2020.

[ACGR20]   Sophie Achard, Marianne Clausel, Irène Gannaz, and François Roueff. New results on approximate Hilbert pairs of wavelet filters with common factors. *Applied and Computational Harmonic Analysis*, 49(3):1025–1045, 2020.

[ADH15]   Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk locality and quantum error correction in AdS/CFT. *Journal of High Energy Physics*, 2015(4):1–34, 2015.

[ADM19]   Vincenzo Alba, Jerome Dubail, and Marko Medenjak. Operator entanglement in interacting integrable quantum systems: the case of the rule 54 chain. *Physical Review Letters*, 122(25):250603, 2019.

[AE15]   Martin Ammon and Johanna Erdmenger. *Gauge/gravity duality: Foundations and applications*. Cambridge University Press, 2015.

[AEMM19]   Ahmed Almheiri, Netta Engelhardt, Donald Marolf, and Henry Maxfield. The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. *Journal of High Energy Physics*, 2019(12):1–47, 2019.

[AEPW02]   Koenraad Audenaert, Jens Eisert, Martin B Plenio, and Reinhard F Werner. Entanglement properties of the harmonic chain. *Physical Review A*, 66(4):042327, 2002.

[AF04]   Robert Alicki and Mark Fannes. Continuity of quantum conditional information. *Journal of Physics A: Mathematical and General*, 37(5):L55, 2004.

[AGZ10]   Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, 2010.

[AHM+20]   Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. Replica wormholes and the entropy of Hawking radiation. *Journal of High Energy Physics*, 2020(5):1–42, 2020.

[AHM+21]   Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. *Reviews of Modern Physics*, 93(3):035002, 2021.

[AL99]   Daniel S Abrams and Seth Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Physical Review Letters*, 83(24):5162, 1999.

[AMPS13]   Ahmed Almheiri, Donald Marolf, Joseph Polchinski, and James Sully. Black holes: complementarity or firewalls? *Journal of High Energy Physics*, 2013(2):1–20, 2013.

[AN12]   Guillaume Aubrun and Ion Nechita. Realigning random states. *Journal of Mathematical Physics*, 53(10):102210, 2012.

[AP20]   Chris Akers and Geoff Penington. Leading order corrections to the quantum extremal surface prescription. *arXiv preprint arXiv:2008.03319*, 2020.

[AP22]   Chris Akers and Geoff Penington. Quantum minimal surfaces from quantum error correction, 2022.

[AR19]   Chris Akers and Pratik Rath. Holographic Rényi entropy from quantum error correction. *Journal of High Energy Physics*, 2019(5):1–24, 2019.

[Ara71]   Huzihiro Araki. On quasifree states of CAR and Bogoliubov automorphisms. *Publications of the Research Institute for Mathematical Sciences*, 6:385–442, 1971.

[Arr19]   Pablo Arrighi. An overview of quantum cellular automata. *Natural Computing*, 18(4):885–899, 2019.

[Arv69]   William B Arveson. Subalgebras of C^*-algebras. *Acta Mathematica*, 123(1):141–224, 1969.

[AS17]   Guillaume Aubrun and Stanisław J Szarek. *Alice and Bob meet Banach*, volume 223. American Mathematical Society, 2017.

[ASY12]   Guillaume Aubrun, Stanisław J Szarek, and Deping Ye. Phase transitions for random states and a semicircle law for the partial transpose. *Physical Review A*, 85(3):030302, 2012.

[Aub12]   Guillaume Aubrun. Partial transposition of random states and non-centered semicircular distributions. *Random Matrices: Theory and Applications*, 1(02):1250001, 2012.

[Bat99]   Guy Battle. *Wavelets and renormalization*. World Scientific, 1999.

[BBCC11]   Teodor Banica, Serban Teodor Belinschi, Mireille Capitaine, and Benoit Collins. Free Bessel laws. *Canadian Journal of Mathematics*, 63(1):3–37, 2011.

[BBMC20]   Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin-Lic Chan. Quantum algorithms for quantum chemistry and quantum materials science. *Chemical Reviews*, 120(22):12685–12717, 2020.

[BCC+15] Ning Bao, ChunJun Cao, Sean M Carroll, Aidan Chatwin-Davies, Nicholas Hunter-Jones, Jason Pollack, and Grant N Remmen. Consistency conditions for an AdS multiscale entanglement renormalization ansatz correspondence. *Physical Review D*, 91(12):125036, 2015.

[BCSF21] Johannes Borregaard, Matthias Christandl, and Daniel Stilck França. Noise-robust exploration of quantum matter on near-term quantum devices. *npj Quantum Information*, 7(45), 2021.

[BDPT18] Alessandro Bisio, Giacomo Mauro D'Ariano, Paolo Perinotti, and Alessandro Tosini. Thirring quantum cellular automaton. *Physical Review A*, 97(3):032132, 2018.

[Bek20] Jacob D Bekenstein. Black holes and entropy. In *JACOB BEKENSTEIN: The Conservative Revolutionary*, pages 307–320. World Scientific, 2020.

[BHJ+12] Christoph Brockt, Jutho Haegeman, David Jennings, Tobias J Osborne, and Frank Verstraete. The continuum limit of a tensor network: a path integral representation. *arXiv preprint arxiv:1210.5401*, 2012.

[Bil08] Patrick Billingsley. *Probability and Measure*. John Wiley & Sons, 2008.

[Bla06] Bruce Blackadar. *Operator algebras: theory of C*-algebras and von Neumann algebras*, volume 122. Springer Science & Business Media, 2006.

[BMEK17] Annabelle Bohrdt, Christian B Mendl, Manuel Endres, and Michael Knap. Scrambling and thermalization in a diffusive quantum many-body system. *New Journal of Physics*, 19(6):063001, 2017.

[BOT13] Marc Burger, Narutaka Ozawa, and Andreas Thom. On Ulam stability. *Israel Journal of Mathematics*, 193(1):109–129, 2013.

[BP13] Fatih Bulut and Wayne N Polyzou. Wavelets in field theory. *Physical Review D*, 87(11):116011, 2013.

[BPSW19] Ning Bao, Geoffrey Penington, Jonathan Sorce, and Aron C Wall. Beyond toy models: distilling tensor networks in full AdS/CFT. *Journal of High Energy Physics*, 2019(11):1–63, 2019.

[BR03] Ola Bratteli and Derek W Robinson. *Operator Algebras and Quantum Statistical Mechanics 2*. Springer, 2003.

[BR12] Ola Bratteli and Derek W Robinson. *Operator Algebras and Quantum Statistical Mechanics*, volume 1. Springer Science & Business Media, 2012.

[BRSS15] Gavin K Brennen, Peter Rohde, Barry C Sanders, and Sukhwinder Singh. Multiscale quantum simulation of quantum field theory using wavelets. *Physical Review A*, 92(3):032315, 2015.

[BS10] Zhidong Bai and Jack W Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*, volume 20. Springer, 2010.

[BW20]   Raphael Bousso and Elizabeth Wildenhain. Gravity/ensemble duality. *Physical Review D*, 102(6):066005, 2020.

