MAY 15, 1935 PHYSICAL REVIEW VOLUME 47

Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?

A. EINSTEIN, B. PODOLSKY AND N. ROSEN, *Institute for Advanced Study, Princeton, New Jersey*

(Received March 25, 1935)

In a complete theory there is an element corresponding to each element of reality. A sufficient condition for the reality of a physical quantity is the possibility of predicting it with certainty, without disturbing the system. In quantum mechanics in the case of two physical quantities described by non-commuting operators, the knowledge of one precludes the knowledge of the other. Then either (1) the description of reality given by the wave function in quantum mechanics is not complete or (2) these two quantities cannot have simultaneous reality. Consideration of the problem of making predictions concerning a system on the basis of measurements made on another system that had previously interacted with it leads to the result that if (1) is false then (2) is also false. One is thus led to conclude that the description of reality as given by a wave function is not complete.

1.

ANY serious consideration of a physical theory must take into account the distinction between the objective reality, which is independent of any theory, and the physical concepts with which the theory operates. These concepts are intended to correspond with the objective reality, and by means of these concepts we picture this reality to ourselves.

In attempting to judge the success of a physical theory, we may ask ourselves two questions: (1) "Is the theory correct?" and (2) "Is the description given by the theory complete?" It is only in the case in which positive answers may be given to both of these questions, that the concepts of the theory may be said to be satisfactory. The correctness of the theory is judged by the degree of agreement between the conclusions of the theory and human experience. This experience, which alone enables us to make inferences about reality, in physics takes the form of experiment and measurement. It is the second question that we wish to consider here, as applied to quantum mechanics.

Whatever the meaning assigned to the term *complete*, the following requirement for a complete theory seems to be a necessary one: *every element of the physical reality must have a counterpart in the physical theory*. We shall call this the condition of completeness. The second question is thus easily answered, as soon as we are able to decide what are the elements of the physical reality.

The elements of the physical reality cannot be determined by *a priori* philosophical considerations, but must be found by an appeal to results of experiments and measurements. A comprehensive definition of reality is, however, unnecessary for our purpose. We shall be satisfied with the following criterion, which we regard as reasonable. *If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.* It seems to us that this criterion, while far from exhausting all possible ways of recognizing a physical reality, at least provides us with one

such way, whenever the conditions set down in it occur. Regarded not as a necessary, but merely as a sufficient, condition of reality, this criterion is in agreement with classical as well as quantum-mechanical ideas of reality.

To illustrate the ideas involved let us consider the quantum-mechanical description of the behavior of a particle having a single degree of freedom. The fundamental concept of the theory is the concept of *state*, which is supposed to be completely characterized by the wave function ψ, which is a function of the variables chosen to describe the particle's behavior. Corresponding to each physically observable quantity A there is an operator, which may be designated by the same letter.

If ψ is an eigenfunction of the operator A, that is, if

$$\psi' \equiv A\psi = a\psi, \tag{1}$$

where a is a number, then the physical quantity A has with certainty the value a whenever the particle is in the state given by ψ. In accordance with our criterion of reality, for a particle in the state given by ψ for which Eq. (1) holds, there is an element of physical reality corresponding to the physical quantity A. Let, for example,

$$\psi = e^{(2\pi i/h)p_0 x}, \tag{2}$$

where h is Planck's constant, p_0 is some constant number, and x the independent variable. Since the operator corresponding to the momentum of the particle is

$$p = (h/2\pi i)\partial/\partial x, \tag{3}$$

we obtain

$$\psi' = p\psi = (h/2\pi i)\partial\psi/\partial x = p_0\psi. \tag{4}$$

Thus, in the state given by Eq. (2), the momentum has certainly the value p_0. It thus has meaning to say that the momentum of the particle in the state given by Eq. (2) is real.

On the other hand if Eq. (1) does not hold, we can no longer speak of the physical quantity A having a particular value. This is the case, for example, with the coordinate of the particle. The operator corresponding to it, say q, is the operator of multiplication by the independent variable. Thus,

$$q\psi = x\psi \neq a\psi. \tag{5}$$

In accordance with quantum mechanics we can only say that the relative probability that a measurement of the coordinate will give a result lying between a and b is

$$P(a, b) = \int_a^b \bar{\psi}\psi dx = \int_a^b dx = b - a. \tag{6}$$

Since this probability is independent of a, but depends only upon the difference $b-a$, we see that all values of the coordinate are equally probable.

A definite value of the coordinate, for a particle in the state given by Eq. (2), is thus not predictable, but may be obtained only by a direct measurement. Such a measurement however disturbs the particle and thus alters its state. After the coordinate is determined, the particle will no longer be in the state given by Eq. (2). The usual conclusion from this in quantum mechanics is that *when the momentum of a particle is known, its coordinate has no physical reality.*

More generally, it is shown in quantum mechanics that, if the operators corresponding to two physical quantities, say A and B, do not commute, that is, if $AB \neq BA$, then the precise knowledge of one of them precludes such a knowledge of the other. Furthermore, any attempt to determine the latter experimentally will alter the state of the system in such a way as to destroy the knowledge of the first.

From this follows that either (1) *the quantum-mechanical description of reality given by the wave function is not complete* or (2) *when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality.* For if both of them had simultaneous reality—and thus definite values—these values would enter into the complete description, according to the condition of completeness. If then the wave function provided such a complete description of reality, it would contain these values; these would then be predictable. This not being the case, we are left with the alternatives stated.

In quantum mechanics it is usually assumed that the wave function *does* contain a complete description of the physical reality of the system in the state to which it corresponds. At first

sight this assumption is entirely reasonable, for the information obtainable from a wave function seems to correspond exactly to what can be measured without altering the state of the system. We shall show, however, that this assumption, together with the criterion of reality given above, leads to a contradiction.

2.

For this purpose let us suppose that we have two systems, I and II, which we permit to interact from the time $t=0$ to $t=T$, after which time we suppose that there is no longer any interaction between the two parts. We suppose further that the states of the two systems before $t=0$ were known. We can then calculate with the help of Schrödinger's equation the state of the combined system I+II at any subsequent time; in particular, for any $t>T$. Let us designate the corresponding wave function by Ψ. We cannot, however, calculate the state in which either one of the two systems is left after the interaction. This, according to quantum mechanics, can be done only with the help of further measurements, by a process known as the *reduction of the wave packet*. Let us consider the essentials of this process.

Let $a_1, a_2, a_3, \cdots$ be the eigenvalues of some physical quantity A pertaining to system I and $u_1(x_1), u_2(x_1), u_3(x_1), \cdots$ the corresponding eigenfunctions, where x_1 stands for the variables used to describe the first system. Then Ψ, considered as a function of x_1, can be expressed as

$$\Psi(x_1, x_2) = \sum_{n=1}^{\infty} \psi_n(x_2) u_n(x_1), \tag{7}$$

where x_2 stands for the variables used to describe the second system. Here $\psi_n(x_2)$ are to be regarded merely as the coefficients of the expansion of Ψ into a series of orthogonal functions $u_n(x_1)$. Suppose now that the quantity A is measured and it is found that it has the value a_k. It is then concluded that after the measurement the first system is left in the state given by the wave function $u_k(x_1)$, and that the second system is left in the state given by the wave function $\psi_k(x_2)$. This is the process of reduction of the wave packet; the wave packet given by the infinite series (7) is reduced to a single term $\psi_k(x_2)u_k(x_1)$.

The set of functions $u_n(x_1)$ is determined by the choice of the physical quantity A. If, instead of this, we had chosen another quantity, say B, having the eigenvalues $b_1, b_2, b_3, \cdots$ and eigenfunctions $v_1(x_1), v_2(x_1), v_3(x_1), \cdots$ we should have obtained, instead of Eq. (7), the expansion

$$\Psi(x_1, x_2) = \sum_{s=1}^{\infty} \varphi_s(x_2) v_s(x_1), \tag{8}$$

where φ_s's are the new coefficients. If now the quantity B is measured and is found to have the value b_r, we conclude that after the measurement the first system is left in the state given by $v_r(x_1)$ and the second system is left in the state given by $\varphi_r(x_2)$.

We see therefore that, as a consequence of two different measurements performed upon the first system, the second system may be left in states with two different wave functions. On the other hand, since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system. This is, of course, merely a statement of what is meant by the absence of an interaction between the two systems. Thus, *it is possible to assign two different wave functions* (in our example ψ_k and φ_r) *to the same reality* (the second system after the interaction with the first).

Now, it may happen that the two wave functions, ψ_k and φ_r, are eigenfunctions of two noncommuting operators corresponding to some physical quantities P and Q, respectively. That this may actually be the case can best be shown by an example. Let us suppose that the two systems are two particles, and that

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} e^{(2\pi i/h)(x_1 - x_2 + x_0)p} dp, \tag{9}$$

where x_0 is some constant. Let A be the momentum of the first particle; then, as we have seen in Eq. (4), its eigenfunctions will be

$$u_p(x_1) = e^{(2\pi i/h)px_1} \tag{10}$$

corresponding to the eigenvalue p. Since we have here the case of a continuous spectrum, Eq. (7) will now be written

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \psi_p(x_2) u_p(x_1) dp, \tag{11}$$

where

$$\psi_p(x_2) = e^{-(2\pi i/h)(x_2 - x_0)p}. \tag{12}$$

This ψ_p however is the eigenfunction of the operator

$$P = (h/2\pi i)\partial/\partial x_2, \tag{13}$$

corresponding to the eigenvalue $-p$ of the momentum of the second particle. On the other hand, if B is the coordinate of the first particle, it has for eigenfunctions

$$v_x(x_1) = \delta(x_1 - x), \tag{14}$$

corresponding to the eigenvalue x, where $\delta(x_1 - x)$ is the well-known Dirac delta-function. Eq. (8) in this case becomes

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \varphi_x(x_2) v_x(x_1) dx, \tag{15}$$

where

$$\varphi_x(x_2) = \int_{-\infty}^{\infty} e^{(2\pi i/h)(x - x_2 + x_0)p} dp$$

$$= h\delta(x - x_2 + x_0). \tag{16}$$

This φ_x, however, is the eigenfunction of the operator

$$Q = x_2 \tag{17}$$

corresponding to the eigenvalue $x + x_0$ of the coordinate of the second particle. Since

$$PQ - QP = h/2\pi i, \tag{18}$$

we have shown that it is in general possible for ψ_k and φ_r to be eigenfunctions of two noncommuting operators, corresponding to physical quantities.

Returning now to the general case contemplated in Eqs. (7) and (8), we assume that ψ_k and φ_r are indeed eigenfunctions of some noncommuting operators P and Q, corresponding to the eigenvalues p_k and q_r, respectively. Thus, by measuring either A or B we are in a position to predict with certainty, and without in any way disturbing the second system, either the value of the quantity P (that is p_k) or the value of the quantity Q (that is q_r). In accordance with our criterion of reality, in the first case we must consider the quantity P as being an element of reality, in the second case the quantity Q is an element of reality. But, as we have seen, both wave functions ψ_k and φ_r belong to the same reality.

Previously we proved that either (1) the quantum-mechanical description of reality given by the wave function is not complete or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality. Starting then with the assumption that the wave function does give a complete description of the physical reality, we arrived at the conclusion that two physical quantities, with noncommuting operators, can have simultaneous reality. Thus the negation of (1) leads to the negation of the only other alternative (2). We are thus forced to conclude that the quantum-mechanical description of physical reality given by wave functions is not complete.

One could object to this conclusion on the grounds that our criterion of reality is not sufficiently restrictive. Indeed, one would not arrive at our conclusion if one insisted that two or more physical quantities can be regarded as simultaneous elements of reality *only when they can be simultaneously measured or predicted*. On this point of view, since either one or the other, but not both simultaneously, of the quantities P and Q can be predicted, they are not simultaneously real. This makes the reality of P and Q depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this.

While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible.

Entanglement Temperature and Entanglement Entropy of Excited States

Gabriel Wong,[1] Israel Klich,[1] Leopoldo A. Pando Zayas,[2] and Diana Vaman[1]

[1]*Department of Physics, University of Virginia, Box 400714, Charlottesville, Virginia 22904, USA*
[2]*Michigan Center for Theoretical Physics, The University of Michigan, Ann Arbor, MI 48109, USA*

(Dated: August 29, 2013)

We derive a general relation between the ground state entanglement Hamiltonian and the physical stress tensor within the path integral formalism. For spherical entangling surfaces in a CFT, we reproduce the *local* ground state entanglement Hamiltonian derived by Casini, Huerta and Myers. The resulting reduced density matrix can be characterized by a spatially varying "entanglement temperature." Using the entanglement Hamiltonian, we calculate the first order change in the entanglement entropy due to changes in conserved charges of the ground state, and find a local first law-like relation for the entanglement entropy. Our approach provides a field theory derivation and generalization of recent results obtained by holographic techniques. However, we note a discrepancy between our field theoretically derived results for the entanglement entropy of excited states with a non-uniform energy density and current holographic results in the literature. Finally, we give a CFT derivation of a set of constraint equations obeyed by the entanglement entropy of excited states in any dimension. Previously, these equations were derived in the context of holography.

I. INTRODUCTION AND MOTIVATION

For a given quantum state of a many-body system with density matrix ρ, measurements of observables O_A supported inside a spatial subregion A are determined by the reduced density matrix ρ_A, defined by

$$\mathrm{Tr}(\rho O_A) = \mathrm{Tr}_A(\rho_A O_A). \tag{I.1}$$

The relation above is defined to hold for all operators O_A in A. It follows that $\rho_A = \mathrm{Tr}_B(\rho)$, where the trace is taken over the complement $B = A^c$. Since an observer in A has no direct access to degrees of freedom in B, he/she suffers a loss of information that can be quantified by the entanglement entropy:

$$S_A = -\mathrm{Tr}_A(\rho_A \ln \rho_A). \tag{I.2}$$

S_A provides a measure of the entanglement between A and B, since increasing the entanglement between A and B will increase the loss of information upon restriction to A. The study of entanglement entropy was originally motivated by attempts to interpret black hole entropy as information loss by an observer restricted to the outside of the event horizon [1]. More recently, entanglement entropy has become an important tool in condensed matter physics, where it plays a role as a diagnostic of many body states. Indeed, the scaling of entanglement entropy characterizes the amenability of systems to numerical simulations such as the density matrix renormalization algorithm (DMRG) in 1d, and the nature of the challenge in higher dimensions. An important class of applications of entanglement entropy studies are topological states. Such states have no local observables which reveal their nature, and thus the entanglement entropy may be used in such a situation where no obvious way exists to identify the topological order. For example in [2] the interplay of DMRG and entanglement entropy on a torus has been used to identify the nature of topological degeneracy.

While entanglement entropy provides an important measure of entanglement, the reduced density matrix ρ_A is a more fundamental object. In particular, the study of the entanglement spectrum, i.e. the eigenvalues of ρ_A, has picked up pace as it has been recognized as a tool for probing topological order in a more detailed way. For example, the relation between the entanglement entropy of Quantum Hall wave functions and the edge theory associated with such states has been elaborated in [3–6]. Entanglement spectrum also holds a direct relation to the gluing function as well as the gapless edge modes in topological insulators [7, 8]. These remarkable relations between a bulk property and edge physics highlight the wealth of information encoded in the entanglement spectrum.

As stated above, the entanglement spectrum of quantum systems may reveal a lot about their nature. Even more detailed information is available if one knows the actual eigenstates of ρ_A. Since any ρ_A is Hermitean and positive semidefinite, it may be expressed as:

$$\rho_A = e^{-H_A} \tag{I.3}$$

for some Hermitean operator H_A. If H_A is known, the detailed study of ρ_A follows immediately to the exploration of H_A. Unfortunately, in most cases H_A does not offer a particular simplification or advantage as it is in general a highly nonlocal operator.

However, in particular special cases H_A may become local and simple enough to be used for calculations. The prime example for such a situation has arisen as a result of studies of Hawking and Unruh radiation. According to the Bisognano-Wichmann theorem [9, 10], the causal evolution of a quantum field theory where A is taken to be a half space may be described by a modular operator which is generated by a Lorentz boost. The Minkowski ground state in a causal wedge is then shown to satisfy a Kubo-Martin-Schwinger condition with respect to the boost, establishing H_A as the generator of Lorentz boost.

Recently this result was extended by Casini, Huerta and Myers [11]. For a spherical region A in a CFT, they find that the entanglement Hamiltonian may be written explicitly in a local form using the physical energy density T_{00}:

$$H_A = \int_A \beta(x) T_{00}(x). \tag{I.4}$$

In this paper, we use the locality property of entanglement Hamiltonians such as (I.4) to compute the entanglement entropy of excited states.

The starting point of our story is an elementary derivation of the above formula using the representation of the ground state reduced density matrix $\langle\phi|\rho_A|\phi'\rangle$ as a Euclidean Path integral integral with boundary conditions for the fields ϕ and ϕ' along the cut at A [12].

Deferring the explicit derivation to section II, let us first discuss the basic idea. Treating ρ_A as a propagator, we derive the expression (I.4) by performing the path integral along a Euclidean "time" s that evolves the upper edge of A to the bottom. The resulting path integral may be expressed as:

$$\rho_A = Z_A^{-1} T \exp\{-\int_{s_i}^{s_f} K(s) ds\}, \tag{I.5}$$

where T denotes "time" ordering in s and K is the quantum operator generating s evolution.

If the path integral of our theory is invariant under translations in s, then K is a conserved charge independent of "entanglement time" s. Hence:

$$\rho_A = \exp\left(-(s_f - s_i)K\right). \tag{I.6}$$

A well studied situation is the case where the theory is rotationally invariant, and $A = x^1 > 0$ is a half space. Taking

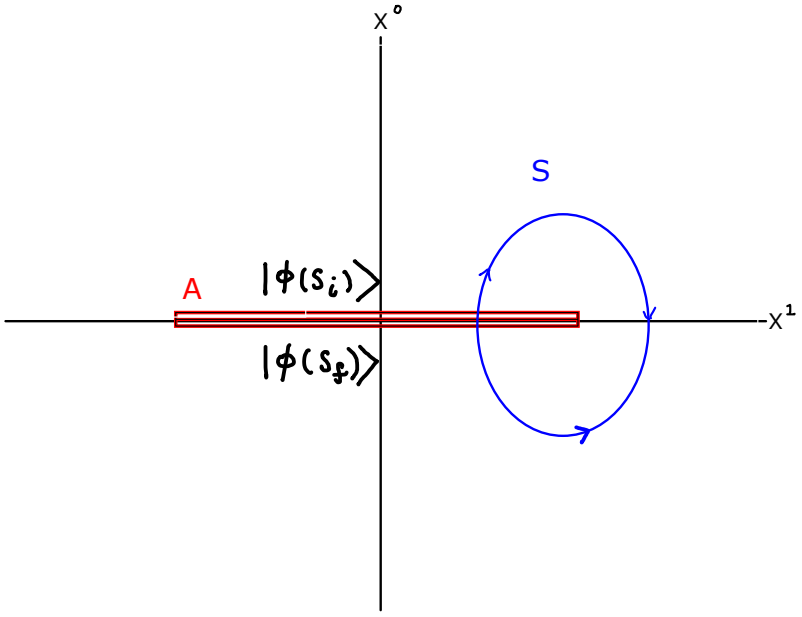


FIG. I.1: **Evaluating ρ_A along Euclidean time s**

s to be the angular variable on the $x^1, x^0 = t_E$ plane, we find the standard result that K is the angular new operator (or the boost generator in Minkowski signature) [13].