[CAB$^+$21] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

[Car86]   John L Cardy. Operator content of two-dimensional conformally invariant theories. *Nuclear Physics B*, 270:186–204, 1986.

[CDF92]   Albert Cohen, Ingrid Daubechies, and J-C Feauveau. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45(5):485–560, 1992.

[CGAH$^+$17] Jordan S Cotler, Guy Gur-Ari, Masanori Hanada, Joseph Polchinski, Phil Saad, Stephen H Shenker, Douglas Stanford, Alexandre Streicher, and Masaki Tezuka. Black holes and random matrices. *Journal of High Energy Physics*, 2017(5):1–54, 2017.

[CGGPG13] Benoît Collins, Carlos E González-Guillén, and David Pérez-García. Matrix product states, random matrix theory and the principle of maximum entropy. *Communications in Mathematical Physics*, 320(3):663–677, 2013.

[CGRPG19] Juan Ignacio Cirac, José Garre-Rubio, and David Pérez-García. Mathematical open problems in projected entangled pair states. *Revista Matemática Complutense*, 32(3):579–599, 2019.

[CGW14]   Andrew M Childs, David Gosset, and Zak Webb. The Bose-Hubbard model is QMA-complete. In *International Colloquium on Automata, Languages, and Programming*, pages 308–319. Springer, 2014.

[CHP$^+$19] Jordan Cotler, Patrick Hayden, Geoffrey Penington, Grant Salton, Brian Swingle, and Michael Walter. Entanglement wedge reconstruction via universal recovery channels. *Physical Review X*, 9(3):031011, 2019.

[Chr75]   Erik Christensen. Perturbations of type I von Neumann algebras. *Journal of the London Mathematical Society*, 2(3):395–405, 1975.

[Chr77a]   Erik Christensen. Perturbation of operator algebras. *Inventiones mathematicae*, 43(1):1–13, 1977.

[Chr77b]   Erik Christensen. Perturbations of operator algebras II. *Indiana University Mathematics Journal*, 26(5):891–904, 1977.

[Chr80]   Erik Christensen. Near inclusions of C^*-algebras. *Acta Mathematica*, 144(1):249–265, 1980.

[CMMN19]   Jordan S Cotler, M Reza Mohammadi Mozaffar, Ali Mollabashi, and Ali Naseh. Entanglement renormalization for weakly interacting fields. *Physical Review D*, 99(8):085005, 2019.

[CN16] Benoit Collins and Ion Nechita. Random matrix techniques in quantum information theory. *Journal of Mathematical Physics*, 57(1):015215, 2016.

[CNŻ10] Benoît Collins, Ion Nechita, and Karol Życzkowski. Random graph states, maximal flow and Fuss–Catalan distributions. *Journal of Physics A: Mathematical and Theoretical*, 43(27):275–303, 2010.

[CNŻ13] Benoît Collins, Ion Nechita, and Karol Życzkowski. Area law for random graph states. *Journal of Physics A: Mathematical and Theoretical*, 46(30):305302, 2013.

[CPGSV17] J Ignacio Cirac, David Perez-Garcia, Norbert Schuch, and Frank Verstraete. Matrix product unitaries: structure, symmetries, and topological invariants. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(8):083105, 2017.

[CPGSV21] J Ignacio Cirac, David Perez-Garcia, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4):045003, 2021.

[CPWW] Newton Cheng, Geoff Penington, Michael Walter, and Freek Witteveen. Random tensor networks with nontrivial link spectra. In preparation.

[CR87] Alan Carey and Simon Ruijsenaars. On fermion gauge groups, current algebras and Kac-Moody algebras. *Acta Applicandae Mathematica*, 10:1–86, 1987.

[CRO+19] Yudong Cao, Jonathan Romero, Jonathan P Olson, Matthias Degroote, Peter D Johnson, Mária Kieferová, Ian D Kivlichan, Tim Menke, Borja Peropadre, Nicolas PD Sawaya, et al. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, 2019.

[CRSV17] Rui Chao, Ben W Reichardt, Chris Sutherland, and Thomas Vidick. Overlapping qubits. *arXiv preprint arXiv:1701.01062*, 2017.

[CSS+12] Erik Christensen, Allan M Sinclair, Roger R Smith, Stuart A White, Wilhelm Winter, et al. Perturbations of nuclear C^*-algebras. *Acta mathematica*, 208(1):93–150, 2012.

[CST+21] Andrew M Childs, Yuan Su, Minh C Tran, Nathan Wiebe, and Shuchen Zhu. Theory of Trotter error with commutator scaling. *Physical Review X*, 11(1):011020, 2021.

[CU09] Kunal Narayan Chaudhury and Michael Unser. Construction of Hilbert transform pairs of wavelet bases and Gabor-like transforms. *IEEE Transactions on Signal Processing*, 57(9):3411–3425, 2009.

[CU10] Kunal Narayan Chaudhury and Michael Unser. On the Hilbert transform of wavelets. *IEEE transactions on signal processing*, 59(4):1890–1894, 2010.

[CV09] Philippe Corboz and Guifré Vidal. Fermionic multiscale entanglement renormalization ansatz. *Physical Review B*, 80:165129, 2009.

[CX22] Yu-An Chen and Yijia Xu. Equivalence between fermion-to-qubit mappings in two spatial dimensions. *arXiv preprint arXiv:2201.05153*, 2022.

[DB19] Alexander M Dalzell and Fernando GSL Brandão. Locally accurate MPS approximations for ground states of one-dimensional gapped local Hamiltonians. *Quantum*, 3:187, 2019.

[DBWR14] Frédéric Dupuis, Mario Berta, Jürg Wullschleger, and Renato Renner. One-shot decoupling. *Communications in Mathematical Physics*, 328(1):251–284, 2014.

[DDP18] Blake R Duschatko, Philipp T Dumitrescu, and Andrew C Potter. Tracking the quantized information transfer at the edge of a chiral Floquet phase. *Physical Review B*, 98(5):054309, 2018.

[DF13] Lukas Drescher and Omar Fawzi. On simultaneous min-entropy smoothing. In *2013 IEEE International Symposium on Information Theory*, pages 161–165. IEEE, 2013.

[DH10] Nicolas Dutil and Patrick Hayden. One-shot multiparty state merging. *arXiv preprint arXiv:1011.1974*, 2010.

[DHM19] Xi Dong, Daniel Harlow, and Donald Marolf. Flat entanglement spectra in fixed-area states of quantum gravity. *Journal of High Energy Physics*, 2019(10):1–25, 2019.

[DMW21] Xi Dong, Sean McBride, and Wayne W Weng. Replica wormholes and holographic entanglement negativity. *arXiv preprint arXiv:2110.11947*, 2021.

[DNBD19] Jacopo De Nardis, Denis Bernard, and Benjamin Doyon. Diffusion in generalized hydrodynamics and quasiparticle scattering. *SciPost Physics*, 6:049, 2019.

[Don16] Xi Dong. The gravity dual of Rényi entropy. *Nature communications*, 7(1):1–6, 2016.

[DP17] Giacomo Mauro D'Ariano and Paolo Perinotti. Quantum cellular automata and free quantum field theory. *Frontiers of Physics*, 12(1):1–11, 2017.

[DQW21] Xi Dong, Xiao-Liang Qi, and Michael Walter. Holographic entanglement negativity and replica symmetry breaking. *arXiv preprint arXiv:2101.11029*, 2021.

[Dut11] Nicolas Dutil. Multiparty quantum protocols for assisted entanglement distillation. *arXiv preprint arXiv:1105.4657*, 2011.

[DW19] Ronald De Wolf. Quantum computing: Lecture notes. *arXiv preprint arXiv:1907.09415*, 2019.

[ECP10] Jens Eisert, Marcus Cramer, and Martin B Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82(1):277, 2010.

[EV09a] Glen Evenbly and Guifré Vidal. Algorithms for entanglement renormaliza-tion. *Physical Review B*, 79(14):144108, 2009.

[EV09b] Glen Evenbly and Guifré Vidal. Entanglement renormalization in two spatial dimensions. *Physical Review Letters*, 102(18):180406, 2009.

[EV10a] G Evenbly and G Vidal. Entanglement renormalization in free bosonic systems: real-space versus momentum-space renormalization group trans-forms. *New Journal of Physics*, 12:025007, 2010.

[EV10b] Glen Evenbly and Guifré Vidal. Entanglement renormalization in noninter-acting fermionic systems. *Physical Review B*, 81(23):235102, 2010.

[EV13] Glen Evenbly and Guifré Vidal. Quantum criticality with the multi-scale en-tanglement renormalization ansatz. In *Strongly Correlated Systems*, pages 99–130. Springer, 2013.

[EW15] Netta Engelhardt and Aron C Wall. Quantum extremal surfaces: holo-graphic entanglement entropy beyond the classical regime. *Journal of High Energy Physics*, 2015(1):1–27, 2015.

[EW16] Glen Evenbly and Steven R White. Entanglement renormalization and wavelets. *Physical Review Letters*, 116:140403, 2016.

[EW18] Glen Evenbly and Steven R White. Representation and design of wavelets using unitary circuits. *Physical Review A*, 97(5):052314, 2018.

[F$^+$82] Richard P Feynman et al. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6/7), 1982.

[Far20] Terry Farrelly. A review of quantum cellular automata. *Quantum*, 4:368, 2020.

[FH20] Michael Freedman and Matthew B Hastings. Classification of quantum cellular automata. *Communications in Mathematical Physics*, 376(2):1171–1222, 2020.

[FHH20] Lukasz Fidkowski, Jeongwan Haah, and Matthew B Hastings. Exactly solv-able model for a 4+ 1 d beyond-cohomology symmetry-protected topologi-cal phase. *Physical Review B*, 101(15):155124, 2020.

[FHH22] Michael Freedman, Jeongwan Haah, and Matthew B Hastings. The group structure of quantum cellular automata. *Communications in Mathematical Physics*, pages 1–26, 2022.

[FLM13] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum correc-tions to holographic entanglement entropy. *Journal of High Energy Physics*, 2013(11):1–18, 2013.

[FMS12] Philippe Francesco, Pierre Mathieu, and David Sénéchal. *Conformal Field Theory*. Springer Science & Business Media, 2012.

[FPPV19] Lukasz Fidkowski, Hoi Chun Po, Andrew C Potter, and Ashvin Vishwanath. Interacting invariants for floquet phases of fermions in two dimensions. *Physical Review B*, 99(8):085115, 2019.

[FRV17] Adrián Franco-Rubio and Guifré Vidal. Entanglement and correlations in the continuous multi-scale entanglement renormalization ansatz. *Journal of High Energy Physics*, 2017(12):129, 2017.

[Fuc95] Jürgen Fuchs. *Affine Lie algebras and quantum groups: An introduction, with applications in conformal field theory*. Cambridge University Press, 1995.

[GAE07] David Gross, Koenraad Audenaert, and Jens Eisert. Evenly distributed unitaries: On the structure of unitary designs. *Journal of Mathematical Physics*, 48(5):052104, 2007.

[GHKV18] Sarang Gopalakrishnan, David A Huse, Vedika Khemani, and Romain Vasseur. Hydrodynamics of operator spreading and quasiparticle diffusion in interacting integrable systems. *Physical Review B*, 98(22):220303, 2018.

[GJ12] James Glimm and Arthur Jaffe. *Quantum physics: a functional integral point of view*. Springer Science & Business Media, 2012.

[GNP21] Zongping Gong, Adam Nahum, and Lorenzo Piroli. Coarse-grained entanglement and operator growth in anomalous dynamics. *arXiv preprint arXiv:2109.07408*, 2021.

[GNVW12] David Gross, Vincent Nesme, Holger Vogts, and Reinhard F Werner. Index theory of one dimensional quantum walks and cellular automata. *Communications in Mathematical Physics*, 310(2):419–454, 2012.

[GPC21] Zongping Gong, Lorenzo Piroli, and J Ignacio Cirac. Topological lower bound on quantum chaos by entanglement growth. *Physical Review Letters*, 126(16):160601, 2021.

[GRV17] Martin Ganahl, Julián Rincón, and Guifre Vidal. Continuous matrix product states for quantum fields: An energy minimization algorithm. *Physical Review Letters*, 118(22):220402, 2017.

[GSSC20] Zongping Gong, Christoph Sünderhauf, Norbert Schuch, and J Ignacio Cirac. Classification of matrix-product unitaries with symmetries. *Physical Review Letters*, 124(10):100402, 2020.

[GWSG20] Koen Groenland, Freek Witteveen, Kareljan Schoutens, and Rene Gerritsma. Signal processing techniques for efficient compilation of controlled rotations in trapped ions. *New Journal of Physics*, 22(6):063006, 2020.

[Haa12] Rudolf Haag. *Local quantum physics: Fields, particles, algebras*. Springer Science & Business Media, 2012.

[Haa21] Jeongwan Haah. Clifford quantum cellular automata: Trivial group in 2d and witt group in 3d. *Journal of Mathematical Physics*, 62(9):092202, 2021.

[Har16] Daniel Harlow. Jerusalem lectures on black holes and quantum information. *Reviews of Modern Physics*, 88(1):015002, 2016.

[Har17] Daniel Harlow. The Ryu–Takayanagi formula from quantum error correction. *Communications in Mathematical Physics*, 354(3):865–912, 2017.

[Har18] Daniel Harlow. TASI Lectures on the Emergence of the Bulk in AdS/CFT. *arXiv preprint arXiv:1802.01040*, 2018.

[Has07] Matthew B Hastings. An area law for one-dimensional quantum systems. *Journal of statistical mechanics: theory and experiment*, 2007(08):P08024, 2007.

[Has10] Matthew B Hastings. Locality in quantum systems. *Quantum Theory from Small to Large Scales*, 95:171–212, 2010.