From a more general perspective, K can be viewed as a Killing energy that can be written in terms of the energy momentum tensor. For any constant s slice Σ we can write

$$K = \int_\Sigma T_{ab} k^a d\Sigma^b, \qquad H_A = (s_f - s_i)K, \tag{I.7}$$

where $k^a = \frac{dx^a}{ds}$ is the Killing vector for the boost and $\{x^a\}$ is a set of flat space coordinates. Choosing to evaluate K on $\Sigma = A$ we find $k^a \sim \delta^a_0$ and $\Sigma^a \sim \delta^a_0$, which reproduces the relation (I.4).

Given a spherically symmetric region A' in a Euclidean CFT of any dimension, we will determine the entanglement Hamiltonian for $\rho_{A'}$ by making use of a conformal map u taking A to A', which induces a mapping $\rho_A \to \rho_{A'} = U\rho_A U^{-1}$ [1]. The vector field $k'^a = \frac{dx^a}{ds'}$ for the new entanglement time s', is just the image of k under u. Thus, the entanglement Hamiltonian for A' is given by (I.4) with

$$\beta(x) = 2\pi k'^0(x), \tag{I.8}$$

where $x \in A$ and the factor of 2π is simply $s_f - s_i$. We will interpret $\beta(x)$ as a local "entanglement" temperature, that is determined by the shape of A and the background geometry of the CFT. In this interpretation, equation (I.4) resembles a density matrix for the original, physical system in local thermal equlibrium with temperature $\beta(x)$. The entanglement entropy is the thermal entropy of this system. It must be emphasized, however, that the appearance of $\beta(x)$ does not correspond to a "real" temperature in the sense that all inertial observers will find that local observables are at their vacuum values in accordance with (I.1)[2]. However, the point of view of a local "entanglement temperature" is appealing: indeed $\beta(x)$ must vanish at boundary of the region, signaling a high effective temperature close to the boundary. This behavior may be understood as the statement that the degrees of freedom close to the boundary are the ones most entangled with the external region, and thus have a larger entropy.

Consistent with this interpretation, we have checked that for two dimensional CFT's in various backgrounds with central charge c, the ground state entanglement entropy can be obtained by integrating the equilibrium thermal entropy per unit length

$$\frac{dS_{thermal}}{dx} = \frac{c\pi}{3\beta(x)} \tag{I.9}$$

over the region A using (I.8). Moreover, for excited states $\beta(x)$ relates the increase in entanglement entropy to an increase in energy inside A via a *local* first law-like equation:

$$d\delta S_A(x) = \beta(x) Tr_A(\delta\rho_A T_{00}) dx, \tag{I.10}$$

Here $\frac{d\delta S_A}{dx}(x)$ is the local entanglement entropy density[3] relative to the ground state and $\delta\rho_A$ is the variation in the reduced density matrix due to the increase in energy. To first order in $\delta\rho_A$, the total increase in entanglement entropy is obtained by integrating (I.10) over A.

Under a *general* variation of the ground state $\rho_A \to \rho_A + \delta\rho_A$ we find that the first order change in entanglement entropy is

$$\delta S_A = Tr_A(\delta\rho_A H_A). \tag{I.11}$$

For ground states with other conserved charges Q_a that preserve conformal invariance (e.g. momentum in 1+1 D) , the corresponding charge densities q_a and the associated chemical potentials μ_a will appear in the form

$$H_A = \int_A \beta(x)(T_{00} - \mu_a q_a), \tag{I.12}$$

leading to a generalized first law:

$$d\delta S_A(x) = \beta(x)\delta\langle T_{00}\rangle dx - \beta(x)\mu_a \delta\langle q_a\rangle dx. \tag{I.13}$$

While preparing this manuscript, a paper [15] was posted where a set of constraint equations for δS_A and an expression for "entanglement density" were derived using AdS/CFT. In section (VI) we provide a CFT derivation of those results in two spacetime dimensions and generalize the constraint equations to arbitrary dimensions[4]. We will also comment on the relation between our results and those in calculations in [17] and [18].

---

[1] This is essentially a Euclidean version of the arguments in [11].

[2] However, non-inertial observers whose proper time coincides with s will observe thermal radiation due to the local temperature [13].

[3] This is not to be confused with the "entanglement density", introduced in [14] and discussed later in this paper.

[4] The constraint equation was recently generalized to holographic CFT's in 3 space-time dimensions in [16]

II. PATH INTEGRAL DERIVATION OF THE ENTANGLEMENT HAMILTONIAN

Consider a Euclidean QFT on a manifold M and some spatial region A. The path integral expression for the reduced density matrix on A is similar to the propagator of the theory except that the initial and final states live on the upper and lower edge of a branch cut defined along A. Thus, to switch to a canonical description, it is natural to choose a foliation of M by constant s-slices $\Sigma(s)$ such that the initial/final slice at "time"(s_i, s_f) lie on the branch cut (see Fig. I.1). The manifold M is then parametrized by coordinates (s, y^a) where y^a are coordinates on Σ . The reduced density matrix on A in the Schrödinger picture is

$$\langle\phi_0(s_f)|\rho_A|\phi_0'(s_i)\rangle = \int D[\phi]e^{-\mathcal{S}[\phi]}\delta[\phi(s_f)-\phi_0(s_f)]\delta[\phi(s_i)-\phi_0(s_i)], \tag{II.1}$$

where $\mathcal{S}[\phi]$ is the action functional. To find the entanglement Hamiltonian, we divide the "time" interval $[s_i, s_f]$ into small steps $[s_{n+1}, s_n]$ of size Δs and consider a discretization of the path integral in (II.1). For notational simplicity we will write $\rho_A[s_{n+1}, s_n] = \langle\phi(s_{n+1})|\rho_A|\phi(s_n)\rangle$, so that

$$\langle\phi_0(s_f)|\rho_A|\phi_0'(s_i)\rangle = \int d[\phi(s_{N-1})]...d[\phi(s_2)]\rho_A[s_f, s_{N-1}]...\rho_A[s_{n+1}, s_n]...\rho_A[s_2, s_i]. \tag{II.2}$$

Next we will regard the matrix element $\rho_A[s_{n+1}, s_n]$ as a function of the final time s_f and final field configuration $\phi(s_{n+1}, y)$. We wish to show that this function satisfies a heat equation

$$\frac{\partial}{\partial s_{n+1}}\rho_A(s_{n+1}) = -K(s_{n+1})\rho_A(s_{n+1}) \tag{II.3}$$

and identify the operator $K(s_{n+1})$. For a given field configuration in the path integral we need to evaluate $\frac{\partial}{\partial s_{n+1}}\mathcal{S}[\phi(s_{n+1}, y), s_{n+1}]$ at *fixed* $\phi(s_{n+1}, y)$. One way of doing this is to keep the final time at s_{n+1}, but transform the background metric by a diffeomorphism that enlarges the proper size of the integration region. Explicitly we want a coordinate transformation $s \to s'(s)$ such that

$$\mathcal{S} + d\mathcal{S} = \int_{s_n}^{s_{n+1}+ds} ds \int_{\Sigma(s)} d^{d-1}y\mathcal{L}[g_{ab}, \phi] = \int_{s_n}^{s_{n+1}} ds' \int_{\Sigma(s')} d^{d-1}y\mathcal{L}[g_{ab} + dg_{ab}, \phi], \tag{II.4}$$

where $g_{ab}(s, y)$ is the metric on M. Therefore,

$$d\mathcal{S} = \int_{s_n}^{s_{n+1}} ds' \int_{\Sigma(s')} d^{d-1}y\frac{\delta\mathcal{L}}{\delta g_{ab}}dg_{ab}. \tag{II.5}$$

In a general coordinate system this transformation and the response of the path integral $\rho_($ $s_{n+1})$ is

$$x^a \to x^a = x^{a\prime} - \epsilon^a \tag{II.6}$$

$$d\rho(s_{n+1}) = -\tfrac{1}{2}\textstyle\int_{[s_n, s_{n+1}]\times\Sigma}\langle T_{ab}\rangle\nabla^{(a}\epsilon^{b)}\sqrt{g}d^dx = \int_{\Sigma(s_{n+1})}\langle T_{ab}\rangle\epsilon^b d\Sigma^a. \tag{II.7}$$

Here $\langle\rangle$ refers to the path integral average on $[s_n, s_{n+1}]$. In the last equality we assumed the *quantum* conservation law $\nabla^a\langle T_{ab}\rangle = 0$ and applied the divergence theorem; this means that T_{ab} includes a possible anomalous contribution due to the transformation of the Jacobian in the path integral measure. The coordinate transformation that will satisfy equation II.4 is

$$\epsilon^a = \frac{dx^a}{ds}f(s)ds, \tag{II.8}$$

where the function $f(s)$ smoothly goes from 0 to 1 as s goes from s_n to s_{n+1}. This is so that we do not change the lower endpoint of the s integration.

Defining

$$K(s_{n+1}) = \int_{\Sigma(s_{n+1})}\langle T_{ab}\rangle\frac{dx^b}{ds}d\Sigma^a, \tag{II.9}$$

we find

$$\frac{\partial}{\partial s_{n+1}}\rho_A[s_{n+1},s_n]=\int D[\phi]e^{-S[\phi]}(-K(s_{n+1}))=\langle\phi_0(s_{n+1})|-K(s_{n+1})\rho_A|\phi_0'(s_n)\rangle=-(\hat{K}\rho_A)(s_{n+1}). \tag{II.10}$$

The solution to this heat equation with initial condition $\rho_A(s_n)=0$ is $\rho_A[s_f,s_{N-1}]=\langle\phi(s_{n+1})|1-\Delta sK|\phi(s_n)\rangle$. Inserting this into equation (II.2) gives

$$\langle\phi_0(s_f)|\rho_A|\phi_0'(s_i)\rangle=\int\prod_{n=1}^{N-1}D[\phi(s_n)]\langle\phi(s_{n+1})|1-\Delta sK|\phi(s_n)\rangle \tag{II.11}$$

$$=\langle\phi_0(s_f)|T\exp\left(-\int_{s_i}^{s_f}K(s)ds\right)|\phi_0'(s_i)\rangle. \tag{II.12}$$

This is the most general form of the entanglement Hamiltonian in a QFT. Since equation (II.12) only depends on the geometric data provided by the vector field $\frac{dx^a}{ds}$ which in turn is determined by the region A, it represents a *universal* relation between the entanglement Hamiltonian and the quantum stress tensor.

To recover the local Entanglement Hamitonian (I.4), we consider regions A for which $s\to s+ds$ is a spacetime symmetry of the path integral (II.1) so that $K[s]$ is the corresponding conserved charge. Since K is independent of s, we can evaluate it on any time slice (say at s_i) and the time ordered product in (II.12) reduces to

$$\rho_A=\exp(-(s_f-s_i)K(s_i)). \tag{II.13}$$

Below we will show that $s\to s+ds$ is indeed a spacetime symmetry of the path integral if A is a half space in a rotationally invariant QFT or a spherical region in a CFT, and we will derive the corresponding local entanglement Hamiltonians. Here we would like to note that given a small deformation of the region A away from translational or spherical symmetry, one could perform a systematic expansion of equation (II.12) using the deformed entanglement Hamiltonian $K_0+\epsilon K_1$. To first order in ϵ this would just add a perturbation to the local entanglement Hamiltonian which is localized near the boundary of A. A similar strategy can be applied to deformations of the theory away from rotational or conformal invariance. We leave this for future work.

III. EXAMPLES OF LOCAL ENTANGLEMENT HAMILTONIANS

A. Entanglement Hamiltonians in 2D

To illustrate how to compute K and its entanglement entropy, we first review the case of a rotationally invariant QFT on $\mathbb{R}^2$ with A the region A being the half line $A=\{x^1>0\}$ [19]. Since A is mapped to itself by a 2π rotation, we choose s to be the angular coordinate on the Euclidean plane so that $\Sigma(s)$ are rays emanating from the origin as in Figure III.1.

Then

$$\frac{dx^a}{ds}\partial_a=x^1\partial_0-x^2\partial_1 \tag{III.1}$$

is a Killing vector field generating rotations of the plane. Since the path integral measure is assumed to be rotationally invariant, K is just the angular momentum [20]

$$K=\int_{\Sigma(s=0)}x^1T_{00}-x^0T_{01}=\int_A x^1T_{00}. \tag{III.2}$$

The entanglement Hamiltonian is given by equation (I.4) with the entanglement temperature

$$\beta=2\pi x^1. \tag{III.3}$$

Upon Wick rotating $s\to is$, the circular flow generated by K becomes hyperbolas representing the worldlines of uniformly accelerated observers, and $\beta(x)$ is the proper temperature they experience. Thus in Minkowski signature K is the boost generator. The form of the entanglement Hamiltonian implies that ρ_A represents an ensemble with the physical energy density T_{00} in local thermal equilibrium with local temperature $\beta(x)$; its entanglement entropy is therefore just the thermal entropy, obtained by integrating the thermal entropy density $\frac{dS_{thermal}}{dx}$ over A [20].

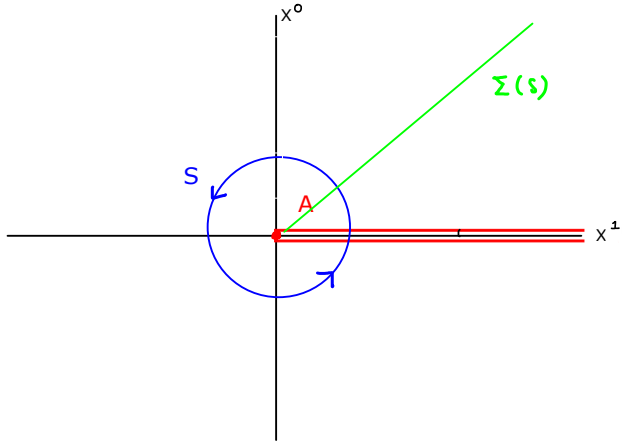


FIG. III.1: **Foliation of the Euclidean plane corresponding to angular quantization**

In particular, for a CFT with central charge c, it is well known that [12]

$$\frac{dS_{thermal}}{dx} = \frac{c\pi}{3\beta(x)} \tag{III.4}$$

so the entanglement entropy is

$$S_A = \int_\delta^L dx \frac{c\pi}{6x} = \frac{c}{6} Log \frac{L}{\delta}, \tag{III.5}$$

where we have introduced a UV and IR cutoff on A restricting the integration to $[\delta, L]$. The local temperature is higher near the boundary of A and diverges at $x = 0$ due to the zero of the vector field, which is also the singularity of the foliation defined by s. As a result, most of the contribution to the entanglement entropy arises from near the edge.

For a CFT on $\mathbb{R}^2$ we can easily generalize the previous results to an arbitrary interval $A' = [u, v]$. Let $z = x^1 + ix^0$ so that $\frac{dz}{ds} = iz$ is the rotational vector field appropriate to the region A discussed previously. The conformal map

$$z = -\frac{w-u}{w-v} \tag{III.6}$$

induces a transformation U on the reduced density matrices:

$$\rho_A \to \rho_{A'} = U \rho_A U^{-1}, \tag{III.7}$$

by transforming the boundary conditions of the path integral. The path integral measure is conformally invariant because there is no anomaly in flat space. Meanwhile, the vector field $\frac{dz}{ds'}$ is mapped to

$$\frac{dw}{ds'} = \frac{dw}{dz}\frac{dz}{ds'} = \frac{i(w-u)(w-v)}{u-v}. \tag{III.8}$$

It is clear that the periodic flow defined by this vector field will evolve $A' \to A'$. Moreover, the transformation $w \to w + \frac{dw}{ds'} ds'$ is a symmetry of the CFT on the w plane, because it can be decomposed into a combination of a conformal transformation between z and w, and an ordinary rotation on the z plane.

Thus, the entanglement Hamiltonian for $\rho_{A'}$ is

$$H_{A'} = \int_A 2\pi \frac{(y^1-u)(y^1-v)}{u-v} T_{00} dy, \tag{III.9}$$

where we defined $w = y^1 + iy^0$ and evaluated the integral along A for convenience. As before, the entanglement entropy is obtained from integrating $\frac{dS_{thermal}}{dx}$ using the entanglement temperature

$$\beta(y) = 2\pi \frac{(y^1-u)(y^1-v)}{u-v}. \tag{III.10}$$

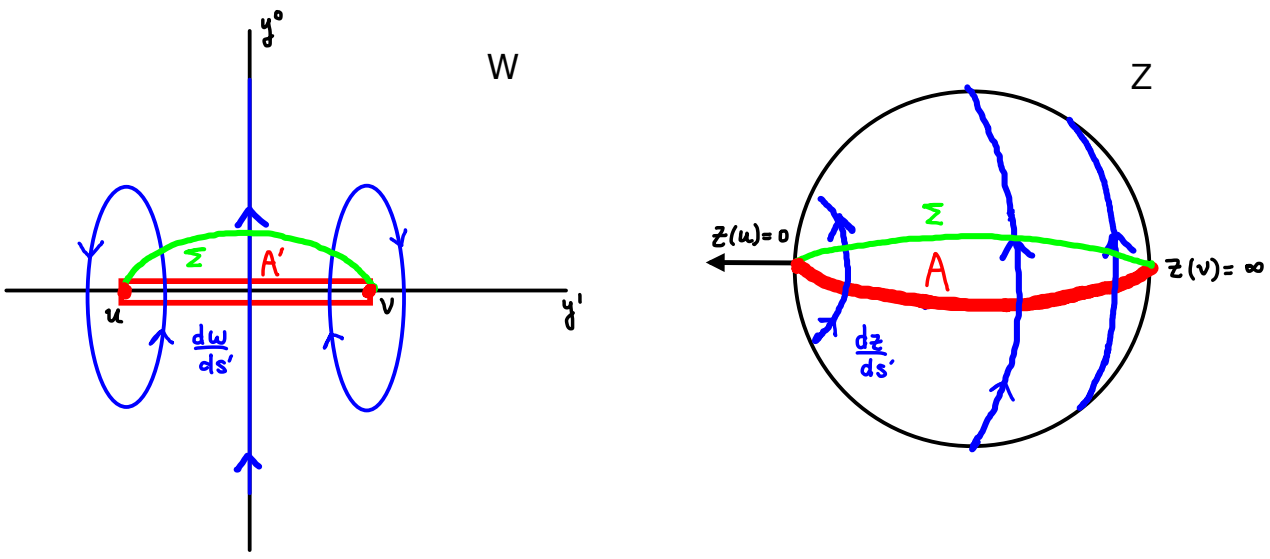


FIG. III.2: **A rotation on the z plane (represented as a Riemann sphere) is mapped to a conformal rotation on the w plane**

This gives

$$S_{A'} = \frac{c}{3} log \frac{v-u}{\delta}, \tag{III.11}$$

as expected[5].

For a CFT at finite temperature ($w \sim w + i\beta'$) or on a spatial circle ($w \sim w + L$), we can similarly derive the entanglement Hamiltonian by finding the conformal map from the z-plane to $\mathbb{R}\times S^1$ or $S^1\times\mathbb{R}$. Given $A' = [-l, l]\times\{0\}$, the conformal map and entanglement temperature for a CFT at the (ordinary) temperature β' is

$$z = \frac{-\exp\left(\frac{2\pi w}{\beta'}\right) + \exp\left(-\frac{2\pi l}{\beta'}\right)}{\exp\left(\frac{2\pi w}{\beta'}\right) - \exp\left(\frac{2\pi l}{\beta'}\right)}, \qquad \beta = 2\beta' \text{csch}\left(\frac{2l\pi}{\beta'}\right)\sinh\left(\frac{\pi(l-y)}{\beta'}\right)\sinh\left(\frac{\pi(l+y}{\beta'}\right). \tag{III.12}$$

The results for a CFT at finite size can be obtained from equation (III.12) by the substitution $\beta' \to iL$. Below we summarize the results for the entanglement temperature and entanglement entropy obtained by integrating the thermal entropy density in various CFT backgrounds.

TABLE I: Entanglement Temperature for $A = [-l, l]$ in different CFT backgrounds

CFT Background	Entanglement Temperature $\beta(y)$	Entanglement Entropy $S_A = \int_A \frac{\pi c}{3\beta(y)}$
Zero Temp. and infinite size: $M = \mathbb{R}^2$	$\frac{2\pi(l^2-y^2)}{2l}$	$\frac{c}{3}\ln\frac{l}{\delta}$
Finite temperature $\beta' : M = \mathbb{R}\times S^1$	$2\beta'\text{csch}\left(\frac{2l\pi}{\beta'}\right)\sinh\left(\frac{\pi(l-y)}{\beta'}\right)\sinh\left(\frac{\pi(l+y)}{\beta'}\right)$	$\frac{c}{3}\ln\left(\frac{\beta'}{\pi\delta}\sinh(\frac{2\pi l}{\beta'})\right)$
Finite Size $L : M = S^1\times\mathbb{R}$	$2L\text{csc}\left(\frac{2l\pi}{L}\right)\sin\left(\frac{\pi(l-y)}{L}\right)\sin\left(\frac{\pi(l+y)}{L}\right)$	$\frac{c}{3}\ln\left((\frac{L}{\pi\delta}\sin(\frac{2\pi l}{L})\right)$

[5] Note that even though $Tr_A(\rho_A log\rho_A)$ is invariant under the similarity transformation (III.7) of the reduced density matrix, we get a different result for the entanglement entropy of $\rho_{A'}$ because we have to transform the regularized boundary of A.

The results for the entanglement entropies were derived previously using the replica trick, [12, 21–23], and serve as a check on our results for the entanglement temperature and Hamiltonian.

B. Entanglement Hamiltonians in higher dimensions

Here we generalize the results of the previous section to spherical entangling surfaces in dimensions $d > 2$. As before, we first consider a rotationally invariant CFT on $\mathbb{R}^d$ with $A = \{x^1 > 0\}$. We choose polar coordinates on the x^1, x^0 plane $x^1 = z\cos(\frac{s}{l})$, $x^0 = z\sin(\frac{s}{l})$, so the flat metric is

$$d\tau^2 = (\frac{z}{l})^2 ds^2 + dz^2 + d\vec{x}^2. \tag{III.13}$$

At this point l is an arbitrary length parameter introduced to make s dimensionful. Then the result (III.2) for the entanglement Hamiltonian of ρ_A is still valid. Now we map $\mathbb{R}^d \to H^{d-1} \times S^1$, by multiplying the metric above by a conformal factor $(\frac{l}{z})^2$.

$$d\tau^2_{H^{d-1}\times S^1} = ds^2 + (\frac{l}{z})^2(dz^2 + d\vec{x}^2). \tag{III.14}$$

The H^{d-1} factor refers to hyperbolic space, which is the image of the half space A. Thus we see that ρ_A is transformed into a thermal density matrix $\rho_{H^{d-1}}$ on hyperbolic space. Since this conformal map does not change the original coordinates on $\mathbb{R}^d$, the vector field generated by the new entanglement Hamiltonian is just $\frac{\partial}{\partial s}$.

Now consider a new reduced density matrix $\rho_{A'}$ for a ball of radius l. We will obtain the entanglement Hamiltonian $H_{A'}$ by mapping $\rho_{H^{d-1}} \to \rho_{A'}$ as follows. First we choose coordinates (u, Ω_{d-2}, s) on $H^{d-1} \times S^1$ and spherical coordinates (r, Ω_{d-2}, t) on $\mathbb{R}^d$ so that the metrics are

$$d\tau^2_{H^{d-1}\times S^1} = ds^2 + R^2(du^2 + \sinh(u)^2 d\Omega^2_{d-2}), \tag{III.15}$$

$$d\tau^2_{R^d} = dt^2 + dr^2 + r^2 d\Omega^2_{d-2}. \tag{III.16}$$

Then, defining complex coordinates $\sigma = u + i\frac{s}{l}$ and $w = r + it$ on the respective two dimensional slices, we consider the mapping introduced in [24]

$$e^{-\sigma} = \frac{l - w}{l + w}. \tag{III.17}$$

This is an analogue of equation (III.6) mapping $\rho_{A'} \to \rho_{H^{d-1}}$. The entanglement vector field and entanglement Hamiltonian is

$$\frac{dw}{ds} = \frac{dw}{d\sigma}\frac{d\sigma}{ds} = i\frac{l^2 - r^2}{2l}, \qquad H_A = 2\pi \int_A \frac{l^2 - r^2}{2l} T_{00}. \tag{III.18}$$

This agrees with the result of [11], where a Minkowski signature version of the conformal mapping (III.17) was used to derive the entanglement Hamiltonian.

IV. CFT DERIVATION OF ENTANGLEMENT ENTROPY FOR EXCITED STATES

Consider a state $|\psi\rangle$ in a QFT in $\mathbb{R}^{1,d-1}$ with a density matrix $\rho^0 = |\psi\rangle\langle\psi|$. As in [17] we make a small perturbation $\rho = \rho^0 + \delta\rho$ and consider the entanglement entropy of a region A. Expanding to first order in $\delta\rho_A$ we find

$$S_A = -Tr_A(\rho_A \ln \rho_A) = -Tr_A(\rho^0_A \ln \rho^0_A) - Tr_A(\delta\rho^0_A \ln \rho^0_A) - Tr_A(\delta\rho_A), \tag{IV.1}$$

where $\delta\rho_A = Tr_B(\delta\rho)$. The normalization $Tr(\rho_A) = Tr_A(\rho^0_A) = 1$ implies $Tr(\delta\rho_A) = 0$, so the first order change in entanglement entropy due to the perturbation $\delta\rho$ is simply

$$\delta S_A = -Tr_A(\delta\rho_A \ln \rho^0_A) = Tr_A(\delta\rho_A H_A). \tag{IV.2}$$

Note that there is also a term proportional to $\mathrm{Tr}(\delta\rho)$ which vanishes due to the normalization $Tr(\rho) = 1$. When the state ρ^0 is the ground state, we will refer to δS_A as the *renormalized* entanglement entropy[6] [12]. It is just the

[6] This is only a first order approximation to the renormalized entropy, but we will just call it renormalized entropy for short.

increase in "entanglement energy" of the new state, measured according to the ground state entanglement Hamiltonian. However we emphasize that equation (IV.2) applies to an *arbitrary* deformation $\delta\rho$ for *any* initial state ρ^0. When the region A is a half space in a QFT or a spherical ball in a CFT, we can use the entanglement temperatures previously derived to obtain H_A for the ground state as in equation (I.4). From equation (IV.2) we have:

$$\delta S_A = Tr_A(\delta\rho_A \int_A \beta(x) T_{00}(x)) = \int_A \beta(x) Tr(\delta\rho T_{00}(x)) := \int_A \beta(x)\delta\langle T_{00}(x)\rangle \tag{IV.3}$$

In the second to last equality, we noted that the operator $T_{00}(x)$ is only being evaluated inside A so that $\delta\rho_A$ can be replaced with $\delta\rho$. Note that in (IV.2) the operator $\delta\rho_A$ and H_A are defined on a subregion A with boundaries, which implies boundary conditions have to be imposed at ∂A on their quantization. On the other hand, in (IV.3) the operator T_{00} is interpreted as the energy density quantized with the boundary conditions appropriate to the *whole* space; we have merely chosen to *evaluate* it inside A. These two interpretation must agree by the definition of the reduced density matrix. As a check, in appendix B we will show that for a particular excitation of a free scalar field with non-uniform energy density, (IV.3) and (IV.2) do indeed give the same result for δS_A.

When $\delta\langle T_{00}\rangle$ is spatially uniform[7] inside A, we can remove it from the integration, so that

$$\delta S_A = \beta_0 \delta\langle T_{00}\rangle Vol(A) := \beta_0 \delta E_A, \tag{IV.4}$$

where $\delta E_A = \delta\langle T_{00}\rangle \mathrm{Vol}(A)$ is the excitation energy inside region A, and β_0 is the average entanglement temperature inside A

$$\beta_0 = \frac{\int_A \beta(x)}{\mathrm{Vol}(A)}. \tag{IV.5}$$

When the region A has radius l, we find[8] $\beta_0 = \frac{2\pi}{d+1} l$ in agreement with the result of [18]. However, we note that the holographic results of [18] only strictly apply to nonabelian gauge theories with holographic duals, at large N and assuming a small region A (i.e. for small radius l), whereas our result is valid to order $O(\delta\rho)$ for any CFT and any radius l. We also note that there is a discrepancy between our results when $\delta\langle T_{00}\rangle$ is spatially varying. Given a state with $\delta\langle T_{00}\rangle = \sum_{n=0}^{\infty} a_n r^n$ in a $d > 2$ dimensional CFT[9], we find

$$\delta S_A = 2\pi \mathrm{Vol}(S^{d-2}) \sum_{n=0} \frac{a_n l^{d+n}}{(d+n)^2 - 1} \tag{IV.6}$$

which disagrees with the holographic calculation of the same quantity in equation (20) of [18]. In section IV, we will discuss the holographic version of eq IV.2 and speculate on a possible source of the discrepancy. As noted earlier, we have checked in appendix B that our results (IV.2) and (IV.3) are consistent for a *non-uniform* excitation of a free scalar field, where δS_A can be computed explicitly.

V. A GENERALIZED FIRST LAW FOR ENTANGLEMENT ENTROPY

Equation (IV.3) resembles a local first law of thermodynamics inside the region A:

$$d\delta S_A(x) = \beta(x)\delta\langle T_{00}(x)\rangle dx. \tag{V.1}$$

When other conserved charges are present, a generalization of equation (V.1) can be derived as follows. Consider a state at finite temperature T and with conserved charges Q_a that preserve conformal invariance and chemical potentials μ_a weighted with the following density matrix

$$\rho = \frac{\exp\left(-\frac{(H-\mu_a Q_a)}{T}\right)}{Z}. \tag{V.2}$$

---

[7] Since our entanglement Hamiltonian was derived for a CFT on $\mathbb{R}^d$, we will assume the energy density starts to die off somewhere outside A, in order for the energy to be finite.

[8] As already noted in [18], this is also consistent with the computation of δS_A for primary states of a two dimensional CFT which was performed in [25] via the replica trick.

[9] We will explain the restriction to $d > 2$ in the section VI.

After tracing over the complement of A we arrive at a path integral representation of ρ_A similar to the one given in equation (II.1), except that adding the charges has effectively shifted our Hamiltonian from H to $H' = H - \mu_a Q_a$. The corresponding shift in the energy density is $T'_{00} = T_{00} - \mu_a q_a$, where we introduced the charge densities q_a by $Q_a := \int_{space} q_a d^{d-1}x$. Going through the same path integral derivation as in section II, we would reproduce equation (I.4) with T_{00} replaced by T'_{00}. Under a deformation $\delta\rho$ that changes the charge densities and energy inside A, equation (V.1) now becomes

$$d\delta S_A(x) = \beta(x)\delta\langle T'_{00}(x)\rangle dx = \beta(x)\{\delta\langle T_{00}(x)\rangle dx - \mu_a \delta\langle q_a(x)\rangle dx\} \tag{V.3}$$

A simple way to check the above argument for the entanglement Hamiltonian leading to equation (V.3) is to consider a state $\rho \sim \exp[-\beta'(H - \mu P)]$ for a two dimensional CFT with total central charge c. In this case the conserved Virasoro charges are the Hamiltonian $H = L_0 + \overline{L_0} - \frac{c}{12}$ and momentum $P = L_0 - \overline{L_0}$. The entanglement Hamiltonian for an interval $A = [0, l]$ is

$$H_A = \int_0^l \beta(x)(T_{00} - \mu T_{01})dx = \int_0^l \beta(x)(1-\mu)T_{++} + \beta(x)(1+\mu)T_{--}, \tag{V.4}$$

where $T_{\pm\pm} = \frac{1}{2}(T_{00} \pm T_{01})$ are the right and left moving components of the stress tensor, and $\beta(x)$ is the entanglement temperature (III.12) for a CFT at finite temperature[10] β'. The operator in equation (V.4) is the sum of two *commuting* entanglement Hamiltonians corresponding to non-interacting ensembles at finite (ordinary) temperature $\beta_\pm = \beta'(1\pm\mu)$ and with energy density $T_{\pm\pm}$. Assuming that the left and right central charges are equal, each ensemble has an effective central charge of $\frac{c}{2}$. Thus the entanglement entropy is:

$$S_A = \frac{c}{6}\ln\left(\frac{\beta_+}{\pi\delta}\sinh(\frac{\pi l}{\beta_+})\right) + \frac{c}{6}\ln\left(\frac{\beta_-}{\pi\delta}\sinh(\frac{\pi l}{\beta_-})\right). \tag{V.5}$$

This agrees with the result of [26] obtained via the replica trick and holographic calculations.

VI. HOLOGRAPHIC DERIVATION AND DISCUSSION OF RELATED PAPERS

According to the holographic prescription of [27], the entanglement entropy for a state $|\psi\rangle$ in a region A of a d-dimensional CFT with a holographic dual gravity theory is

$$S_A = \frac{Area(\gamma_A)}{4G}, \tag{VI.1}$$

where γ is a minimal surface, anchored on ∂A, in the bulk spacetime representing the gravity dual of the corresponding CFT, G is the bulk Newton's constant. The geometry dual to the ground state in the CFT corresponds to pure AdS

$$d\tau^2 = (\frac{R}{z})^2(-dt^2 + dz^2 + r^2 d\Omega^2_{d-2}), \tag{VI.2}$$

and the minimal surface for $A = \{r = l\}$ is a half sphere extending into the bulk: $\gamma_A = \{r^2 = l^2 - z^2\}$.

For general excited states, it is difficult to find the exact bulk metric and compute the minimal surface. However, just as in the CFT computation of the previous section, a drastic simplification occurs if we consider only the first order deformation of the entanglement entropy, which is proportional to the variation of area functional :

$$\delta Area(\gamma_A) = \delta\int_{\gamma_A}\sqrt{g} = \int_{\gamma_A}\delta\sqrt{g}. \tag{VI.3}$$

In the last equality, we observed that the area variation due to the deformation of the surface γ_A vanishes by the definition of a minimal surface. Thus, the area variation is entirely due to the change in the metric, and there is no need to solve for the minimal surface in the new geometry. Comparing this equation to (IV.4), we see that $\delta\rho_A$ corresponds

---

[10] Technically, to get a discrete spectrum for P we should put the CFT on a spatial S^1 of length L. Here we will assume $\beta' >> L$, so that we can ignore the periodicity along L in computing the entanglement temperature.

to the deformation of the metric while H_A corresponds to the ground state minimal surface. The second fact is less obvious from the usual AdS/CFT correspondence, but it is consistent with ideas proposed in [11]. In reference [11] it was shown that for spherical regions A, there exists a foliation of AdS by hyperbolic slices $\mathcal{H} = H^{d-1} \times R$ such that one of the slices is a causal horizon γ'_Athat is anchored on ∂A. Since a horizon is also a minimal surface, we can identify $\gamma'_A = \gamma_A$. The new foliation of AdS is dual to a CFT on the boundary slice $\mathcal{H}$, which is in a thermal state that is conformally related to ρ_A . It is thus tempting to identify the foliation of AdS and the associated horizon γ_A with the reduced density matrix ρ_A and therefore H_A.

As in [18][11] we consider an excited state with energy density [12] $\langle T_{00}\rangle = \frac{dR^{d-1}m}{16\pi G}$. As established in ref. [29], the holographic stress tensor associated with this energy density and the boundary metric determines the *asymptotic* form of the bulk metric near the boundary at $z \sim 0$ to be:

$$d\tau^2 = (\frac{R}{z})^2(-g^{-1}(z)dt^2 + g(z)dz^2 + r^2 d\Omega^2_{d-2}), \quad \text{with} \quad g(z) = 1 + mz^d + ... \tag{VI.4}$$

where the ellipsis denotes higher order terms in z. In this approximation, the first order variation of the entanglement entropy for spherical regions A is

$$\begin{aligned} \frac{\delta S'_A}{\delta m}\Big|_{m=0} \delta m &= \frac{\delta Area(\gamma_A)}{4G}\Big|_{m=0} = R^{d-1}\Omega_{d-2}\int_0^l \frac{r(z)^{d-2}}{z^{d-1}}\delta\sqrt{g(z)+r'(z)^2} \\ &= \beta_0 \delta E_A, \end{aligned} \tag{VI.5}$$

where we evaluated the integral along the half sphere $r^2 = l^2 - z^2$ corresponding to the ground state at $m = 0$, $\beta_0 = \frac{2\pi}{d+1}l$, and δE_A is defined as in the section IV. The notation $\delta S'_A$ is a reminder of the additional approximation due to the expansion (VI.4), where sub-leading in terms in z were dropped. However, in this case, this approximation (truncation) leads to a result which agrees with the field theoretic one in eq. (IV.5)

Next, we consider a a non-uniform state with energy density $\langle T_{00}\rangle = \frac{dR^{d-1}m}{16\pi G}\sum_{n\geq 0} c_n r^n$ in a $d > 2$ dimensional CFT. Note that this state is not allowed $d = 2$ spacetime dimensions, because the energy density has to satisfy a wave equation, as explained later in this section. The dual metric has the same form as in (VI.4) with

$$g(z) = 1 + mz^d \sum_{n\geq 0} c_n r^n + \ldots, \tag{VI.6}$$

using (VI.3) we find:

$$\delta S'_A = \frac{ml^d R^{d-1}\text{Vol}(S^{d-2})}{8G}\sum_{n\geq 0}\frac{c_n l^n}{1+d+n}. \tag{VI.7}$$

The above expression reproduces and generalizes the results in [18], without recourse to a an explicit evaluation of the minimal surfaces. This time, we note that above $\delta S'_A$ differs from our result (IV.6) for the entropy of a sphere, although both are supposed to represent entropy of a system with the same non-uniform energy density.