[Has13] Matthew B Hastings. Classifying quantum phases with the Kirby torus trick. *Physical Review B*, 88(16):165114, 2013.

[Has17] Matthew B Hastings. The asymptotics of quantum max-flow min-cut. *Communications in Mathematical Physics*, 351(1):387–418, 2017.

[Haw75] Stephen W Hawking. Particle creation by black holes. In *Euclidean quantum gravity*, pages 167–188. World Scientific, 1975.

[Haw76] Stephen W Hawking. Black holes and thermodynamics. *Physical Review D*, 13(2):191, 1976.

[HCO$^+$10] Jutho Haegeman, J Ignacio Cirac, Tobias J Osborne, Henri Verschelde, and Frank Verstraete. Applying the variational principle to (1+ 1)-dimensional quantum field theories. *Physical Review Letters*, 105(25):251601, 2010.

[HFH18] Jeongwan Haah, Lukasz Fidkowski, and Matthew B Hastings. Nontrivial quantum cellular automata in higher dimensions. *arXiv preprint arXiv:1812.01625*, 2018.

[HGPC21] Reza Haghshenas, Johnnie Gray, Andrew C Potter, and Garnet Kin Chan. The variational power of quantum circuit tensor networks. *arXiv preprint arXiv:2107.01307*, 2021.

[HH13] Daniel Harlow and Patrick Hayden. Quantum computation vs. firewalls. *Journal of High Energy Physics*, 2013(6):1–56, 2013.

[HHEG20] Jonas Haferkamp, Dominik Hangleiter, Jens Eisert, and Marek Gluza. Contracting projected entangled pair states is average-case hard. *Physical Review Research*, 2(1):013010, 2020.

[HHKL18] Jeongwan Haah, Matthew Hastings, Robin Kothari, and Guang Hao Low. Quantum algorithm for simulating real time evolution of lattice Hamiltonians. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 350–360. IEEE, 2018.

[HK06] Matthew B Hastings and Tohru Koma. Spectral gap and exponential decay of correlations. *Communications in Mathematical Physics*, 265(3):781–804, 2006.

[HLW06] Patrick Hayden, Debbie W Leung, and Andreas Winter. Aspects of generic entanglement. *Communications in Mathematical Physics*, 265(1):95–117, 2006.

[HM04] Gary T Horowitz and Juan Maldacena. The black hole final state. *Journal of High Energy Physics*, 2004(02):008, 2004.

[HNQ$^+$16] Patrick Hayden, Sepehr Nezami, Xiao-Liang Qi, Nathaniel Thomas, Michael Walter, and Zhao Yang. Holographic duality from random tensor networks. *Journal of High Energy Physics*, 2016(11):1–56, 2016.

[Hoo93] Gerard'T Hooft. Dimensional reduction in quantum gravity. *arXiv preprint gr-qc/9310026*, 1993.

[HOVV13] Jutho Haegeman, Tobias J Osborne, Henri Verschelde, and Frank Verstraete. Entanglement renormalization for quantum fields in real space. *Physical Review Letters*, 110:100402, 2013.

[HP07] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. *Journal of High Energy Physics*, 2007(09):120, 2007.

[HR17] Fenner Harper and Rahul Roy. Floquet topological order in interacting systems of bosons and fermions. *Physical Review Letters*, 118(11):115301, 2017.

[HRT07] Veronika E Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A covariant holographic entanglement entropy proposal. *Journal of High Energy Physics*, 2007(07):062, 2007.

[HSW$^+$18] Jutho Haegeman, Brian Swingle, Michael Walter, Jordan Cotler, Glen Evenbly, and Volkher B Scholz. Rigorous free-fermion entanglement renormalization from wavelet theory. *Physical Review X*, 8:011003, 2018.

[HVDH21] Markus Hauru, Maarten Van Damme, and Jutho Haegeman. Riemannian optimization of isometric tensor networks. *SciPost Phys*, 10:040, 2021.

[Jay57] Edwin T Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620, 1957.

[JKP09] Nathaniel Johnston, David W Kribs, and Vern I Paulsen. Computing stabilized norms for quantum operations via the theory of completely bounded maps. *Quantum Information & Computation*, 9(1):16–35, 2009.

[JM08] Richard Jozsa and Akimasa Miyake. Matchgates and classical simulation of quantum circuits. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 464, pages 3089–3106, 2008.

[Joh88]   Barry E Johnson. Approximately multiplicative maps between Banach algebras. *Journal of the London Mathematical Society*, 2(2):294–316, 1988.

[KFNR21]   Jonah Kudler-Flam, Vladimir Narovlansky, and Shinsei Ryu. Negativity spectra in random tensor networks and holography. *arXiv preprint arXiv:2109.02649*, 2021.

[Kit06]   Alexei Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321(1):2–111, 2006.

[KK17]   Isaac H Kim and Michael J Kastoryano. Entanglement renormalization, quantum error correction, and bulk causality. *Journal of High Energy Physics*, 2017:40, 2017.

[KKR06]   Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local hamiltonian problem. *Siam journal on computing*, 35(5):1070–1097, 2006.

[KLM01]   Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. *nature*, 409(6816):46–52, 2001.

[KR97]   Richard V Kadison and John R Ringrose. *Fundamentals of the theory of operator algebras. Volume I: Elementary Theory*, volume 1. American Mathematical Society, 1997.

[KR05]   Andreas Klappenecker and Martin Rotteler. Mutually unbiased bases are complex projective 2-designs. In *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, pages 1740–1744. IEEE, 2005.

[KS16]   Robert König and Volkher B Scholz. Matrix product approximations to multipoint functions in two-dimensional conformal field theory. *Physical Review Letters*, 117(12):121601, 2016.

[KS17a]   Isaac H Kim and Brian Swingle. Robust entanglement renormalization on a noisy quantum computer. *arXiv preprint arxiv:1711.07500*, 2017.

[KS17b]   Robert König and Volkher B Scholz. Matrix product approximations to conformal field theories. *Nuclear Physics B*, 920:32–121, 2017.

[KSVV02]   Alexei Yu Kitaev, Alexander Shen, Mikhail N Vyalyi, and Mikhail N Vyalyi. *Classical and quantum computation*. American Mathematical Society, 2002.

[KTP20]   Isaac Kim, Eugene Tang, and John Preskill. The ghost in the radiation: Robust encodings of the black hole interior. *Journal of High Energy Physics*, 2020(6):1–65, 2020.

[Lan91]   Rolf Landauer. Information is physical. *Physics Today*, 44(5):23–29, 1991.

[LC17]   Guang Hao Low and Isaac L Chuang. Optimal Hamiltonian simulation by quantum signal processing. *Physical Review Letters*, 118(1):010501, 2017.

[LC19] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.

[LC21] Ryan Levy and Bryan K Clark. Entanglement entropy transitions with random tensor networks. *arXiv preprint arXiv:2108.02225*, 2021.

[Lee17] Ching Hua Lee. Generalized exact holographic mapping with wavelets. *Physical Review B*, 96:245103, 2017.

[Llo96] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[LM13] Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 2013(8):1–29, 2013.

[LMMR14] Jonathan Lux, Jan Müller, Aditi Mitra, and Achim Rosch. Hydrodynamic long-time tails after a quantum quench. *Physical Review A*, 89:053608, 2014.

[LPG21] Cécilia Lancien and David Pérez-García. Correlation length in random MPS and PEPS. *Annales Henri Poincaré*, pages 1–82, 2021.

[LPWV20] Javier Lopez-Piqueres, Brayden Ware, and Romain Vasseur. Mean-field entanglement transitions in random tree tensor networks. *Physical Review B*, 102(6):064202, 2020.

[LR72] Elliott H Lieb and Derek W Robinson. The finite group velocity of quantum spin systems. *Communications in Mathematical Physics*, 28(3):251–257, 1972.

[Lun76] Lars-Erik Lundberg. Quasi-free "second quantization". *Communications in Mathematical Physics*, 50:103–112, 1976.

[LVFL21] Yaodong Li, Romain Vasseur, Matthew Fisher, and Andreas WW Ludwig. Statistical mechanics model for Clifford random tensor networks and monitored quantum circuits. *arXiv preprint arXiv:2110.02988*, 2021.

[LVV15] Zeph Landau, Umesh Vazirani, and Thomas Vidick. A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians. *Nature Physics*, 11(7):566–569, 2015.

[Mal99] Juan Maldacena. The large-N limit of superconformal field theories and supergravity. *International Journal of Theoretical Physics*, 38(4):1113–1133, 1999.

[Mal08] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 2008.

[Mar86] Norman Margolus. Quantum computation. *Annals of the New York Academy of Sciences*, 480:487–497, 1986.

[MB21] Erica Morgan and Fernando GSL Brandão. A classical model correspondence for *G*-symmetric random tensor networks. *Journal of Physics Communications*, 2021.

[MM09] Marc Mezard and Andrea Montanari. *Information, Physics, and omputation*. Oxford University Press, 2009.

[MM11] Cristopher Moore and Stephan Mertens. *The Nature of Computation*. OUP Oxford, 2011.

[MPSW15] Kevin Marshall, Raphael Pooser, George Siopsis, and Christian Weedbrook. Quantum simulation of quantum field theory using continuous variables. *Physical Review A*, 92(6):063825, 2015.

[MS17] James A Mingo and Roland Speicher. *Free Probability and Random Matrices*, volume 35. Springer, 2017.

[MSS16] Juan Maldacena, Stephen H Shenker, and Douglas Stanford. A bound on chaos. *Journal of High Energy Physics*, 2016(8):1–17, 2016.

[MV18a] Ashley Milsted and Guifré Vidal. Geometric interpretation of the multi-scale entanglement renormalization ansatz. *arXiv preprint arXiv:1812.00529*, 2018.

[MV18b] Ashley Milsted and Guifré Vidal. Tensor networks as conformal transformations. *arXiv preprint arXiv:1805.12524*, 2018.

[MVS21] Raimel Medina, Romain Vasseur, and Maksym Serbyn. Entanglement transitions from restricted Boltzmann machines. *Physical Review B*, 104(10):104205, 2021.

[MWW20] Donald Marolf, Shannon Wang, and Zhencheng Wang. Probing phase transitions of holographic entanglement entropy with fixed area states. *Journal of High Energy Physics*, 2020(12):1–41, 2020.

[Naa13] Pieter Naaijkens. *Quantum spin systems on infinite lattices*. Springer, 2013.

[NC02] Michael A Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. American Association of Physics Teachers, 2002.

[NRSR21] Adam Nahum, Sthitadhi Roy, Brian Skinner, and Jonathan Ruhman. Measurement and entanglement phase transitions in all-to-all quantum circuits, on quantum trees, and in Landau-Ginsburg theory. *PRX Quantum*, 2(1):010352, 2021.

[NS06] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.

[NSW13] Bruno Nachtergaele, Volkher B Scholz, and Reinhard F Werner. Local approximation of observables and commutator bounds. In *Operator methods in mathematical physics*, pages 143–149. Springer, 2013.

[NSY19]   Bruno Nachtergaele, Robert Sims, and Amanda Young. Quasi-locality
          bounds for quantum lattice systems. I. Lieb-Robinson bounds, quasi-local
          maps, and spectral flow automorphisms. *Journal of Mathematical Physics*,
          60(6):061101, 2019.

[NW16]   Sepehr Nezami and Michael Walter. Multipartite entanglement in stabilizer
          tensor networks. *arXiv preprint arxiv:1608.02595*, 2016.

[NW20]   Sepehr Nezami and Michael Walter. Multipartite entanglement in stabilizer
          tensor networks. *Physical Review Letters*, 125(24):241602, 2020.

[OP04]   Masanori Ohya and Dénes Petz. *Quantum entropy and its use.* Springer
          Science & Business Media, 2004.

[Orú14]   Román Orús. A practical introduction to tensor networks: Matrix product
          states and projected entangled pair states. *Annals of Physics*, 349:117–158,
          2014.

[OS21a]   Tobias J Osborne and Alexander Stottmeister. Conformal field theory from
          lattice fermions. *arXiv preprint arXiv:2107.13834*, 2021.

[OS21b]   Tobias J Osborne and Alexander Stottmeister. Quantum simulation of
          conformal field theory. *arXiv preprint arXiv:2109.14214*, 2021.

[OT05]   Roberto Oliveira and Barbara M Terhal. The complexity of quantum
          spin systems on a two-dimensional square lattice. *arXiv preprint quant-
          ph/0504050*, 2005.

[Pag93a] Don N Page. Average entropy of a subsystem. *Physical Review Letters*,
          71(9):1291, 1993.

[Pag93b] Don N Page. Information in black hole radiation. *Physical Review Letters*,
          71(23):3743, 1993.

[Par04]   Chun-Gil Park. On an approximate automorphism on a C^*-algebra. *Pro-
          ceedings of the American Mathematical Society*, 132(6):1739–1745, 2004.

[PB20]   Marc Potters and Jean-Philippe Bouchaud. *A First Course in Random Matrix
          Theory: For Physicists, Engineers and Data Scientists.* Cambridge University
          Press, 2020.

[PC20]   Lorenzo Piroli and J Ignacio Cirac. Quantum cellular automata, tensor
          networks, and area laws. *arXiv preprint arXiv:2007.15371*, 2020.

[PEDC05] Martin B Plenio, Jens Eisert, J Dreissig, and Marcus Cramer. Entropy, entan-
          glement, and area: analytical results for harmonic lattice systems. *Physical
          Review Letters*, 94(6):060503, 2005.

[Pen20]   Geoffrey Penington. Entanglement wedge reconstruction and the informa-
          tion paradox. *Journal of High Energy Physics*, 2020(9):1–84, 2020.

[Pet90]   Dénes Petz. *Algebra of the canonical commutation relation.* Citeseer, 1990.