In [18], use of equation (VI.4) was justified by taking the small region limit, that is, the $l \to 0$ limit in which γ_A approaches the $z = 0$ boundary. However, neglect of higher order terms in z, while not affecting the energy density $\langle T_{00}\rangle$, may affect the computed entropy. For example, adding a correction of the form $mz^{d+k}r^\mu$ will yields, using (VI.3), a contribution proportional to $l^{d+k+\mu}$ to the holographic entropy. Neglect of such terms may be the reason that our results agree with those of [18] only for the case of uniform energy density. In this way, our result provides an easy consistency check for the $z \to 0$ limit metric used in holographic calculations.

A. Dynamical equations for entanglement entropy and entanglement density

While this project was being completed, we noticed a recent paper [15] where a set of dynamical equations were derived for δS_A in the case of time dependent excited states by using the holographic formula (VI.1). In $d = 2$

---

[11] see also [28] for an extension of results in [18]

[12] To facilitate comparisons with [18], in this section we write $\delta\langle T_{00}\rangle = \langle T_{00}\rangle$, with the understanding that the energy density in the latter expression is normal ordered so as to subtract the vacuum energy. Note that there is a typo in eq. (2) of [18] where d was replaced with $d - 1$.

spacetime dimensions they are:

$$(\partial_t^2 - \partial_\xi^2)\delta S_A(\xi, l, t) = 0 \tag{VI.8}$$

$$(\frac{\partial_l^2}{4} - \frac{\partial_t^2}{4} - \frac{1}{2l^2})\delta S_A(\xi, l, t) = 0 \tag{VI.9}$$

where $A = [\xi - l, \xi + l]$. In the holographic setting these equations arose from solving Einstein's equations perturbatively to determine the evolution of the metric for the excited state. Here we will provide a simple field theoretic derivation of these equations. First note that in terms of the variable $x' = x - \xi$, the renormalized entanglement entropy for a CFT on a plane is

$$\delta S_A = 2\pi \int_{-l}^{l} dx' \frac{l^2 - x'^2}{2l^2} \langle T_{00} \rangle (x' + \xi, t), \tag{VI.10}$$

so the entanglement temperature is independent of ξ. Thus,

$$(\partial_t^2 - \partial_\xi^2)\delta S_A(\xi, l, t) = 2\pi \int_{-l}^{l} dx' \frac{l^2 - x'^2}{2l^2} (\partial_t^2 - \partial_\xi^2)\langle T_{00} \rangle (x' + \xi, t) = 0, \tag{VI.11}$$

where in the last equality we used the fact that in $d = 2$ the conservation of the energy momentum tensor combined with its tracelessness imply that $T_{00} = T_{++} + T_{--}$ is a sum of left and right movers, and therefore satisfy the wave equation. The second equation (VI.8)can be obtained straightforwardly by applying the differential operator to (VI.10) and integrating by parts using $\partial_t^2 T_{00} = -\partial_\xi^2 T_{00} = -\partial_{x'}^2 T_{00}$. As in [15], we can also generalize and (VI.8) to the case when we couple an operator $O(x, t)$ to a source $J(x, t)$ so that our physical Hamiltonian is deformed to $H' = H - \int JOd^{d-1}x$. Provided that $O(x, t)$ preserves conformal symmetry, this deformation changes the ground state Hamiltonian by deforming the energy density $T_{00} \to T'_{00} = T_{00} - JO$ in I.4. The equations (VI.8) are now modified by source terms that arise form the differential operators hitting $J(x, t)O(x, t)$. Thus

$$(\partial_t^2 - \partial_\xi^2)\delta S_A(\xi, l, t) = \int_{-l}^{l} \beta(x', l)(\partial_t^2 - \partial_\xi^2)(J(x' + \xi, t)\langle O(x' + \xi, t)\rangle_J), \tag{VI.12}$$

$$(\frac{\partial_l^2}{4} - \frac{\partial_t^2}{4} - \frac{1}{2l^2})\delta S_A(\xi, l, t) = -\int_{-l}^{l} \beta(x', l)\frac{\partial_t^2}{4}(J(x' + \xi, t)\langle O(x' + \xi, t)\rangle_J), \tag{VI.13}$$

with

$$\beta(x', l) = 2\pi \frac{l^2 - x'^2}{2l^2}. \tag{VI.14}$$

To facilitate a comparison with the result of [15], we take the Fourier transform of $\langle O(x' + \xi, t)\rangle_J$ and make explicit the dependence of $J(x' + \xi, t)$ on $\langle O(k_1, w_1)\rangle_J$:

$$\langle O(x, t)\rangle_J = \int d\omega_1 \int dk_1 \langle O(k_1, \omega_1)\rangle_J e^{i(k_1\xi + \omega_1 t)} e^{ik_1 x'}, \tag{VI.15}$$

$$J(x' + \xi, t) = \int d\omega_2 \int dk_2 f(k_2, \omega_2)\langle O(k_2, \omega_2)\rangle_J e^{i(k_2\xi + \omega_2 t)} e^{ik_2 x'}. \tag{VI.16}$$

Above we chose the source J corresponding to the perturbation of the bulk scalar given in equation (3.17) of [15]. Inserting these in (VI.12) and integrating over x' gives equations of the form

$$(\partial_t^2 - \partial_\xi^2)\delta S_A(\xi, l, t) = \int d\omega_1 \int d\omega_2 \int dk_1 \int dk_2 F(k_1, k_2, \omega_1, \omega_2, l)\langle O(k_1, \omega_2)\rangle_J \langle O(k_2, \omega_2)\rangle_J e^{i((k_1+k_2)\xi + (\omega_1+\omega_2)t)}, \tag{VI.17}$$

and similarly (VI.13). These equations have the same form as (3.22) and (3.23) of [15], which were interpreted as the holographic dual to the perturbative Einstein's equations with the right hand side serving as the matter source.

In general dimensions,we can derive a constraint equation similar to VI.9 for a ball A of radius l centered on $\vec{\xi}$:

$$(\partial_l^2 - (d-2)\frac{\partial_l}{l} - \nabla_\xi^2 - \frac{d}{l^3})\delta S_A(\vec{\xi}, l, t) = 0 \tag{VI.18}$$

As in the case of 2 dimensions , this can be verified straightforwardly by applying the differential operator above to the expression for δS_A in (VI.10) and integrating by parts after noting that:

$$\int_A \beta(r)\nabla^2_{\vec{\xi}} T_{00}(\vec{\xi}+\vec{r})drd\Omega = \int_A \beta(r)\nabla^2_{\vec{r}} T_{00}(\vec{\xi}+\vec{r})drd\Omega = -\int_A \nabla\beta(r)\cdot\nabla_{\vec{r}} T_{00}(\vec{\xi}+\vec{r})drd\Omega \tag{VI.19}$$

For $d = 3$, [16] recently derived the same equation holographically. In[30], a general argument was proposed explaining why (VI.10) leads to the perturbative Einstein's equations via the holographic entanglement entropy formula (VI.1). In addition, a quantity called entanglement density was introduced in [15]. In $d = 2$, for an interval $A = [u, v]$ of length $l = v - u$ and midpoint ξ, this is defined as

$$n(\xi, l, t) = \frac{1}{2}\frac{\delta^2 S_A}{\delta u \delta v}, \qquad \Delta n(\xi, l, t) = \frac{1}{2}\frac{\delta^2 \Delta S_A}{\delta u \delta v}, \tag{VI.20}$$

where in the second equality we present the shifted entanglement density in terms of the renormalized entanglement entropy ΔS_A. Writing ΔS_A in terms of u and v as in equation (III.9) and computing the derivatives gives

$$l^2 \Delta n(\xi, l, t) + \Delta S_A = 0, \tag{VI.21}$$

$$\lim_{l\to 0} \Delta n(\xi, l, t) = T_{00}(\xi) \lim_{l\to 0} 2\pi \int_{-l}^{l} dx' \frac{l^2 - x'^2}{2l^2} = \frac{\pi}{3} T_{00}(\xi). \tag{VI.22}$$

which agrees with the holographic results of [15]. Finally, we note some overlap with [17]. The author of [17] considered a gravitational theory on Rindler space and derived the change in entanglement entropy across the Rindler horizon as in equation (IV.2) due to a metric perturbation $g_{ab} \to g_{ab} + h_{ab}$. There the entanglement Hamiltonian (I.7), was evaluated along the event horizon H and was shown to be equal to an operator $\hat{A}_H$ that measures the area of the event horizon. The crucial ingredient deriving this relation was the universal coupling $\int h_{ab} T^{ab}$ of the graviton with the energy momentum tensor, which results in a perturbative Einstein's equation that relates T_{ab} to $\Box h_{ab}$. Thus, the renormalized horizon entanglement entropy was found to be

$$\delta S_H = \frac{Tr(A_H \delta \rho_H)}{4G} = \frac{\delta Area(H)}{4G}. \tag{VI.23}$$

Even though this equation was not derived from AdS/CFT, there is an obvious parallel here with equation (VI.3), where the minimal surface γ_A is identified with the horizon H.

VII. CONCLUSION

In this paper, we employed path integral methods to find a universal relation between the ground state entanglement Hamiltonian for an arbitrary region A and the physical stress tensor. For spherical entangling surfaces in a CFT we find, as in [11], that the entanglement Hamiltonian is the integral of a local density against a local entanglement temperature. We further generalize this result to include states with conserved charges preserving conformal invariance and derive new expressions for the entanglement Hamiltonians in various cylindrical backgrounds in 2 dimensions. Along the way, we show that the standard results for entanglement entropy in $d = 2$ dimensions that are traditionally derived from the replica trick can be obtained easily by evaluating the *thermal* entropy density using the entanglement temperature, and integrating over A. While completing this paper, we became aware that the same method was used in [31] to obtain the leading area law behavior of entanglement entropy for a half space A in a $d+1$-dimensional CFT and to derive the exact result for a finite interval A in a $d = 2$ CFT on the plane. It was also argued there that at high temperatures the entanglement entropy for theories with a mass gap m can be estimated by cutting off the size of the integration region A at $x^1 = \frac{1}{m}$, and indeed this gives the exact result for $d = 2$. In this paper, we made the additional observation that the entanglement temperature relates the change in entanglement entropy to changes in conserved charges of the ground state via equation (V.3). However, we should note that the spatially varying entanglement temperature is not physical in the sense that it does not determine the expectation value of local observables such as T_{00} (Indeed, $\langle T_{00}\rangle$ is a constant.). This is because the entanglement Hamiltonian (I.4) is an integral over operators that do not commute, so the reduced density matrix does not factorize. Indeed the entanglement temperature is not even conformally invariant; however equation V.3 shows that in a fixed conformal frame, it gives a universal relation between the expectation value of physical charges inside a region A and the renormalized entanglement entropy.

The relation (I.4) between the entanglement Hamiltonian and the stress tensor, when combined with our CFT expression (IV.2) for renormalized entanglement entropy provides a direct connection between the expectation value

of the stress tensor and the increase in entanglement entropy, as was first noted in the holographic calculations of [18] and further generalized in [15]. We also want to point out that it was recently observed in [32] that for spherical and cylindrical regions A, the holographic prescription [27] for the ground state entanglement entropy coincide with setting the *finite* part of 00 component of the holographic stress tensor to zero on a 4D slice of the bulk spacetime. The idea is that given a parametrization $r = f(z)$ of the 4D slice, setting

$$\langle T^h_{00} \rangle = \langle K_{ab} - h_{ab} K \rangle = 0, \tag{VII.1}$$

where h_{ab} is the induced metric and K_{ab} is the extrinsic curvature gives a differential equation for $f(z)$ that is identical to the minimal area equation. A heuristic field theory justification might go as follows. Demanding that a state in a CFT has the same entanglement entropy of the ground state corresponds to setting

$$\delta S_A = \int_A \beta_A \langle T^{CFT}_{00} \rangle = 0. \tag{VII.2}$$

If we follow [11] we identify the region casual development of A with a curved 4D slice of AdS, then identifying $\langle T^{CFT}_{00} \rangle$ with $\langle T^h_{00} \rangle$ would gives equation VII.1.

Acknowledgments

We thank Phil Szepietowski for participation in the initial stages of this project. G.W. would like to thank Peter Arnold, Thomas Mark, and Yifei Shi for helpful discussions and the Michigan Center for Theoretical Physics for hospitality. L.PZ. would like to acknowledge the hospitality of University of Virginia through the Physics Theory Visitor program and University of Texas, Austin. This work was supported by the DoE Grant #DE-SC0007859 to the University of Michigan. G.W. was supported by the Presidential Fellowship from the University of Virginia. IK would like to acknowledge financial support from NSF CAREER award No. DMR-0956053. DV and GW were supported in part by DOE grant #DE-SC0007984.

Appendix A: Evaluating the ground state entanglement entropy from the entanglement Hamiltonian

In this section we would like to point out a subtlety in evaluating the ground state entanglement entropy directly from equation (I.4). The discussion will also serve to provide some background for the calculation in appendix B. Given the normalized reduced density, $\rho_A = \exp(-H_A)/Z_A$, the entanglement entropy is

$$S_A = -Tr_A(\rho_A \ln \rho_A) = Tr_A(\rho_A H_A) + lnZ_A. \tag{A.1}$$

Equation (I.4) implies the entanglement energy vanishes:

$$Tr(\rho_A H_A) = \int_A \beta(x) Tr(\rho T_{00}) = \int_A \beta(x) \langle 0| : T_{00} : |0\rangle = 0. \tag{A.2}$$

In the last equality, we have normal ordered T_{00} with respect to the usual Minkowski annihilation operators, so S_A comes entirely from the "free energy " term[13] $ln : Z_A :$.

However there is an alternative way to evaluate the entanglement energy by conformally mapping ρ_A to a thermal density matrix with uniform temperature [11]. In the case of a a free scalar field in 2 spacetime dimensions and for $A = \{x > 0\}$, $H_A = 2\pi \int_{x \geq 0} x[(\partial_x \phi)^2 + (\partial_t \phi)^2] dx$ is the Rindler Hamiltonian [20]. In terms of Rindler coordinates

$$x = e^{\xi} \cosh(\eta), \qquad t = e^{\xi} \sinh(\eta), \tag{A.3}$$

it can be written as

$$H_A = 2\pi \int_{-\infty}^{\infty} (\partial_\xi \phi)^2 + (\partial_\eta \phi)^2 d\xi. \tag{A.4}$$

---

[13] Even though Z_A is not an operator, we use the normal ordering symbol to highlight the fact that its value depends on normal ordering.

Thus H_A can be quantized by expanding the field in terms of plane waves in *Rindler* coordinates [13],

$$\phi_k = \int \frac{dk}{\sqrt{4\pi k}} b_k e^{ik(\xi-\eta)} + c.c., \qquad H_A = \int dk[b_k^\dagger b_k + (1/2)\delta(0)]k. \tag{A.5}$$

The delta function term represents the Casimir energy and is removed by normal ordering with respect to the *Rindler* annihilation operator b_k. It is well known that under Rindler normal ordering, the Minkowski vacuum is thermal [13] so that

$$Tr[\rho_A H_A] = \langle 0|\dot{:}H_A\dot{:}|0\rangle = \frac{1}{12}\ln\frac{L}{\delta} \tag{A.6}$$

where L and δ are IR and UV cutoff's so that $A = [\delta, L]$. This result can be obtained by a standard computation of the average thermal energy for a free relativistic gas of massless (Rindler) particles at temperature $\frac{1}{2\pi}$, subject to Dirichlet boundary conditions in the *Rindler* spatial coordinate ξ. Note that this differs from equation (A.2) due to the difference in Rindler mode vs. Minkowski mode normal ordering, which we denote by 3 and 2 dots respectively. We can also obtain the corresponding Rindler free energy by usual statistical mechanics arguments:

$$\dot{:}\ln Z_A\dot{:} = \frac{1}{12}\ln\frac{L}{\delta}. \tag{A.7}$$

Adding this term to the entanglement energy (A.6) gives

$$S_A = \frac{1}{6}\ln\frac{L}{\delta}, \tag{A.8}$$

which is consistent with the known result [23]. Since adding a normal ordering constant a to H_A corresponds to a shift $lnZ_A \to lnZ_A - a$, (A.7), (A.6) and (A.2) implies : $\ln Z_A := \frac{1}{6}\ln\frac{L}{\delta}$, which is the same as S_A as it should be.

The lesson here is that while S_A is conformally invariant, neither the entanglement energy or free energy is.

To drive home this point we can derive the same result in a two dimensional Euclidean CFT, in the same spirit as [11] and [12]. The Euclidean version of the coordinate change from Minkowski to Rindler coordinates is the conformal map

$$w = logz, \qquad z = x + it, \qquad w = \xi + i\theta, \tag{A.9}$$

where z and w are the Euclideanised Minkowski/Rindler coordinates respectively and θ is the angular coordinate on the z plane. The z plane is mapped to a strip of length 2π and the Entanglement hamiltonian on the z plane is mapped to physical hamiltonian H_θ that evolves states along the θ direction[14]. For the ground state on the plane, $T(z) = 0$ so that the transformation of the stress tensor [15] gives $T(w) = c/24$ [33] . Integrating along ξ to gives the expectation value of the Hamiltonian on the z plane:

$$\langle H_\theta\rangle = 2\pi\int_{w(A)} d\xi \frac{(\langle T(w) + T(\bar{w})\rangle)}{2\pi} = \frac{c}{12}\ln\frac{L}{\delta} \tag{A.10}$$

which is the desired result[16]. In the last equality we have again set $A = [\delta, L]$ on the $t = 0$ slice of the z plane, so that it is mapped to $w(A) = [\ln\delta, \ln(L)]$.