[PEV09] Robert NC Pfeifer, Glen Evenbly, and Guifré Vidal. Entanglement renormalization, scale invariance, and quantum criticality. *Physical Review A*, 79(4):040301, 2009.

[PFM⁺16] Hoi Chun Po, Lukasz Fidkowski, Takahiro Morimoto, Andrew C Potter, and Ashvin Vishwanath. Chiral Floquet phases of many-body localized bosons. *Physical Review X*, 6(4):041070, 2016.

[PR13] Kyriakos Papadodimas and Suvrat Raju. An infalling observer in AdS/CFT. *Journal of High Energy Physics*, 2013(10):1–73, 2013.

[PR14] Kyriakos Papadodimas and Suvrat Raju. State-dependent bulk-boundary maps and black hole complementarity. *Physical Review D*, 89(8):086010, 2014.

[Pre92] John Preskill. Do black holes destroy information? In *Proceedings of the International Symposium on Black Holes, Membranes, Wormholes and Superstrings, SKalara and DV Nanopoulos, eds.(World Scientific, Singapore, 1993) pp*, pages 22–39. World Scientific, 1992.

[Pre98] John Preskill. Lecture notes for physics 229: Quantum information and computation. *California Institute of Technology*, 16(1):1–8, 1998.

[Pre18] John Preskill. Quantum computing in the NISQ era and beyond. *arXiv preprint arxiv:1801.00862*, 2018.

[PSSY19] Geoff Penington, Stephen H Shenker, Douglas Stanford, and Zhenbin Yang. Replica wormholes and the black hole interior. *arXiv preprint arXiv:1911.11977*, 2019.

[PTWW14] Francesc Perera, Andrew Toms, Stuart White, and Wilhelm Winter. The Cuntz semigroup and stability of close C^*-algebras. *Analysis & PDE*, 7(4):929–952, 2014.

[PYHP15] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015(6):1–55, 2015.

[Qi13] Xiao-Liang Qi. Exact holographic mapping and emergent space-time geometry. *arXiv preprint arxiv:1309.6282*, 2013.

[Ren08] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6(01):1–127, 2008.

[RT06a] Shinsei Ryu and Tadashi Takayanagi. Aspects of holographic entanglement entropy. *Journal of High Energy Physics*, 2006(08):045, 2006.

[RT06b] Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from the anti–de Sitter space/conformal field theory correspondence. *Physical Review Letters*, 96(18):181602, 2006.

[RT17] Mukund Rangamani and Tadashi Takayanagi. Holographic entanglement entropy. In *Holographic Entanglement Entropy*, pages 35–47. Springer, 2017.

[RWW20] Daniel Ranard, Michael Walter, and Freek Witteveen. A converse to Lieb-Robinson bounds in one dimension using index theory. *arXiv:2012.00741*, 2020.

[Sár17] Gábor Sárosi. AdS_2 holography and the SYK model. *arXiv preprint arXiv:1711.08482*, 2017.

[SB16] Sukhwinder Singh and Gavin K Brennen. Holographic construction of quantum field theory using wavelets. *arXiv preprint arxiv:1606.05068*, 2016.

[SBK05] Ivan W Selesnick, Richard G Baraniuk, and Nick C Kingsbury. The dual-tree complex wavelet transform. *IEEE signal processing magazine*, 22(6):123–151, 2005.

[SCD+22] Wilbur Shirley, Yu-An Chen, Arpit Dua, Tyler D Ellison, Nathanan Tantivasadakarn, and Dominic J. Williamson. Three-dimensional quantum cellular automata from chiral semion surface topological order and beyond. *arXiv preprint arXiv:2202.05442*, 2022.

[Sch11] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, 2011.

[Sel01] Ivan W Selesnick. Hilbert transform pairs of wavelet bases. *IEEE Signal Processing Letters*, 8:170–173, 2001.

[Sel02] Ivan W Selesnick. The design of approximate Hilbert transform pairs of wavelet bases. *IEEE Transactions in Signal Processing*, 50(5):1144–1152, 2002.

[SJ21] Troy J Sewell and Stephen P Jordan. Preparing renormalization group fixed points on nisq hardware. *arXiv preprint arXiv:2109.09787*, 2021.

[Smi02] Jan Smit. *Introduction to quantum fields on a lattice*. Cambridge University Press, 2002.

[SMMT20] Alexander Stottmeister, Vincenzo Morinelli, Gerardo Morsella, and Yoh Tanimoto. Operator-algebraic renormalization and wavelets. *arXiv preprint arxiv:2002.01442*, 2020.

[SN96] Gilbert Strang and Truong Nguyen. *Wavelets and filter banks*. SIAM, 1996.

[SNBV+19] David T Stephen, Hendrik Poulsen Nautrup, Juani Bermejo-Vega, Jens Eisert, and Robert Raussendorf. Subsystem symmetries, quantum cellular automata, and computational phases of quantum matter. *Quantum*, 3:142, 2019.

[SS08] Yasuhiro Sekino and Leonard Susskind. Fast scramblers. *Journal of High Energy Physics*, 2008(10):065, 2008.

[SS14] Stephen H Shenker and Douglas Stanford. Black holes and the butterfly effect. *Journal of High Energy Physics*, 2014(3):1–25, 2014.

[ŞSBC18] M Burak Şahinoğlu, Sujeet K Shukla, Feng Bi, and Xie Chen. Matrix product representation of locality preserving unitaries. *Physical Review B*, 98(24):245122, 2018.

[SSS19] Phil Saad, Stephen H Shenker, and Douglas Stanford. JT gravity as a matrix integral. *arXiv preprint arXiv:1903.11115*, 2019.

[SSSY21] Phil Saad, Stephen H Shenker, Douglas Stanford, and Shunyu Yao. Wormholes without averaging. *arXiv preprint arXiv:2103.16754*, 2021.

[ST08] Kostas Skenderis and Marika Taylor. The fuzzball proposal for black holes. *Physics Reports*, 467(4-5):117–171, 2008.

[Sti55] W Forrest Stinespring. Positive functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216, 1955.

[Sus95] Leonard Susskind. The world as a hologram. *Journal of Mathematical Physics*, 36(11):6377–6396, 1995.

[SV96] Andrew Strominger and Cumrun Vafa. Microscopic origin of the Bekenstein-Hawking entropy. *Physics Letters B*, 379(1-4):99–104, 1996.

[SV09] Norbert Schuch and Frank Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nature physics*, 5(10):732–735, 2009.

[SW04] Benjamin Schumacher and Reinhard F Werner. Reversible quantum cellular automata. *arXiv preprint quant-ph/0405174*, 2004.

[Swi12a] Brian Swingle. Constructing holographic spacetimes using entanglement renormalization. *arXiv preprint arXiv:1209.3304*, 2012.

[Swi12b] Brian Swingle. Entanglement renormalization and holography. *Physical Review D*, 86:065007, 2012.

[SWVC07] Norbert Schuch, Michael M Wolf, Frank Verstraete, and J Ignacio Cirac. Computational complexity of projected entangled pair states. *Physical Review Letters*, 98(14):140506, 2007.