Appendix B: Non-uniform excitation of 2D free scalar field

In this appendix we provide an explicit evaluation of the entanglement energy $\delta S_A = Tr(\delta\rho_A H_A)$ in equation (IV.3) for a spatially non-uniform excitation of a 2D free scalar field and show that it is indeed equal to eq (IV.2). First note that normal ordering is irrelevant in this case because shifting H_A by a constant does not change the entanglement

---

[14] This is the d=2 dimensional analogue of the conformal transformation to the hyperbolic space $\mathcal{H}$ III.17 for the half infinite line A.
[15] To conform with the conventions of [33] T_{ab} is defined so the Hamiltonian is $H = \frac{1}{2\pi}T_{00}$.
[16] One of the 2π's are from the length of the strip and the other from the definition of H in terms of $T(w)$.

energy due to the normalization condition $Tr(\delta\rho_A) = 0$. Now following [34] we consider a particular excitation labelled by a positive Rindler momentum k:

$$d_k^\dagger|0\rangle = \int_0^\infty dp D(k,p) a_p^\dagger|0\rangle, \qquad D(k,p) = (2kSinh(\pi k))^{1/2}\Gamma(-ik)|p|^{ik-\frac{1}{2}}, \qquad k > 0 \tag{B.1}$$

where $a_p^\dagger$ are the conventional Minkowski creation operators. It is then straight forward to compute the (unnormalized) energy density by quantizing the energy density $T_{00} = \frac{1}{2}\{(\partial_\xi\phi)^2 + (\partial_\eta\phi)^2\}$ in terms of Minkowski modes:

$$\langle 0|d_k : T_{00} : d_k^\dagger|0\rangle = \frac{(-1+e^{2k\pi})k\pi^2\mathrm{csch}^2(k\pi)}{2\pi x^2}. \tag{B.2}$$

Now we compute the δS_A for the half space A using the entanglement Hamiltonian in equation (III.2). Dividing by the (infinite) normalization constant $N = \langle 0|d_k d_k^\dagger|0\rangle = 2\pi\int_0^\infty \frac{dp}{p}$ and inserting into the equation (IV.3) gives[17]

$$\delta S_A = \pi k(1 + \coth(k\pi)). \tag{B.3}$$

Alternatively, we can evaluate δS_A using equation (IV.2) via an explicit representation of the reduced density matrix $\delta\rho_A$ corresponding to the state in equation (B.1). If we define the following reduced density matrices for the kth mode,

$$\rho_0(k) = \sum_{n_k=0}^\infty e^{-2\pi nk}(1-e^{2\pi k})|n_k\rangle\langle n_k|, \tag{B.4}$$

$$\rho_1(k) = \sum_{n_k=0}^\infty (4n_k \sinh^2(\pi k))|n_k\rangle\langle n_k|, \tag{B.5}$$

where $|n_k\rangle$ denotes the occupational number basis for the Rindler particles, then results of [34] imply that $\delta\rho_A = \rho_1(k)\prod_{l\neq k}\rho_0(l) - \prod_l \rho_0(l)$. Inserting this into (IV.3) and evaluating the trace using the Rindler Hamiltonian (A.5) gives

$$\begin{aligned}\delta S_A &= 2\pi\sum_{n\geq 1} e^{-2\pi nk}(4n\sinh^2(\pi k) - (1-e^{-2\pi k})\langle n_k|\vdots H_A\vdots|n_k\rangle \\ &= 2\pi\sum_{n\geq 1} e^{-2\pi nk}(4n\sinh^2(\pi k) - (1-e^{-2\pi k}))nk \\ &= \pi k(1+\coth(\pi k)),\end{aligned} \tag{B.6}$$

which is the same result as equation (B.3).

[1] L. Bombelli, R. Koul, J. Lee, and R. Sorkin, *Quantum source of entropy for black holes*, *Physical Review D* **34** (1986), no. 2 373–383.

[2] H.-C. Jiang, Z. Wang, and L. Balents, *Identifying topological order by entanglement entropy*, *Nature Physics* **8** (2012), no. 12 902–905.

[3] H. Li and F. Haldane, *Entanglement Spectrum as a Generalization of Entanglement Entropy: Identification of Topological Order in Non-Abelian Fractional Quantum Hall Effect States*, *Physical Review Letters* **101** (2008), no. 1 10504.

[4] A. M. Läuchli, E. J. Bergholtz, J. Suorsa, and M. Haque, *Disentangling entanglement spectra of fractional quantum hall states on torus geometries*, *Physical review letters* **104** (2010), no. 15 156404.

[5] A. Chandran, M. Hermanns, N. Regnault, and B. A. Bernevig, *Bulk-edge correspondence in entanglement spectra*, *Physical Review B* **84** (2011), no. 20 205136.

[6] X. Qi, H. Katsura, and A. Ludwig, *General relationship between the entanglement spectrum and the edge state spectrum of topological quantum states*, *Arxiv preprint arXiv:1103.5437* (2011).

[17] The logarithmically divergent integral over x is cancelled by the normalization N.

[7] A. M. Turner, Y. Zhang, and A. Vishwanath, *Entanglement and inversion symmetry in topological insulators*, *Physical Review B* **82** (2010), no. 24 241102.

[8] L. Fidkowski, T. Jackson, and I. Klich, *Model characterization of gapless edge modes of topological insulators using intermediate brillouin-zone functions*, *Physical review letters* **107** (2011), no. 3 036601.

[9] J. J. Bisognano and E. H. Wichmann, *On the duality condition for a hermitian scalar field*, *Journal of Mathematical Physics* **16** (1975) 985.

[10] J. J. Bisognano and E. H. Wichmann, *On the duality condition for quantum fields*, *Journal of mathematical physics* **17** (1976) 303.

[11] H. Casini, M. Huerta, and R. C. Myers, *Towards a derivation of holographic entanglement entropy*, *Journal of High Energy Physics* **2011** (2011), no. 5 1–41.

[12] C. Holzhey, F. Larsen, and F. Wilczek, *Geometric and renormalized entropy in conformal field theory*, *Nucl.Phys.* **B424** (1994) 443–467.

[13] W. G. Unruh and R. M. Wald, *What happens when an accelerating observer detects a rindler particle*, *Physical Review D* **29** (1984), no. 6 1047–1056.

[14] M. Nozaki, T. Numasawa, and T. Takayanagi, *Holographic Local Quenches and Entanglement Density*, *arXiv: 1302.5703* (2013).

[15] M. Nozaki, T. Numasawa, A. Prudenziati, and T. Takayanagi, *Dynamics of Entanglement Entropy from Einstein Equation*, *arXiv:1304.7100* (2013).

[16] J. Bhattacharya and T. Takayanagi, *Entropic counterpart of perturbative einstein equation*, *arXiv preprint arXiv:1308.3792* (2013).

[17] E. Bianchi and R. C. Myers, *On the Architecture of Spacetime Geometry*, *arXiv: 1212.5183* (2012).

[18] J. Bhattacharya, M. Nozaki, T. Takayanagi, and T. Ugajin, *Thermodynamical Property of Entanglement Entropy for Excited States*, *Phys. Rev. Lett. 110,* **091602** (2013).

[19] W. G. Unruh, *Notes on black-hole evaporation*, *Phys. Rev. D* **14** (Aug, 1976) 870–892.

[20] L. Susskind and J. Lindesay, *Black holes, information and the string theory revolution*, *World Scientific* (2005).

[21] P. Calabrese and J. Cardy, *Entanglement entropy and quantum field theory*, *Journal of Statistical Mechanics: Theory and Experiment* **2004** (2004), no. 06 P06002.

[22] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *Journal of Physics A: Mathematical and Theoretical* **42** (2009), no. 50 504005.

[23] F. Larsen and F. Wilczek, *Geometric entropy, wave functionals, and fermions*, *Annals Phys.* **243** (1995) 280–298, [{\tthep-th/9408089}].

[24] L.-Y. Hung, R. C. Myers, M. Smolkin, and A. Yale, *Holographic calculations of renyi entropy*, *Journal of High Energy Physics* **2011** (2011), no. 12 1–61.

[25] M. I. Berganza, F. C. Alcaraz, and G. Sierra, *Entanglement of excited states in critical spin chians*, *J.Stat.Mech.* **1201** (2012) P01016.

[26] V. E. Hubeny, M. Rangamani, and T. Takayanagi, *A covariant holographic entanglement entropy proposal*, *Journal of High Energy Physics* **2007** (2007), no. 07 062.

[27] S. Ryu and T. Takayanagi, *Aspects of holographic entanglement entropy*, *Journal of High Energy Physics* **2006** (2006), no. 08 045.

[28] D. Allahbakhshi, M. Alishahiha, and A. Naseh, *Entanglement thermodynamics*, *arXiv preprint arXiv:1305.2728* (2013).

[29] S. de Haro, S. N. Solodukhin, and K. Skenderis, *Holographic reconstruction of space-time and renormalization in the ads / cft correspondence*, *Commun.Math.Phys.* **217** (2001) 595–622.

[30] N. Lashkari, M. B. McDermott, and M. V. Raamsdonk, *Gravitational dynamics from entanglement "thermodynamics"*, *arXiv 1308.3716* (08, 2013).

[31] B. Swingle, *Structure of entanglement in regulated Lorentz invariant field theories*, *arXiv:1304.6402* (2013).

[32] A. Bhattacharyya and A. Sinha, *Entanglement entropy from the holographic stress tensor*, *arXiv preprint arXiv:1303.1884* (2013).

[33] J. L. Cardy, *Conformal invariance and statistical mechanics*, *Les Houches, Session XLIX* (1988) 169–245.

[34] E. Benedict and S.-Y. Pi, *Entanglement entropy of nontrivial states*, *arXiv preprint hep-th/9505121* (1995).

Detecting topological order in a ground state wave function

Michael Levin and Xiao-Gang Wen
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A large class of topological orders can be understood and classified using the string-net condensation picture. These topological orders can be characterized by a set of data $(N, d_i, F^{ijk}_{lmn}, \delta_{ijk})$. We describe a way to detect this kind of topological order using only the ground state wave function. The method involves computing a quantity called the "topological entropy" which directly measures the quantum dimension $D = \sum_i d_i^2$.

PACS numbers: 11.15.-q, 71.10.-w

Introduction: Until recently, the only known physical characterizations of topological order [1] involved properties of the Hamiltonian - e.g. quasiparticle statistics [2], ground state degeneracy [3, 4], and edge excitations [1]. In this paper, we demonstrate that topological order is manifest not only in these dynamical properties but also in the basic entanglement of the ground state wave function. We hope that this characterization of topological order can be used as a theoretical tool to classify trial wave functions - such as resonating dimer wave functions [5], Gutzwiller projected states, [6–10] or quantum loop gas wave functions [11]. In addition, it may be useful as a numerical test for topological order. Finally, it demonstrates definitively that topological order is a property of a wave function, not a Hamiltonian. The classification of topologically ordered states is nothing but a classification of complex functions of thermodynamically large numbers of variables.

Main Result: We focus on the $(2+1)$ dimensional case (though the result can be generalized to any dimension). Let Ψ be an arbitrary wave function for some two dimensional lattice model. For any subset A of the lattice, one can compute the associated quantum entanglement entropy S_A.[12] The main result of this paper is that one can determine the "quantum dimension" D of Ψ by computing the entanglement entropy S_A of particular regions

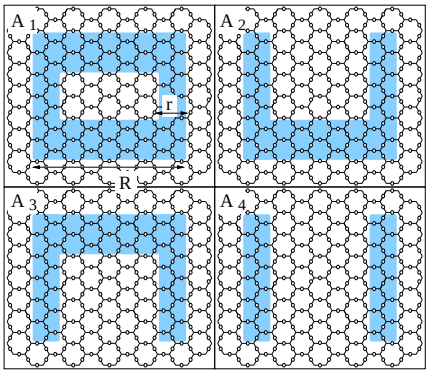


FIG. 1: One can detect topological order in a state Ψ by computing the von Neumann entropies S_1, S_2, S_3, S_4 of the above four regions, A_1, A_2, A_3, A_4, in the limit of $R, r \to \infty$. Here the four regions are drawn in the case of the honeycomb lattice. The geometry ensures that the number of sites n_1, n_2, n_3, n_4 along the boundaries of the 4 regions obey the relation $n_1 - n_2 = n_3 - n_4$.

A in the plane. Normal states have $D = 1$ while topologically ordered states have $D \neq 1$ (e.g. in the case of topological orders described by discrete gauge theories, D is equal to the number of elements in the gauge group). Thus, D provides a way to distinguish topologically ordered states from normal states.

More specifically, consider the four regions A_1, A_2, A_3, A_4 drawn in Fig. 1. Let the corresponding entanglement entropies be S_1, S_2, S_3, S_4. Consider the linear combination $(S_1 - S_2) - (S_3 - S_4)$, computed in the limit of large, thick annuli, $R, r \to \infty$. The main result of this paper is that

$$(S_1 - S_2) - (S_3 - S_4) = -\log(D^2) \qquad (1)$$

where D is the quantum dimension of the topological field theory associated with Ψ.

We call the quantity $(S_1 - S_2) - (S_3 - S_4)$ the "topological entropy", $-S_{\text{top}}$, since it measures the entropy associated with the (non-local) topological entanglement in Ψ. The above result implies that S_{top} is an universal number associated with each topological phase. It is invariant under smooth deformations of Ψ.

Physical picture: The idea behind (1) is that topologically ordered states contain nonlocal entanglement. Consider, for example, a spin-1/2 model with spins located on the links $\boldsymbol{i}$ of the honeycomb lattice and with a Hamiltonian realizing a Z_2 lattice gauge theory. [13–15] The ground state wave function Ψ is known exactly. The easiest way to describe Ψ is in terms of strings. One can think of each spin state as a string state, where a $\sigma^x_{\boldsymbol{i}} = -1$ spin corresponds to a link occupied by a string and a $\sigma^x_{\boldsymbol{i}} = 1$ spin corresponds to an empty link. In this language, Ψ is very simple: $\Psi(X) = 1$ for all string states X where the strings form closed loops, and Ψ vanishes otherwise.

All local correlations $\langle \sigma^x_{\boldsymbol{i}} \sigma^x_{\boldsymbol{j}} \rangle$ vanish for this state. However, Ψ contains *nonlocal* correlations or entanglement. To see this, imagine drawing a curve C in the plane (see Fig. 2). There is a nonlocal correlation between the spins on the links crossing this curve: $\langle W(C) \rangle = \langle \prod_{\boldsymbol{i} \in C} \sigma^x_{\boldsymbol{i}} \rangle = 1$. This correlation originates from the fact that the number of strings crossing the curve is always even. Similar correlations exist for more general states that contain virtual string-breaking fluctuations. In the general case, the nonlocal correlations can be captured by "fattened" string operators $W_{\text{fat}}(C)$ that act on spins within some distance l of C where l is the length scale for string breaking.

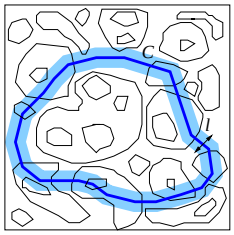


FIG. 2: The state Ψ contains nonlocal correlations originating from the fact that closed strings always cross a closed curve C an even number of times. These correlations can be measured by a string operator $W(C)$ (thin blue curve). For more general states, a fattened string operator $W_{\rm fat}(C)$ (light blue region) is necessary.

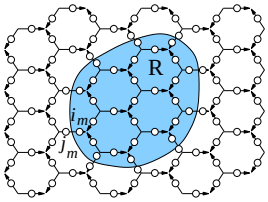


FIG. 3: A simply connected region R in the honeycomb lattice. We split the sites on the boundary links into two sites labeled $\boldsymbol{i_m}$ and $\boldsymbol{j_m}$, where $m = 1, ..., n$.

To determine whether a state is topologically ordered, one has to determine whether the state contains such nonlocal correlations or entanglement. While it is difficult to find the explicit form of the fattened string operators $W_{\rm fat}$,[16] one can establish their existence or nonexistence using quantum information theory. The idea is that if the string operators exist, then the entropy of an annular region (such as A_1 in Fig. 1) will be lower than one would expect based on local correlations.

The combination $(S_1-S_2)-(S_3-S_4)$ measures exactly this anomalous entropy. To see this, notice that $(S_1 - S_2)$ is the amount of additional entropy associated with closing the region A_2 at the top. Similarly, $(S_3 - S_4)$ is the amount of additional entropy associated with closing the region A_4 at the top. If Ψ has only local correlations with correlation length ξ then these two quantities are the same up to corrections of order $O(e^{-R/\xi})$, since A_2, A_4 only differ by the region at the bottom. For such states, $\lim_{R\to\infty}(S_1-S_2)-(S_3-S_4) = 0$. Thus, a *nonzero* value for $S_{\rm top}$ signals the presence of nonlocal correlations and topological order.

The universality of $S_{\rm top}$ can also be understood from this picture. Small deformations of Ψ will typically modify the form of the string operators $W_{\rm fat}$ and change their width l. However, as long as l remains finite, $(S_1 - S_2) - (S_3 - S_4)$ will converge to the same universal number when the width r of the annular region is larger than l.

A simple example: Let us compute the topological entropy of the ground state wave function Ψ of the Z_2 model and confirm (1) in this case. We will first compute the entanglement entropy S_R for an arbitrary region R. To make the boundary more symmetric, we split the sites on the boundary links into two sites (see Fig. 3). The wave function Ψ generalizes to the new lattice in the natural way. The new wave function (still denoted by Ψ) has the same entanglement entropy.

We will decompose Ψ into $\Psi = \sum_l \Psi^{\rm in}_l \Psi^{\rm out}_l$ where $\Psi^{\rm in}_l$ are wave functions of spins inside R, $\Psi^{\rm out}_l$ are wave functions of spins outside R, and l is a dummy index. A simple decomposition can be obtained using the string picture. For any $q_1, ..., q_n$, with $q_m = 0, 1$, and $\sum_m q_m$ even, we can define a wave function $\Psi^{\rm in}_{q_1,...,q_n}$ on the spins inside of R: $\Psi^{\rm in}_{q_1,...,q_n}(X) = 1$ if (a) the strings in X form closed loops and (b) X satisfies the boundary condition that there is a string on $\boldsymbol{i_m}$ if $q_m = 1$, and no string if $q_m = 0$. Similarly, we can define a set of wave functions $\Psi^{\rm out}_{r_1...,r_n}$ on the spins outside of R.

If we glue Ψ^{in} and Ψ^{out} together - setting $q_m = r_m$ for all m - the result is Ψ. Formally, this means that

$$\Psi = \sum_{q_1+...+q_l \text{ even}} \Psi^{\rm in}_{q_1,...,q_n} \Psi^{\rm out}_{q_1...,q_n} \qquad (2)$$

It is not hard to see that the functions $\{\Psi^{\rm in}_{q_1,...,q_n} : \sum_m q_m \text{ even}\}$, and $\{\Psi^{\rm out}_{r_1...,r_n} : \sum_m r_m \text{ even}\}$ are orthonormal up to an irrelevant normalization factor. Therefore, the density matrix for the region R is an equal weight mixture of all the $\{\Psi^{\rm in}_{q_1...,q_n} : \sum_m q_m \text{ even}\}$. There are 2^{n-1} such states. The entropy is therefore $S_R = (n-1)\log 2$. [12]

This formula applies to simply connected regions like the one in Fig. 3. The same argument can be applied to general regions R and leads to $S_R = (n-j)\log 2$, where n is the number of spins along ∂R, and j is the number of disconnected boundary curves in ∂R,

We are now ready to calculate the topological entropy associated with Ψ. According to (1) we need to calculate the entropy associated with the four regions shown in Fig. 1. From $S_R = (n-j)\log 2$, we find $S_1 = (n_1 - 2)\log 2$, $S_2 = (n_2 - 1)\log 2$, $S_3 = (n_3 - 1)\log 2$, and $S_4 = (n_4 - 2)\log 2$, where n_1, n_2, n_3, n_4 are the number of spins along the boundaries of the four regions. The topological entropy is therefore $-S_{\rm top} = (n_1 - n_2 - n_3 + n_4 - 2)\log 2$. But the four regions are chosen such that $(n_1 - n_2) = (n_3 - n_4)$. Thus the size dependent factor cancels out and we are left with $-S_{\rm top} = -2\log 2 = -\log(2^2)$. This is in agreement with (1) since the quantum dimension of Z_2 gauge theory is $D = 2$.

General string-net models: To derive (1) in the general case, we compute the the topological entropy for the exactly soluble string-net models discussed in Ref. [17]. The ground states of these models describe a large class of $(2 + 1)$ dimensional topological orders. The models and the associated topological orders are characterized by several pieces of data: (a) An integer N - the number of string types. (b) A completely symmetric tensor δ_{ijk} where $i, j, k = 0, 1, ..., N$ and δ_{ijk} only takes on the values 0 or 1. This tensor represents the branching rules: three string types i, j, k are allowed to meet at a point if and only if $\delta_{ijk} = 1$. (c) A dual string type i^* corresponding to each string type i. This dual string type corresponds to the same string, but with the opposite orientation. (d) A real tensor d_i and a complex tensor

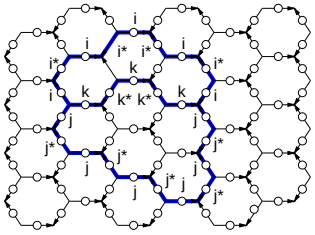
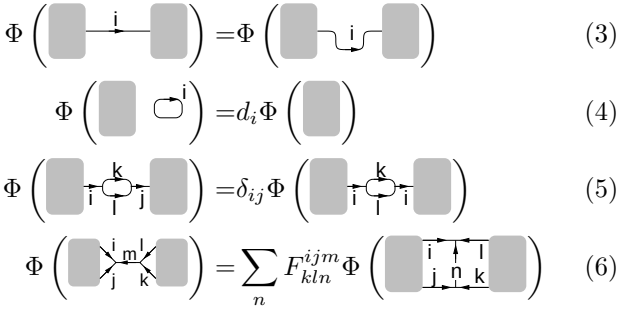


FIG. 4: A typical string-net state on the honeycomb lattice. The empty links correspond to spins in the $i = 0$ state. The orientation conventions on the links are denoted by arrows.

F^{ijm}_{kln} satisfying certain algebraic relations [17]. For each set of $F^{ijm}_{kln}, d_i, \delta_{ijk}$ satisfying these relations, there is a corresponding exactly soluble topologically ordered spin model.

The spins in the model are located on the links $\boldsymbol{k}$ of the honeycomb lattice. However, the spins are not usual spin-1/2 spins. Each spin can be in $N+1$ different states which we will label by $i = 0, 1, ..., N$. Each spin state can be thought of as a string-net state. To do this, one first needs to pick an orientation for each link on the honeycomb lattice. When a spin is in state i, we think of the link as being occupied by a type-i string oriented in the appropriate direction. If a spin is in state $i = 0$, then we think of the link as empty. In this way spin states correspond to string-net states (see Fig. 4). The Hamiltonian of the model involves a 12 spin interaction [17]. The model is known to be gapped and topologically ordered and all the relevant quantities - ground state degeneracies, quasiparticle statistics, etc., can be calculated explicitly.

The ground state wave function Φ of the model is also known exactly. It is easiest to describe in terms of the string-net language. If a spin configuration $\{i_{\boldsymbol{k}}\}$ corresponds to an invalid string-net configuration - that is, a string-net configurations that doesn't obey the branching rules defined by δ_{ijk} - then $\Phi(\{i_{\boldsymbol{k}}\}) = 0$. On the other hand, if $\{i_{\boldsymbol{k}}\}$ corresponds to a valid string-net configuration then the amplitude is in general nonzero. We would like to have an explicit formula for these amplitudes. Unfortunately, this is not possible in general. However, we can write down linear relations that determine the amplitudes uniquely. These relations relate the amplitudes of string-net configurations that only differ by small local transformations. The relations are given by



where the shaded areas represent other parts of the string-nets that are not changed. Also, the type-0 string is interpreted as the no-string (or vacuum) state. The first relation (3) is drawn schematically. The more precise statement of this rule is that any two string-net configurations on the honeycomb lattice that can be continuously deformed into each other have the same amplitude. In other words, the string-net wave function Φ only depends on the topologies of the network of strings.

By applying these relations multiple times, one can compute the amplitude for any string-net configuration (on the honeycomb lattice) in terms of the amplitude of the vacuum configuration. Thus, (3-6) completely specify the ground state wave function Φ.

Let us first compute the von Neumann entropy S_R of the exact ground state wave function Φ for a simply connected region R (see Fig. 3). Again we split the site on the boundary links into two sites. We decompose Φ into $\Phi = \sum_l \Phi^{\text{in}}_l \Phi^{\text{out}}_l$ where Φ^{in}_l are wave functions of spins inside R, Φ^{out}_l are wave functions of spins outside R, and l is some dummy index.

A wave function Φ^{in} on the spins inside of R can be defined as follows. Let $\{i_{\boldsymbol{k}}\}$ be some spin configuration inside of R. If $\{i_{\boldsymbol{k}}\}$ doesn't correspond to a valid string-net configuration - that is one that obeys the branching rules, then we define $\Phi^{\text{in}}(\{i_{\boldsymbol{k}}\}) = 0$. If $\{i_{\boldsymbol{k}}\}$ does correspond to a valid string-net configuration, then we define $\Phi^{\text{in}}(\{i_{\boldsymbol{k}}\})$ using the same graphical rules (3-6) that we used for Φ.

However, there is an additional subtlety. Recall that in the case of Φ, the graphical rules could be used to reduce any string-net configuration to the vacuum configuration. To fix Φ, we defined $\Phi(\text{vacuum}) = 1$.

In this case, since we are dealing with a region R *with a boundary*, string-net configurations cannot generally be reduced to the vacuum configuration. However, they can be reduced to the tree-like diagrams $X_{\{q,s\}}$ shown in Fig. 5a. Thus, to define Φ^{in}, we need to specify the amplitude for all of these basic configurations. There are multiple ways of doing this and hence multiple possibilities for Φ^{in}. Here, we will consider all the possibilities. For any labeling $q_1, ..., q_n, s_1, ..., s_{n-3}$ of the string-net in Fig. 5(a), we define a wave function $\Phi^{\text{in}}_{\{q,s\}}$ by $\Phi^{\text{in}}_{\{q,s\}}(X_{\{q',s'\}}) = \delta_{\{q\},\{q'\}}\delta_{\{s\},\{s'\}}$. Starting from these amplitudes and using the graphical rules (3-6) we can determine $\Phi^{\text{in}}_{\{q,s\}}(X)$ for all other string-net configurations. In the same way, we can define wave functions $\Phi^{\text{out}}_{\{r,t\}}$ on the spins outside of R through $\Phi^{\text{out}}_{\{r,t\}}(Y_{\{r',t'\}}) = \delta_{\{r\},\{r'\}}\delta_{\{t\},\{t'\}}$, where the $Y_{\{r,t\}}$ are shown in Fig. 5(b).

Now consider the product wave functions $\Phi^{\text{in}}_{\{q,s\}}\Phi^{\text{out}}_{\{r,t\}}$. These are wave functions on the all the spins in the system - both inside and outside R. They can be generated from the amplitudes for the string-net configurations $Z_{\{q,s,r,t\}}$ in Fig. 6:

$$\Phi^{\text{in}}_{\{q,s\}}\Phi^{\text{out}}_{\{r,t\}}(Z_{\{q',s',r',t'\}}) = \delta_{\{q\},\{q'\}}\delta_{\{s\},\{s'\}}\delta_{\{r\},\{r'\}}\delta_{\{t\},\{t'\}}$$

On the other hand, it is not hard to show that for the ground state wave function Φ, the amplitude for $Z_{\{q,s,r,t\}}$ is

$$\Phi(Z_{\{q,s,r,t\}}) = \delta_{\{q\},\{r\}}\delta_{\{s\},\{t\}} \prod_m (\sqrt{d_{q_m}})$$

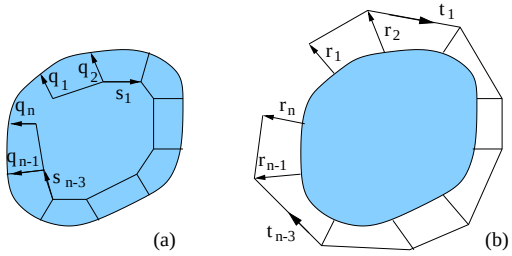


FIG. 5: The basic string-net configurations (a) $X_{\{q,s\}}$ for inside R and (b) $Y_{\{r,t\}}$ for outside R.

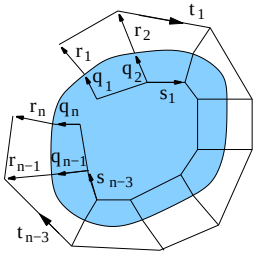


FIG. 6: The string-net configuration $Z_{\{q,s,r,t\}}$ obtained by "gluing" the configuration $X_{\{q,s\}}$ to the configuration $Y_{\{r,t\}}$ in Fig. 5.

Comparing the two, we see that

$$\Phi = \sum_{\{q,s,r,t\},} \Phi^{\rm in}_{\{q,s\}} \Phi^{\rm out}_{\{r,t\}} \delta_{\{q\},\{r\}} \delta_{\{s\},\{t\}} \prod_m (\sqrt{d_{q_m}}) \quad (7)$$

It turns out that the wave functions $\{\Phi^{\rm in}_{\{q,s\}}\}$ are orthonormal, as are the $\{\Phi^{\rm out}_{\{r,t\}}\}$ (up to an irrelevant normalization constant). This means that we can use them as a basis. If we denote $\Phi^{\rm in}_{\{q,s\}} \Phi^{\rm out}_{\{r,t\}}$ by $|\{q,s,r,t\}\rangle$, then in this basis, the wave function Φ is

$$\langle \{q,s,r,t\}|\Phi\rangle = \delta_{\{q\},\{r\}} \delta_{\{s\},\{t\}} \prod_m (\sqrt{d_{q_m}}) \quad (8)$$

The density matrix for the region R can now be obtained by tracing out the spins outside of R, or equivalently, tracing out the spin states $|\{r,t\}\rangle$:

$$\langle \{q',s'\}|\rho_R|\{q,s\}\rangle = \delta_{\{q\},\{q'\}} \delta_{\{s\},\{s'\}} \prod_m d_{q_m} \quad (9)$$

Since the density matrix is diagonal, we can easily obtain the entanglement entropy for S_R. Normalizing ρ_R so that ${\rm Tr}(\rho_R) = 1$, and taking $-{\rm Tr}\rho_R \log \rho_R$, we find

$$S_R = -\sum_{\{q,s\}} \frac{\prod_m d_{q_m}}{D^{n-1}} \log \left(\frac{\prod_l d_{q_l}}{D^{n-1}} \right) \quad (10)$$

where $D = \sum_k d_k^2$. The sum can be evaluated explicitly (with the help of the relations in [17]). The result is

$$S_R = -\log(D) - n \sum_{k=0}^{N} \frac{d_k^2}{D} \log \left(\frac{d_k}{D} \right) \quad (11)$$

This result applies to simply connected regions like the one shown in Fig. 1. The same argument can be applied to general regions R. In the general case, we find

$$S_R = -j \log(D) - n \sum_{k=0}^{N} \frac{d_k^2}{D} \log \left(\frac{d_k}{D} \right) \quad (12)$$

where n is the number of spins along ∂R, and j is the number of disconnected boundary curves in ∂R.

We can now calculate the topological entropy associated with Φ. Applying (12), we find $S_1 = -2\log D - n_1 s_0$, $S_2 = -\log D - n_2 s_0$, $S_3 = -\log D - n_3 s_0$, and $S_4 = -2\log D - n_4 s_0$ where n_1, n_2, n_3, n_4 are the numbers of spins along the boundaries of the four regions, and $s_0 = \sum_{k=0}^{N} \frac{d_k^2}{D} \log \left(\frac{d_k}{D} \right)$. The topological entropy is therefore $-S_{\rm top} = -2\log D + (n_1 - n_2 - n_3 + n_4) s_0 = -2\log D$ in agreement with (1).

Near the completion of this paper, we become aware of a similar result, obtained independently in the recent paper, Ref. [18]. This research is supported by NSF Grant No. DMR–04–33632 and by ARO Grant No. W911NF-05-1-0474.

[1] X.-G. Wen, Advances in Physics **44**, 405 (1995).
[2] D. Arovas, J. R. Schrieffer, and F. Wilczek, Phys. Rev. Lett. **53**, 722 (1984).
[3] F. D. M. Haldane and E. H. Rezayi, Phys. Rev. B **31**, 2529 (1985).
[4] X.-G. Wen and Q. Niu, Phys. Rev. B **41**, 9377 (1990).
[5] D. S. Rokhsar and S. A. Kivelson, Phys. Rev. Lett. **61**, 2376 (1988).
[6] G. Baskaran and P. W. Anderson, Phys. Rev. B **37**, 580 (1988).
[7] I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774 (1988).
[8] X.-G. Wen, F. Wilczek, and A. Zee, Phys. Rev. B **39**, 11413 (1989).
[9] N. Read and S. Sachdev, Phys. Rev. Lett. **66**, 1773 (1991).
[10] X.-G. Wen, Phys. Rev. B **44**, 2664 (1991).
[11] M. Freedman, C. Nayak, K. Shtengel, K. Walker, and Z. Wang, Ann. Phys. (NY) **310**, 428 (2004).
[12] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Rev. **A71**, 022315 (2005).
[13] F. Wegner, J. Math. Phys. **12**, 2259 (1971).
[14] J. B. Kogut, Rev. Mod. Phys. **51**, 659 (1979).
[15] A. Y. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003).
[16] M. B. Hastings and X.-G. Wen, Phys. Rev B **72**, 045141 (2005).
[17] M. Levin and X.-G. Wen, Phys. Rev. B **71**, 045110 (2005).
[18] A. Kitaev and J. Preskill, hep-th/0510092 (2005).

Topological entanglement entropy

Alexei Kitaev[1,2] and John Preskill[1]
[1] *Institute for Quantum Information, California Institute of Technology, Pasadena, CA 91125, USA*
[2] *Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA*

We formulate a universal characterization of the many-particle quantum entanglement in the ground state of a topologically ordered two-dimensional medium with a mass gap. We consider a disk in the plane, with a smooth boundary of length L, large compared to the correlation length. In the ground state, by tracing out all degrees of freedom in the exterior of the disk, we obtain a marginal density operator ρ for the degrees of freedom in the interior. The von Neumann entropy $S(\rho)$ of this density operator, a measure of the entanglement of the interior and exterior variables, has the form $S(\rho) = \alpha L - \gamma + \cdots$, where the ellipsis represents terms that vanish in the limit $L \to \infty$. The coefficient α, arising from short wavelength modes localized near the boundary, is nonuniversal and ultraviolet divergent, but $-\gamma$ is a universal additive constant characterizing a global feature of the entanglement in the ground state. Using topological quantum field theory methods, we derive a formula for γ in terms of properties of the superselection sectors of the medium.

PACS numbers: 03.65.Ud, 71.10.Pm, 73.43.Nq

In a quantum many-body system at zero temperature, a *quantum phase transition* may occur as a parameter varies in the Hamiltonian of the system. The two phases on either side of a quantum critical point may be characterized by different types of *quantum order*; the quantum correlations among the microscopic degrees of freedom have qualitatively different properties in the two phases. Yet in some cases, the phases cannot be distinguished by any local order parameter.

For example, in two spatial dimensions a system with a mass gap can exhibit *topological order* [1]. The quantum entanglement in the ground state of a topologically ordered medium has global properties with remarkable consequences. For one thing, the quasiparticle excitations of the system (*anyons*) exhibit an exotic variant of indistinguishable particle statistics. Furthermore, in the infinite-volume limit the ground-state degeneracy depends on the genus (number of handles) of the closed surface on which the system resides.

While it is clear that these unusual properties emerge because the ground state is profoundly entangled, up until now no firm connection has been established between topological order and any quantitative measure of entanglement. In this paper we provide such a connection by relating topological order to von Neumann entropy, which quantifies the entanglement of a bipartite pure state.

Specifically, we consider a disk in the plane, with a smooth boundary of length L, large compared to the correlation length. In the ground state, by tracing out all degrees of freedom in the exterior of the disk, we obtain a marginal density operator ρ for the degrees of freedom in the interior. The von Neumann entropy $S(\rho) \equiv -\mathrm{tr}\rho\log\rho$ of this density operator, a measure of the entanglement of the interior and exterior variables, has the form

$$S(\rho) = \alpha L - \gamma + \cdots \ , \tag{1}$$

where the ellipsis represents terms that vanish in the limit $L \to \infty$. The coefficient α, arising from short wavelength modes localized near the boundary, is nonuniversal and ultraviolet divergent [2], but $-\gamma$ (where γ is nonnegative) is a universal additive constant characterizing a global feature of the entanglement in the ground state. We call $-\gamma$ the *topological entanglement entropy*.

This universal quantity reflects topological properties of the entanglement that survive at arbitrarily long distances, and therefore can be studied using an effective field theory that captures the far-infrared behavior of the medium, namely a topological quantum field theory (TQFT) that describes the long-range Aharonov-Bohm interactions of the medium's massive quasiparticle excitations. We find

$$\gamma = \log \mathcal{D} \ , \tag{2}$$

where $\mathcal{D} \geq 1$ is the *total quantum dimension* of the medium, given by

$$\mathcal{D} = \sqrt{\sum_a d_a^2} \ ; \tag{3}$$

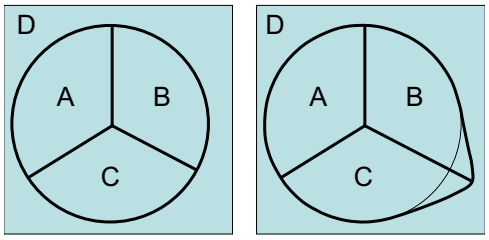
here the sum is over all the superselection sectors of the medium, and d_a is the *quantum dimension* of a particle with charge a.

Any abelian anyon has quantum dimension $d = 1$; therefore, for a model of abelian anyons, $\mathcal{D}^2$ is simply the number of superselection sectors. Thus for a Laughlin state [3] realized in a fractional quantum Hall system with filling factor $\nu = 1/q$ where q is an odd integer, we have $\mathcal{D} = \sqrt{q}$. For the toric code [4], which has four sectors, the topological entropy is $\gamma = \log 2$, as has already been noted in [5].