[Tak03] Masamichi Takesaki. *Theory of operator algebras III*, volume 125. Springer Science & Business Media, 2003.

[TGS+19] Minh C Tran, Andrew Y Guo, Yuan Su, James R Garrison, Zachary Eldredge, Michael Foss-Feig, Andrew M Childs, and Alexey V Gorshkov. Locality and digital quantum simulation of power-law interactions. *Physical Review X*, 9(3):031006, 2019.

[Tom15]   Marco Tomamichel.   *Quantum Information Processing with Finite Resources: Mathematical Foundations*, volume 5. Springer, 2015.

[TW05]   Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Physical Review Letters*, 94(17):170201, 2005.

[Ula60]   Stanislaw M Ulam. A collection of mathematical problems. *New York*, 29, 1960.

[VC06]   Frank Verstraete and J Ignacio Cirac. Matrix product states represent ground states faithfully. *Physical Review B*, 73(9):094423, 2006.

[VC10]   Frank Verstraete and J Ignacio Cirac. Continuous matrix product states for quantum fields. *Physical Review Letters*, 104:190405, 2010.

[Vid07]   Guifré Vidal.   Entanglement renormalization.   *Physical Review Letters*, 99:220405, 2007.

[Vid08]   Guifré Vidal. Class of quantum many-body states that can be efficiently simulated. *Physical Review Letters*, 101(11):110501, 2008.

[VPYL19] Romain Vasseur, Andrew C Potter, Yi-Zhuang You, and Andreas WW Ludwig. Entanglement transitions from holographic random tensor networks. *Physical Review B*, 100(13):134203, 2019.

[Was98]   Antony Wassermann. Operator algebras and conformal field theory III. fusion of positive energy representations of $LSU(N)$ using bounded operators. *Inventiones Mathematicae*, 133(3):467–538, 1998.

[Wat18]   John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[Whi92]   Steven R White. Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, 69(19):2863, 1992.

[Wil13]   Mark M Wilde. *Quantum information theory*. Cambridge University Press, 2013.

[Win16]   Andreas Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 347(1):291–313, 2016.

[WM07]   Daniel F Walls and Gerard J Milburn. *Quantum Optics*. Springer Science & Business Media, 2007.

[Woj97]   Przemyslaw Wojtaszczyk. *A Mathematical Introduction to Wavelets*. Cambridge University Press, 1997.

[WSSW22] Freek Witteveen, Volkher Scholz, Brian Swingle, and Michael Walter. Quantum circuit approximations and entanglement renormalization for the Dirac field in 1+ 1 dimensions. *Communications in Mathematical Physics*, 389(1):75–120, 2022.

[WW20]   Henrik Wilming and Albert H Werner. Lieb-Robinson bounds imply locality of interactions. *arXiv preprint arXiv:2006.10062*, 2020.

[WW21a]   Michael Walter and Freek Witteveen. Hypergraph min-cuts from quantum entropies. *Journal of Mathematical Physics*, 62(9):092203, 2021.

[WW21b]   Michael Walter and Freek Witteveen. Jupyter notebook. https://github.com/amsqi/pyfermions/blob/master/notebooks/bosonic-mera.ipynb, 2021.

[WW21c]   Freek Witteveen and Michael Walter. Bosonic entanglement renormalization circuits from wavelet theory. *SciPost Physics*, 10(6):143, 2021.

[YHQ16]   Zhao Yang, Patrick Hayden, and Xiao-Liang Qi. Bidirectional holographic codes and sub-AdS locality. *Journal of High Energy Physics*, 2016(1):1–24, 2016.

[YLFC21] Zhi-Cheng Yang, Yaodong Li, Matthew Fisher, and Xiao Chen. Entanglement phase transitions in random stabilizer tensor networks. *arXiv preprint arXiv:2107.12376*, 2021.

[YO05]   Runyi Yu and Huseyin Ozkaramanli. Hilbert transform pairs of orthogonal wavelet bases: Necessary and sufficient conditions. *IEEE Transactions on Signal Processing*, 53(12):4723–4725, 2005.

[ZCZ$^+$19] Bei Zeng, Xie Chen, Duan-Lu Zhou, Xiao-Gang Wen, et al. *Quantum information meets quantum matter*. Springer, 2019.

[ZFFM20] Zoltán Zimborás, Terry Farrelly, Szilárd Farkas, and Lluis Masanes. Does causal dynamics imply local interactions? *arXiv preprint arXiv:2006.10707*, 2020.

[ZGV19]   Yijian Zou, Martin Ganahl, and Guifre Vidal. Magic entanglement renormalization for quantum fields. *arXiv preprint arxiv:1906.04218*, 2019.

[ZL20]   Carolyn Zhang and Michael Levin. Classification of interacting Floquet phases with $U(1)$ symmetry in two dimensions. *arXiv preprint arXiv:2010.02253*, 2020.

[ZW18]   Modjtaba Shokrian Zini and Zhenghan Wang. Conformal field theories as scaling limit of anyonic chains. *Communications in Mathematical Physics*, 363:877–953, 2018.

Abstract

In this dissertation we study the fascinating interaction between *quantum information theory* and *many-body physics*. Many-body physics broadly are physical systems which are made up of a large number of subsystems. While the laws and principles of quantum mechanics are well known and can be formulated in compact equations, understanding many-body physics and the emergent phenomena related to it gives rise to a whole new set of challenges. In this dissertation we contribute to this field by showing how in three different domains perspectives from quantum information theory and computation can be useful in the study of many-body physics.

In Part I we investigate quantum systems in one spatial dimension. We show that the ground state, the lowest energy state, of a class of free systems can be prepared efficiently on a quantum computer using a certain circuit structure. Free quantum systems can be simulated efficiently by classical computers; so the possibility of preparing such ground states on a quantum computer is well-known. However, we show that this is possible using a method which implements an important physical principle: real-space renormalization. This method is known as *entanglement renormalization*. There is an associated tensor network ansatz, the multiscale entanglement renormalization ansatz (MERA). We show that entanglement renormalization in free systems is closely related to the theory of wavelets. Wavelets allow a decomposition of functions into a basis set of functions, similar to Fourier analysis. However, in the case of wavelets, the basis functions are not plane waves but localized 'wave packets'. Quantization of appropriately chosen wavelets yields an entanglement renormalization scheme. We explain how to find such wavelets, and we show that if one has wavelets with the right properties one can obtain accurate MERA approximation. We show that this has a natural continuum limit, which is related to free quantum field theories.

In Part II we study a different aspect of many-body quantum mechanics. Whereas in the first part we focussed on finding ground states, here we are interested in the *dynamics* of one-dimensional quantum systems. This means that we have a system with a state changing over time. Closed systems have *unitary* time evolution. Moreover, in physical systems with local interactions, one finds that time evolution also conserves a certain amount of *locality*. Based on these two general principles we dynamics which consist of a single time step which are both unitary and (approximately) locality preserving, which we call *approximately locality preserving unitary* (ALPU). If one imposes strict locality this is known as a quantum cellular automaton. In one spatial dimension these have been classified by an index which measures an information flow. We show that this classification extends to approximately local dynamics, using tools from the theory of operator algebras. We prove that a one-dimensional ALPU can be connected by a continuous path of ALPUs to the identity if and only if the ALPU has no net information flow to the left or the right.