However, nonabelian anyons have quantum dimension greater than one. The significance of d_a (which need not be a rational number) is that the dimension $N_{aaa\cdots a}$ of

the fusion vector space spanned by all the distinguishable ways in which n anyons of type a can be glued together to yield a trivial total charge grows asymptotically like the nth power of d_a. For example, in the $SU(2)_k$ Chern-Simons theory, we have

$$\mathcal{D}^{-1} = \sqrt{\frac{2}{k+2}} \sin\left(\frac{\pi}{k+2}\right) \ . \tag{4}$$



(*a*) (*b*)

FIG. 1. (*a*) The plane is divided into four regions, labeled A, B, C, D, that meet at double and triple intersections. (*b*) Moving the triple intersection where B, C, D meet deforms the regions as shown.

To justify computing the entropy using effective field theory, we require that the boundary have no "sharp" features that might be sensitive to short-distance physics; yet for a boundary drawn on a lattice, sharp corners are unavoidable. Furthermore there is an inherent ambiguity in separating the term that scales with the length L from the constant term. We can circumvent these difficulties by the following construction. We now divide the plane into four regions, all large compared to the correlation length, labeled A, B, C, D as in Fig. 1a. Let S_A denote the von Neumann entropy of the density operator ρ_A that is obtained from the ground state by tracing out the degrees of freedom outside region A, let S_{AB} denote the von Neumann entropy of the density operator ρ_{AB} obtained by tracing out the degrees of freedom outside region $AB \equiv A \cup B$, etc. Then we define the topological entropy $S_{\rm topo}$ as

$$S_{\rm topo} \equiv S_A + S_B + S_C - S_{AB} - S_{BC} - S_{AC} + S_{ABC} \ . \tag{5}$$

This linear combination of entropies has been strategically chosen to ensure that the dependence on the length of the boundaries of the regions cancels out. For example, the term proportional to the length of the double intersection of A and D appears in S_A and S_{ABC} with a + sign, and in $-S_{AB}$ and $-S_{AC}$ with a minus sign. Similarly, the double intersection of A and B appears in S_A and S_B with a + sign, and in $-S_{AC}$ and in $-S_{BC}$ with a minus sign. (The observation that the ultraviolet divergent terms cancel in a suitably constructed linear combination of entropies has also been exploited in [6] and applied there to (1+1)-dimensional systems.)

Assuming the behavior eq. (1) in each term, we find $S_{\rm topo} = -\gamma$. But the advantage of defining $S_{\rm topo}$ using a division into four regions is that we can then argue persuasively that $S_{\rm topo}$ is a topological invariant (dependent only on the topology of how the regions join and not on their geometry) and a universal quantity (unchanged by smooth deformations of the Hamiltonian unless a quantum critical point is encountered).

To see that $S_{\rm topo}$ is topologically invariant, first consider deforming the boundary between two regions, far from any triple point where three regions meet. Deforming the boundary between C and D, say, has no effect on regions A, B, and AB; therefore if all regions are large compared to the correlation length, we expect the changes in S_A, S_B, and S_{AB} to all be negligible. Thus the change in $S_{\rm topo}$ can be expressed as

$$\Delta S_{\rm topo} = (\Delta S_{ABC} - \Delta S_{BC}) - (\Delta S_{AC} - \Delta S_C) \ . \tag{6}$$

We expect, though, that if the regions are large compared to the correlation length, then appending region A to BC should have a negligible effect on the *change* in the entropy, since A is far away from where the deformation is occuring; similarly, appending A to C should not affect the change in the entropy. Thus both terms on the right-hand side of eq. (6) vanish, and $S_{\rm topo}$ is unchanged. The same reasoning applies to the deformation of any other boundary between two regions.

Next consider deforming the position of a triple point, such as the point where B, C, and D meet as in Fig. 1b. Again we may argue that S_A is unchanged by the deformation. We recall that for a bipartite pure state (like the ground state), the marginal density operators for both subsystems have the same nonzero eigenvalues and therefore the same entropy; thus $S_{ABC} = S_D$ and $S_{BC} = S_{AD}$. We see that the change in $S_{\rm topo}$ can be expressed as

$$\begin{aligned} \Delta S_{\rm topo} = &(\Delta S_B - \Delta S_{AB}) \\ &+(\Delta S_C - \Delta S_{AC}) \\ &+(\Delta S_D - \Delta S_{AD}) \ . \end{aligned} \tag{7}$$

All three terms on the right-hand side of eq. (7) vanish because appending A does not affect the change in the entropy. The same reasoning applies when any other triple point moves; we conclude that $S_{\rm topo}$ is unchanged by any deformation of the geometry of the regions that preserves their topology, as long as all regions remain large compared to the correlation length.

Now, what happens to $S_{\rm topo}$ as the Hamiltonian of the system is smoothly deformed? We assume that the Hamiltonian is a sum of local terms, and that the correlation length remains finite during the deformation (and in fact that the correlation length stays small compared to the size of regions A, B, C, D). If the Hamiltonian changes locally in a region far from any boundary, then this change has a negligible effect on the ground state in the vicinity of the boundary, and therefore does not

affect $S_{\rm topo}$. If the Hamiltonian changes locally close to a boundary, we can exploit the topological invariance of $S_{\rm topo}$ to first move the boundary far way, then deform the Hamiltonian, and finally return the boundary to its original location. Thus we see that $S_{\rm topo}$ is a universal quantity characteristic of a particular kind of topological order, which remains invariant if no quantum critical point is encountered as the Hamiltonian varies.

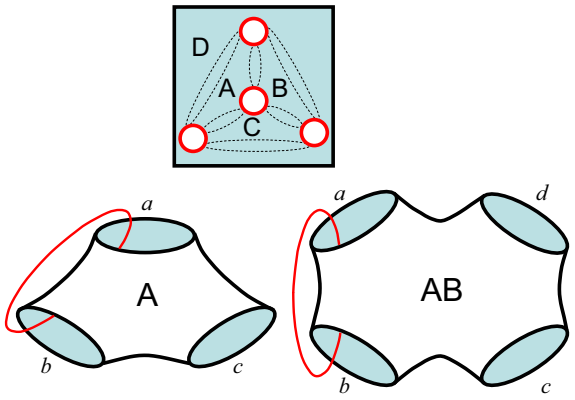


FIG. 2. The planar medium is glued at spatial infinity to its time-reversal conjugate, and wormholes are attached that connect the two conjugate media at the locations of the triple intersections, creating a sphere with four handles. Each region, together with its image, becomes a sphere with three punctures, and each union of two regions, together with its image, becomes a sphere with four punctures. The punctures carry charges labeled a, b, c, d. Anyons that wind around a cycle enclosing a wormhole throat as shown detect a trivial charge.

To facilitate the computation of $S_{\rm topo}$, it is convenient to imagine joining together the planar medium we wish to study with its time-reversal conjugate. We glue together the medium and its conjugate at spatial infinity, and then attach "wormholes" that connect the two planes at the positions of the four triple intersections, as indicated in Fig. 2. The resulting closed surface has the topology of a sphere with four handles. If an isolated planar medium is punctured, then massless chiral modes propagate around the edge of the puncture, but the edge states of the medium and its conjugate have opposite chirality, so the edge states acquire masses when the two surfaces are coupled; therefore, the wormholes can be created adiabatically without destroying the mass gap. No anyons are produced during this adiabatic process, so that the mouth of each wormhole carries trivial anyonic charge.

The boundaries that separate regions in the plane and in its double can be joined through the wormholes as in Fig. 2; then each region of the doubled surface has the topology of a sphere with three punctures, and each union of two adjacent regions becomes a sphere with four punctures. The topological entanglement entropy of the medium and its conjugate are both equal to $S_{\rm topo}$, so that the topological entanglement entropy of the doubled surface is twice $S_{\rm topo}$. The entropy of a region depends only on its topology, so for the doubled surface we have

$$2S_{\rm topo} = 4S_3 - 3S_4 \ , \tag{8}$$

where S_3 denotes the entropy for the sphere with three punctures and S_4 denotes the entropy for the sphere with four punctures.

The quantities S_3 and S_4 can be computed using the appropriate effective field theory, a TQFT [7]. We use the property that no charge is detected by an anyon that winds around the throat of a wormhole. A cycle that encloses a puncture in the double of (say) region A is complementary to a cycle that winds around the wormhole throat; it follows that the puncture carries charge a with probability

$$p_a = |\mathcal{S}_1^a|^2 = d_a^2/\mathcal{D}^2 \ , \tag{9}$$

where $\mathcal{S}_b^a$ is the topological S-matrix of the TQFT, and 1 denotes the trivial charge. To find the joint probability distribution p_{abc} governing the charges a, b, c on the punctures of the sphere with three punctures, we may use standard TQFT methods to compute the probability $p_{ab\to\bar{c}}$ that when charges a and b fuse the total charge is $\bar{c}$. The result is

$$p_{ab\to\bar{c}} = N_{abc} d_c/d_a d_b \ , \tag{10}$$

where N_{abc} is the dimension of the fusion vector space spanned by all the distinguishable ways in which charges a, b, and c can fuse to yield trivial total charge; it follows that

$$p_{abc} = p_a p_b \cdot p_{ab\to\bar{c}} = N_{abc} d_a d_b d_c/\mathcal{D}^4 \ . \tag{11}$$

Evaluating the entropy in the basis in which each puncture has a definite charge, and summing over all the distinguishable fusion states that occur for specified values of the charges, we find

$$\begin{aligned} S_3 &= \sum_{a,b,c} \sum_{\mu=1}^{N_{abc}} -\frac{p_{abc}}{N_{abc}} \log\left(\frac{p_{abc}}{N_{abc}}\right) \\ &= 4\log\mathcal{D} - \sum_{a,b,c} p_{abc} \log\left(d_a d_b d_c\right) \\ &= 4\log\mathcal{D} - 3\sum_a p_a \log d_a \ . \end{aligned} \tag{12}$$

For the sphere with four punctures, a similar calculation yields

$$p_{abcd} = p_a p_b p_c \cdot p_{abc\to\bar{d}} = N_{abcd} d_a d_b d_c d_d/\mathcal{D}^6 \tag{13}$$

and

$$S_4 = 6\log\mathcal{D} - 4\sum_a p_a \log d_a \ . \tag{14}$$

Plugging into eq. (8), we find

$$S_{\rm topo} = 2S_3 - \frac{3}{2}S_4 = -\log \mathcal{D} \equiv -\gamma \ . \qquad (15)$$

Eq. (15) is our main result. Note that it follows if we use eq. (1) to evaluate the entropy of each region, since γ appears four times in the expression for $S_{\rm topo}$ with a negative sign and three times with a positive sign. We also observe that $S_{\rm topo}$ actually depends on the topology of the regions A, B, C. For example, consider the arrangement shown in Fig. 3, in which B and AC both have two connected components, and ABC is not simply connected. Since regions B, AC, and ABC each have boundaries with two components, now γ appears six times with a negative sign and four times with a positive sign, so that $S_{\rm topo} = -2\gamma$.

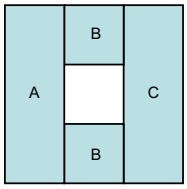


FIG. 3. If the regions A, B, C have the topology shown, then $S_{\rm topo} = -2\gamma$.

Using a different approach, we can formulate a simpler but more heuristic derivation of the formula for γ. First we write the marginal density operator ρ for the disk as $\rho = e^{-\beta H}$. This is just a definition of H and has no other content in itself; furthermore the parameter $T = \beta^{-1}$ is arbitrary — so we are free to choose it to be small compared to the bulk energy gap of the two-dimensional medium. Now we make a natural but nontrivial assumption: that H can be regarded as the Hamiltonian of a $(1+1)$-dimensional conformal field theory (CFT). This CFT ignores short-distance properties of the bulk medium, and therefore will not account correctly for the term in the entropy proportional to L, but it should reproduce correctly the universal constant term.

To compute the entropy for the case of a disk that contains an anyon with charge a (far from the boundary), we evaluate the partition function $Z_a = {\rm tr}_a e^{-\beta H}$ for the associated conformal block of the CFT. Z_a can be expressed as a path integral on a torus of length β in the Euclidean time direction and length L in the spatial direction, in the presence of a Wilson loop carrying anyon charge a that winds through the interior of the torus in the timelike direction. After a modular transformation, we have

$$Z_a = \sum_b \mathcal{S}_a^b \tilde{Z}_b, \qquad (16)$$

where $\tilde{Z}_b$ is the partition function for the b block on a torus of length L in the Euclidean time direction and length β in the spatial direction, and $\mathcal{S}$ is the modular S-matrix of the CFT, which matches the topological S-matrix of the anyon model. In the limit $L \to \infty$, the sum is dominated by the trivial block $\tilde{Z}_1$, and we find

$$\log Z_a \approx \log \left(\mathcal{S}_a^1 \tilde{Z}_1 \right) \approx \log \mathcal{S}_a^1 + \frac{\pi}{12}(c+\bar{c})L/\beta \ , \qquad (17)$$

where c and $\bar{c}$ are the holomorphic and antiholomorphic central charges of the CFT, and $S_a^1 = d_a/\mathcal{D}$ is a topological S-matrix element. Applying the thermodynamic identity $S = -\partial F/\partial T$ (where $F = -T \log Z$ is the free energy), we then find

$$S(\rho) = \frac{\partial}{\partial T}\left(T \log Z\right) = \alpha L - \log\left(\mathcal{D}/d_a\right) \ . \qquad (18)$$

Thus when a is the trivial charge and $d_a = 1$, we recover the result of eq. (1) and (2). While this derivation is not on so firm a footing as the derivation leading to eq. (15), it is more transparent and it generalizes readily to the case where the disk contains an anyon.

We have found an intriguing connection between entanglement entropy and topological order in two dimensions. We note that there are close mathematical ties between the topological entanglement entropy and the (1+1)-dimensional boundary entropy discussed in [8], and we expect that further insights can be derived from studying higher-dimensional analogs of $S_{\rm topo}$. We also hope that our results can provide guidance for the important task of constructing explicit microscopic models that realize topological order.

Results similar to ours have been obtained independently by Levin and Wen [9].

We thank Anton Kapustin for discussions. This work has been supported in part by: the Department of Energy under Grant No. DE-FG03-92-ER40701, the National Science Foundation under Grant No. PHY-0456720, the Army Research Office under Grant Nos. W911NF-04-1-0236 and W911NF-05-1-0294, and the Caltech MURI Center for Quantum Networks under ARO Grant No. DAAD19-00-1-0374.

[1] X.-G. Wen and Q. Niu, Phys. Rev. B **41**, 9377–9396 (1990).
[2] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373–383 (1986); M. Srednicki, Phys. Rev. Lett. **71**, 666–669 (1993).
[3] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395–1398 (1983).
[4] A. Kitaev, Ann. Phys. **303**, 2–30 (2003).
[5] A. Hamma, R. Ionicioiu, and P. Zanardi, Phys. Rev. A **71**, 022315 (2005).
[6] H. Casini and M. Huerta, Phys. Lett. B **600**, 142–150 (2004).
[7] E. Witten, Comm. Math. Phys. **121**, 351–399 (1989).
[8] I. Affleck and A. Ludwig, Phys. Rev. Lett. **67**, 161–164 (1991); J. Cardy, Nucl. Phys. B **328**, 581–596 (1989).
[9] M. Levin and X.-G. Wen, arXiv:cond-mat/0510613 (2005).

Topological Phases of One-Dimensional Fermions: An Entanglement Point of View

Ari M. Turner
Department of Physics, University of California, Berkeley CA 94720, USA

Frank Pollmann
Department of Physics, University of California, Berkeley CA 94720, USA and
Institute of Physics, Academia Sinica, Taipei 11529, Taiwan

Erez Berg
Department of Physics, Harvard University, Cambridge, MA 02138, USA
(Dated: March 20, 2012)

The effect of interactions on topological insulators and superconductors remains, to a large extent, an open problem. Here, we describe a framework for classifying phases of one-dimensional interacting fermions, focusing on spinless fermions with time-reversal symmetry and particle number parity conservation, using concepts of entanglement. In agreement with an example presented by Fidkowski *et. al.* (Ref. [1]), we find that in the presence of interactions there are only eight distinct phases, which obey a $\mathbb{Z}_8$ group structure. This is in contrast to the $\mathbb{Z}$ classification in the non-interacting case. Each of these eight phases is characterized by a unique set of bulk invariants, related to the transformation laws of its entanglement (Schmidt) eigenstates under symmetry operations, and has a characteristic degeneracy of its entanglement levels. If translational symmetry is present, the number of distinct phases increases to 16.

I. INTRODUCTION

Topological phases of matter are not characterized by a broken symmetry, but rather by an underlying topological structure that distinguishes them from other, topologically trivial phases. Such phases have attracted a great deal of attention recently, especially since the theoretical prediction[2–5] and subsequent experimental observation[6,7] of both two- and three-dimensional realizations of time-reversal invariant topological insulators. These (as well as their predecessor, the integer quantum Hall effect) can be thought of as band insulators characterized by the topological structure of their Bloch bands. Similarly, topological superconductors[8–10] are characterized by the topological nature of their fermionic quasi-particle spectrum. All these systems can be understood from a non-interacting point of view. A complete classification of all topological phases of non-interacting fermions, given their symmetries, has been given in Refs. 9–11.

In the presence of electron-electron interactions, the Hamiltonian cannot be reduced to a single particle matrix. Therefore, strictly speaking, the above classification scheme of topological phases cannot be used. Nevertheless, in some classes of topological insulators, the topological order has been argued to be robust even in the presence of interactions, by generalizing the corresponding topological invariant to the many-body case[11–13]. In other classes, however, the situation in the interacting case remains unclear.

In a recent breakthrough, Fidkowski and Kitaev studied a one-dimensional model of spinless superconductors with time reversal symmetry.[1] They found that in the presence of interactions, the free-fermion classification breaks down from $\mathbb{Z}$ to $\mathbb{Z}_8$, *i.e.* there are only eight distinct phases that survive in the presence of interactions (as opposed to an infinite number without interactions). To the best of our knowledge, this is the first case where the non-interacting picture in a class of topological phases is found to be radically modified by interactions. Ref. 1 constructs an explicit path in Hamiltonian space through which phases with different $\mathbb{Z}$ numbers mod(8) can be connected, and also discusses the stability of the edge states. However, a more general understanding of the classification of distinct phases in the presence of interactions (in particular, in terms of *bulk* properties of the ground state wavefunction) is left open.

In this paper, we develop a framework for classifying phases of interacting fermions in one dimension based on bipartite entanglement of the ground states wave function. The fact that entanglement is a useful quantity to probe topological properties of wave functions has been shown in several recent publications, see for example Refs. 14–17. Our technique is based on a method which was introduced in Ref. 18 for classifying phases in spin systems spins. This method has also been developed more fully and shown to give a *complete* 1D classification by Ref. 19 (at least when translational symmetry is not required). Here we generalize the method to fermionic systems. We find that the eight phases found in Ref. 1 are indeed topologically distinct, and characterize them in terms of a set of invariants. These phases cannot be continuously connected by any kind of interaction as long as time-reversal symmetry and fermion parity conservation are preserved.