Finally, Part III is inspired by the interaction between quantum information theory and gravity. Developing a theory of quantum gravity which can describe our universe is one the great outstanding challenges of theoretical physical. An important physical phenomenon where such a theory would be relevant are black holes. Based on general physical principles, it appears that black holes have an entropy (that is, a number of degrees of freedom) corresponding to their *area*. This has lead to the development of *holographic* quantum gravity, where a $d+1$-dimensional gravitational theory is dual to

a d-dimensional non-gravitational quantum theory, which lives on the boundary of the gravitational spacetime. We study a simple *toy model* for this phenomenon: while it is certainly not describing a theory of quantum gravity, it shows in a surprisingly accurate way certain mechanisms which are crucial in holographic gravity. The model is a tensor network model where we choose uniformly random tensors. We generalize this model by adapting the usual PEPS model to use different link states and we deduce properties of the entanglement spectrum. These results closely mirror conjectured properties of holographic systems.

Samenvatting

Quantuminformatie in systemen met veel deeltjes

In dit proefschrift bestuderen we de fascinerende interactie tussen *quantuminformatietheorie* en *veel-deeltjes-fysica*. Veel-deeltjes-fysica is een verzamelnaam voor fysische systemen die bestaan uit een groot aantal subsystemen. Hoewel de wetten van de quantummechanica in principe goed begrepen zijn en compact geformuleerd kunnen worden, brengt het begrijpen van systemen die uit een groot aantal quantumdeeltjes bestaan een geheel nieuwe verzameling aan uitdagingen met zich mee. In dit proefschrift bestuderen we, op drie uiteenlopende onderdelen, hoe perspectieven uit quantuminformatietheorie een rol kunnen spelen bij een beter begrip van dit soort systemen.

In Deel I kijken we naar eendimensionale systemen. We tonen aan dat de *grondtoestand* van een klasse vrije systemen op een quantumcomputer kan worden geprepareerd middels een circuit met een bepaalde structuur. Het is weinig verbazend dat dit mogelijk is; vrije quantumsystemen zijn eveneens te simuleren op een klassieke computer. We laten echter zien dat dit mogelijk is op een manier die een belangrijk fysisch principe implementeert: renormalizatie. Deze methode heet *verstrengelingsrenormalizatie*. We tonen aan dat verstrengelingsrenormalizatie in vrije systemen nauw verband houdt met de theorie van *wavelets*. Wavelets bieden een manier om een functie te ontbinden als een som van gelocaliseerde golffuncties. We demonstreren dat dit tevens leidt tot een natuurlijke interpretatie van de continue limiet van het model, een quantumveldentheorie.

In Deel II van het proefschrift bekijken we een ander aspect van de quantummechanica. Waar we ons in het eerste deel richten op het probleem van het beschrijven van grondtoestanden, bestuderen we hier de *dynamica* van eendimensionale quantumsystemen. Hiermee bedoelen we hoe het systeem in de loop van de tijd verandert. Ten eerste geldt dat in een gesloten system de tijdsevolutie van een quantumsysteem *unitair* is. Ten tweede geldt dat in fysieke systemen waar de interacties lokaal zijn dat de tijdsevolutie ook een bepaalde mate van lokaliteit bewaart. Op basis van deze twee gegevens bestuderen we dynamica die in een enkele tijdsstap plaatsvindt en die zowel unitair is, als bij benadering lokaal. We generalizeren eerder werk met betrekking to strikt lokale dynamica en classificeren zulk dynamica in eendimensionale systemen.

In Deel III, het laatste deel van het proefschrift, bestuderen we een aspect van de interactie tussen quantuminformatie en zwaartekracht. Het blijkt dat zwarte gaten, als quantummechanische objecten, kunnen worden beschreven op het *oppervlak* dat het zwarte gat begrenst. Dit is aanleiding geweest voor het ontwikkelen van *holografische quantumzwaartekracht*, waarin een universum met een theorie van quantumzwaartekracht equivalent is aan een reguliere quantumtheorie (zonder zwaartekracht) op de rand (en dus in een dimensie lager) van de ruimte leeft. Wij bestuderen een model dat bedoeld is als een *speelgoedmodel* voor dit fenomeen: hoewel het geenszins een realistich model voor zwaartekracht laat het op een accurate manier bepaalde mechanismes zien die cruciaal zijn in holographische zwaartekracht. Het model is een tensornetwerk (een bepaalde manier om quantumtoestanden te construeren) met uniform willekeurige tensoren. We laten zien dat onze generalizatie van dit model informatietheoretische aspecten van holografische quantumzwaartekracht reproduceert.

Acknowledgements

First of all, I am indebted to Michael for his fantastic supervision during the last four years. It was highly motivating to collaborate, both in research and teaching. His intelligence and deep knowledge of every subfield of mathematics, computer science or physics we encountered are unmatched and never ceased to amaze me. Every meeting with Michael led to new research questions and potential projects. I also appreciate his genuine care for the well-being of all his students.

I would also like to thank my second advisor, Eric. While we did not collaborate extensively, it was always a pleasure to talk and he made me feel welcome at the KdVI.

I have been fortunate to collaborate with great scientists during the last four years. I would like to thank in particular my coauthors Newton Cheng, Koen Groenland, Rene Gerritsma, Geoff Penington, Volkher Scholz, Kareljan Schoutens and Brian Swingle. Collaboration with Daniel Ranard was especially pleasant, and I would like to thank him for sharing many of his nights with our days. Thanks to Koen for letting me use his thesis template, and to Jonas for writing feedback.

I would like to thank the staff at CWI, QuSoft and the KdVI for creating such a well-organized and pleasant working place; in particular Doutzen, Evelien, Marieke and Suzanne. Unfortunately, due to the pandemic I spent two years mostly away from the office. Upon returning I felt clearly that I missed the place, and especially my colleagues at QuSoft. Especially the ever growing group of fellow PhD students created a great atmosphere, both on a social and scientific level. Thanks Alvaro, András, Arjan, Bas, Dimitri, Dyon, Farrokh, Fran, Galina, Garazi, Harold, Jan, Jana, Joran, Joris, Jordi, Koen, Lynn, Marten, Niels, Peter, Philip, Quinten, Rene, Sebastian, Srinivasan, Subhasree, Tom, Yanlin and Yfke! It has been a joy to be part of this research community and I hope to return.

Outside of academia I felt greatly supported by my family. I know my parents and Marit would have been proud of this thesis. The last four years I have been lucky to live in different constellations with my amazing friends Isaac, Marte, Arin, Koen and Sam. Together with all my other wonderful friends they lighten up my life. I would like to thank Moos for many long walks and Sjors for comfortable evenings on the couch. There are more important things in life than science. Rasila, you always show me those things; I am deeply grateful for our time together.