The basic idea is to examine the behavior of the entanglement (Schmidt) eigenstates of a segment in the bulk of the system under the symmetry group of the system. Topologically nontrivial phases can be recognized by the presence of "fractionalized" modes in the entanglement

spectrum, which transform differently under the symmetry group from the constituent microscopic degrees of freedom of the system (analogous to the half-integer spins at the ends of the spin one Heisenberg chain). The character of the entanglement spectrum cannot change without a bulk phase transition, at which the nature of the ground state changes abruptly or the correlation length diverges.

The behavior of the entanglement modes reflects the character of the *physical* topologically-protected modes at the boundary of the system. However, unlike the edge modes, the entanglement spectrum represent a truly *bulk* property of the ground state wavefunction, and as such, it is not sensitive to symmetry-breaking perturbations at the surface.

We start in Sec. II by introducing fermionic Hamiltonians with pairing terms, through the example of a single Majorana chain model. The general framework to classify topological phases based on symmetry properties of the entanglement eigenstates is presented in Section III. We apply it to fermionic systems with time-reversal invariance and fermion number parity conservation, and derive the invariants characterizing the eight distinct phases and the degeneracies in their entanglement spectrum. These phases are shown in the next section to have a $\mathbb{Z}_8$ group structure, defined through the rules for combining phases with different invariants. In Sec. IV, we demonstrate how to construct each phase by combining single chains. In Sec. V, we discuss the additional phases which arise if translational symmetry is imposed. The results are summarized and discussed in Sec. VI.

II. FERMIONIC MODELS WITH PAIRING TERMS

We will investigate time-reversal invariant one-dimensional systems of spinless fermions, in which the particle number is conserved modulo two. (The classification of topological phases is most interesting in this case.) Such a situation can be realized in a system in contact with a superconductor. As a simple example, consider the following Hamiltonian[20]:

$$H_0 \;=\; -\frac{t}{2}\sum_j \left(c_j^\dagger c_{j+1}^\dagger + c_j^\dagger c_{j+1} + \text{H.c.}\right) + u\sum_j c_j^\dagger c_j, \tag{1}$$

with $t, u \geq 0$. The operators $c_j^\dagger$ (c_j) create (annihilate) a spinless fermion on site j. The first term comprises hopping of fermions as well as the creation and annihilation of pairs of fermions while the second term acts as a chemical potential. The fermion parity operator

$$Q = e^{i\pi \sum_j n_j}$$

with $n_j = c_j^\dagger c_j$ commutes with H_0 as the total number of fermions N_{total} modulo two is conserved. Furthermore, the Hamiltonian is time-reversal symmetric. (For spinless fermions, time reversal is represented by complex conjugation.)

Let us begin by considering only the conservation of the fermion number parity. This symmetry of H_0 allows us to distinguish two different phases. The system undergoes phase transitions at $t = u$, but *no* local order parameter can be used to distinguish the two phases on either side of the transition. However, they can be distinguished by their edge states. In the phase $u > t$, the ground state for an open chain is unique while it is two-fold degenerate for $t > u$.[20] If $u = 0$, one can check that these states are given by the equal weighted superposition of all configurations with fixed fermion parity (i.e., an even number or an odd number of particles). The ground state degeneracy in this case can be understood in term of degrees of freedom at the two ends. The two ground states cannot be distinguished by *any* local observable in the bulk, because in any finite region of either state, the parity can be either even or odd. However, the two states *are* distinguishable when the opposite ends are compared to one another: $c_N^\dagger c_1 + c_N^\dagger c_1^\dagger + \text{H.c.}$ has a different eigenvalue for the two states. Furthermore, we can transform the two ground states into each other by acting with either $c_1 + c_1^\dagger$ or $c_N - c_N^\dagger$ on the two ends of the chain. In other words, there is a single fermionic state that is split between the ends of the chain, and the observable described above measures its occupation number. The two states are degenerate because the only distinction between them is long-range, while energy measures only local correlations. In the phase $t < u$, however, there is no such degeneracy. The picture provided remains true even if we include interactions as the arguments can be stated in a way that only requires the Hamiltonian to conserve the fermion parity.

The edge properties have a simple explanation in a different representation defined by the following transformation:

$$a_j \;=\; c_j + c_j^\dagger \tag{2}$$

$$b_j \;=\; -i(c_j - c_j^\dagger). \tag{3}$$

a_j and b_j are Majorana operators; they obey the relations $\{a_i, a_j\} = \{b_i, b_j\} = \delta_{ij}$, $\{a_i, b_j\} = 0$, $a_i = a_i^\dagger$ and $b_i = b_i^\dagger$. The fermion parity, $1 - 2n_j$ of a site, is given by $ib_j a_j$. Using these operators, H_0 can be written (up to a constant) as

$$H_0 = \frac{i}{2}\left(t\sum_j b_j a_{j+1} + u\sum_j a_j b_j\right). \tag{4}$$

Observe that each unit cell now contains two operators. In the case where $t = 0, u = 1$, the ground state is described by $ia_j b_j = -1$, i.e., each site is vacant. In terms of these variables, the phase $t = 1$, $u = 0$ is also simple: $ib_j a_{j+1} = -1$ when t is large. (This can be regarded as the parity of a fermion shared between sites j and

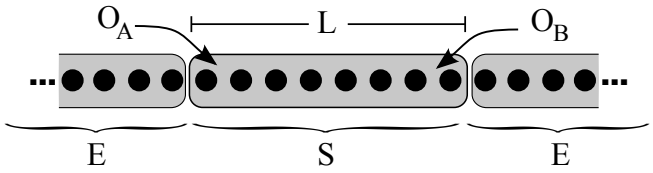


FIG. 1: Illustration of a bipartition of a 1D chain into a segment (S) of length L and a surrounding environment (E). The operators $\mathcal{O}_A$ and $\mathcal{O}_B$ act on the edges of the segment.

$j+1$.) This requirement does not completely determine the ground state wave function in an open chain, though, because it leaves a_1 and b_N free. There are therefore two degenerate states characterized by the occupation of the fermion shared between the ends, ib_Na_1.

The presence of time-reversal symmetry leads to additional distinctions between phases. Quadratic, time reversal invariant fermionic Hamiltonians with conservation of the fermion number mod (2) have been shown[9,10] to support phases classified by an integer $n \in \mathbb{Z}$ (class BDI, according to Ref. 9). Each phase is characterized by having n gapless Majorana modes at each edge, and the different phases cannot be smoothly connected to each other without closing the bulk gap. It was later found[1] that in the presence of interactions this classification breaks down to $\mathbb{Z}_8$. Now we will begin the main discussion, whose goal is to show how the eight distinct phases in the general (interacting) case can be understood and classified according to properties of their entanglement eigenstates under the symmetries of the system, namely time-reversal and fermion number parity conservation.

III. CLASSIFYING PHASES BY SYMMETRY PROPERTIES OF THE ENTANGLEMENT EIGENSTATES

In the preceding section, we discussed physical edge properties to characterize different phases. Below, we present an alternative method of classifying the phases, which *involves only the bulk*. This is achieved by examining the "entanglement spectrum"[16,18,21–23] of a finite (but arbitrarily large) segment[14,15], embedded in an infinite system. The transformation law of the entanglement (or Schmidt) states under the symmetry group of the system can be used to distinguish between different phases, as we describe below.

We consider a bipartition of a 1D chain with periodic boundary conditions into a segment (S) of length L and a surrounding environment (E) of length $N \gg L$ as shown in Figure 1. For the segment S, the reduced density matrix of the ground state wave function $|\psi\rangle$ is given by

$$\rho_S = \mathrm{tr}_E \left(|\psi\rangle\langle\psi| \right). \tag{5}$$

It is convenient to define an "entanglement Hamiltonian" $\mathcal{H}_S$ such that

$$\rho_S = e^{-\mathcal{H}_S}. \tag{6}$$

Then, the low "energy" states of $\mathcal{H}_S$ are the most likely states of the segment S, when the entire system is in its ground state. We call the eigenstates $|\phi_\gamma\rangle_S$ of $\mathcal{H}_S$ "entanglement eigenstates" and the eigenvalues E_γ "entanglement eigenvalues". (These are the same as the Schmidt states defined by $|\psi\rangle = \sum e^{-\frac{E_\gamma}{2}} |\phi_\gamma\rangle_S |\chi_\gamma\rangle_E$, where $|\chi_\gamma\rangle_E$ are the corresponding Schmidt states of the environment.)

Our approach is based on an important observation for the entanglement Hamiltonian $\mathcal{H}_S$: The low "entanglement energy" excitations of $\mathcal{H}_S$ in a d-dimensional system may be well described by a $d-1$ dimensional effective Hamiltonian (see also Ref. 24 for an interesting discussion of this concept). We consider the entanglement spectrum of a sufficiently large one-dimensional segment S, and focus on the "low-lying" entanglement states with $E_\gamma < E_{\mathrm{cut}}$ (where E_{cut} is an arbitrary constant).

We now make a crucial observation: in a gapped system with a finite correlation length ξ,[25] these states can only be distinguished from each other by their behavior within a certain distance of the ends of the segment S. This is justified by the following argument: suppose that we measure a correlation function $C(\ell) = \langle \mathcal{O}_E \mathcal{O}_S \rangle_\psi$, where $\mathcal{O}_E$ acts on sites in the environment E and $\mathcal{O}_S$ acts on sites in the segment (far away from the edges), respectively. ℓ is the minimal distance between the sites on which $\mathcal{O}_E$ and $\mathcal{O}_S$ act. We expect that as ℓ becomes large, $C(\ell) \to \langle \mathcal{O}_E \rangle_\psi \langle \mathcal{O}_S \rangle_\psi$. Let us take for $\mathcal{O}_E$ an operator that projects onto a particular Schmidt state of the environment, say $|\chi_\gamma\rangle_E$. ($\mathcal{O}_S$ can be any operator as long as it acts far from the edges of S.) When applied to the ground state, $\mathcal{O}_E$ projects (through the entanglement between E and S) also onto the corresponding entanglement eigenstate $|\phi_\gamma\rangle_S$ of S. Thus, we have $\langle \mathcal{O}_S \rangle_{\phi_\gamma} \approx \langle \mathcal{O}_S \rangle_\psi$. That is, $\langle \mathcal{O}_S \rangle$ in each eigenstate of $\mathcal{H}_S$ is the same as in the ground state; *i.e.*, far enough from the edge of S, *any eigenstate of $\mathcal{H}_S$ behaves essentially like the ground state.* Therefore, an operator acting far from the edge of S cannot distinguish different Schmidt states; its expectation value must be the same in all of them. (A simple generalization of the above argument, using an off-diagonal $\mathcal{O}_E$, shows that $\mathcal{O}_S$ cannot connect different low-energy entanglement eigenstates if it is sufficiently far from the boundary.)

Furthermore, one can show (see App. A for the case of bosons) that any linear transformation applied to the subspace of low-entanglement energy eigenstates of S can be represented by a local operator acting on sites within a distance ℓ from either boundary of S, with an accuracy that improves exponentially with ℓ. In particular, every *symmetry operation* of the system is a unitary operation acting on the entanglement eigenstates,[26] and thus can be represented (within the low-entanglement basis) as a product of two operators $\mathcal{O}_A$, $\mathcal{O}_B$ (Fig. 1), acting on sites near the left and right edges of S, respectively. (Note

that $\mathcal{O}_A$, $\mathcal{O}_B$ are specific for a particular ground state $|\psi\rangle$.) Thus, $\mathcal{O}_A$ and $\mathcal{O}_B$ form a (projective) representation of the symmetry group. Classifying the representations formed by $\mathcal{O}_A$, $\mathcal{O}_B$ distinguishes different phases, which cannot be adiabatically connected unless a phase boundary is crossed. At the phase boundary, either the character of the ground state changes discontinuously, or else the correlation length diverges and the two ends are no longer independent.

Let us demonstrate these principles through the case of an SU(2)-symmetric integer spin chain. If the segment S is sufficiently long, there can be situations in which entanglement states come in degenerate sets[18] This can be seen from the fact that, according to the argument above, any SU(2) rotation can be represented accurately (within the low-entanglement energy subspace) in terms of two generators, $\mathbf{S}_A$ and $\mathbf{S}_B$, which act only within a distance ℓ of the left and right edges of S, respectively. Within this subspace, $\mathbf{S}_{tot} \sim \mathbf{S}_A + \mathbf{S}_B$, where $\mathbf{S}_{tot}$ is the total spin. Both $\mathbf{S}_A$, $\mathbf{S}_B$ can be block-diagonalized into irreducible representations of SU(2) with well-defined angular momenta. Now, since $\mathbf{S}_{tot}$ is an integer spin, there are two possibilities: either all the blocks in $\mathbf{S}_A$ and $\mathbf{S}_B$ form integer representations, or they all form half-integer representations. This distinguishes two phases: a "trivial" phase in which $\mathbf{S}_A$, $\mathbf{S}_B$ are both integer (*e.g.*, a fully dimerized chain), and a "non-trivial" phase in which they are half-integer (such as the Haldane phase of the spin-one chain[27]). In the latter phase, all the entanglement energies of S must have a degeneracy of at least four[18], due to the even degeneracy at each end, guaranteed by the presence of the half-integer spin operators $\mathbf{S}_A$ and $\mathbf{S}_B$. This is an alternative explanation of the symmetry–protection of the Haldane phase, discussed in Refs. 18,28–31.

A. Fermion parity

We now turn to discuss fermionic systems. Let us consider a Hamiltonian H that conserves the fermion parity Q as defined in (II) with $Q^2 = \mathbb{1}$. We show that one can distinguish two phases, a "trivial" phase and a "non-trivial" one. In the "non-trivial" phase, a segment's entanglement spectrum is doubly degenerate. The double degeneracy is related to a single fermionic degree of freedom, which is split between the opposite ends of the segment.

Any eigenstate of H, and in particular the ground state $|\psi\rangle$, is also an eigenstate of Q. Hence the resulting reduced density matrix ρ_S and the entanglement Hamiltonian $\mathcal{H}_S$ both commute with Q. The eigenstates $|\phi_\gamma\rangle$ of $\mathcal{H}_S$ may therefore be classified by their Q-eigenvalues ($q^S = +1$ if the fermion number N_F is even and $q^S = -1$ if N_F is odd.)

We now posit that it is possible to find an effective expression for Q within the low-entanglement energy subspace, of the form $Q \propto Q^A Q^B$, where Q^A and Q^B are local operators which act near the left and right edges of S, respectively. This is analogous to the example of spin described in the previous section, in which the total spin can be represented as a sum of operators acting on the left and right edges. [In the fermionic case, the decoupling of Q is multiplicative, since Q is itself a unitary symmetry rather than an SU(2) generator.]

Now, Q^A and Q^B can have interesting relationships: Q^A and Q^B may be two fermionic operators (that is, each contains an odd number of creation or annihilation operators) or they may both be bosonic. Note that Q^A or Q^B cannot contain a sum of bosonic and fermionic terms, becasue $Q \propto Q^A Q^B$ is bosonic.[42]

If Q^A, Q^B contain an odd number of fermionic operators, $Q^A Q^B = -Q^B Q^A$. To emphasize this "statistical correlation" between the two ends define an angle $\mu = 0$ or π to distinguish between the two cases, so that

$$Q^A Q^B = e^{i\mu} Q^B Q^A. \tag{7}$$

For $\mu = 0$ (π), $Q = Q^A Q^B$ ($-iQ^A Q^B$). The factor of i must be introduced in the latter case for consistency with the anticommutation of Q^A and Q^B, given that $Q^2 = \left(Q^A\right)^2 = \left(Q^B\right)^2 = \mathbb{1}$.

For example, consider the Hamiltonian in Eq. 1 with $u = 0$, $t = 1$. In this case, it is not difficult to show that there are only two entanglement eigenstates on the segment S (with a nonzero weight in the density matrix), the same two states as the ground states of the original Hamiltonian restricted to this segment. The fermion parity of these states is given by

$$Q = -iQ^A Q^B, \tag{8}$$

where

$$\begin{aligned} Q^A &= c_1 + c_1^\dagger \\ Q^B &= -i(c_L - c_L^\dagger). \end{aligned} \tag{9}$$

We see immediately that Q^A and Q^B anti-commute, and therefore $\mu = \pi$. Furthermore, $\left[\mathcal{H}_S, Q^A\right] = \left[\mathcal{H}_S, Q^B\right] = 0$.

For any system with $\mu = \pi$, the above commutation relations imply that all the eigenvalues of $\mathcal{H}_S$ come in degenerate pairs. To see this, note that Q and $\mathcal{H}_S$ can be diagonalized simultaneously. Then, if $\mathcal{H}_S|\phi_\lambda\rangle = E_\lambda|\phi_\lambda\rangle$ and $Q|\phi_\lambda\rangle = q_\lambda|\phi_\lambda\rangle$ (where $q_\lambda = \pm 1$), then the state $|\psi_\lambda\rangle = Q^A|\phi_\lambda\rangle$ is such that $\mathcal{H}_S|\psi_\lambda\rangle = E_\lambda|\psi_\lambda\rangle$ and $Q|\psi_\lambda\rangle = -q_\lambda|\psi_\lambda\rangle$, *i.e.* $|\psi_\lambda\rangle$ is an independent eigenstate with eigenvalue E_λ. Indeed, for the Hamiltonian (1) with $u = 0$, we find a doubly degenerate entanglement level with $E = \log 2$. (All other entanglement levels in that system have $E = \infty$.)

Note that, unlike the Haldane chain example in Sec. III, the entanglement spectrum is two-fold (rather than four-fold) degenerate. This is a consequence of the fact that the degeneracy is not associated with either Q^A or Q^B alone; it is related to the occupation of the fermionic level formed by combining $Q^A + iQ^B$, *i.e.*, it is *shared* between the two edges.

In a bosonic system, the states of the entanglement eigenstates can be represented by $|\phi_\gamma\rangle = |\alpha\beta\rangle$. Here α, β describe the states of the left and right ends of the chain; that is, they enumerate the eigenvalues of certain low-"energy" combinations of observables that are functions of ℓ sites at the respective ends. This factorization is possible for fermionic chains with $\mu = 0$ as well. However, for chains with $\mu = \pi$, the extra q_S variable describing the two-fold degeneracy cannot be written in terms of local observables belonging to either end (*i.e.*, the fermionic degrees of freedom cannot be measured independently at the two ends of the segment). Therefore, the entanglement states should be labelled by three variables, $|\alpha\beta q_S\rangle$, with the fermion parity q_S of the entire chain represented explicitly.[43]

In a noninteracting system, the entanglement Hamiltonian is also noninteracting.[32,33] It can be represented in terms of entanglement modes. The only subtlety is that some of these modes may be Majorana modes, which satisfy $m^2 = 1$; $m^\dagger = m$. There are two topologically distinct phases, depending on whether there are an even or an odd number of Majorana modes at each end of the segment. The Q^A and Q^B operators defined above can be found explicitly. Given that the left edge has N_f low–energy fermionic entanglement modes, $f_{A,\alpha}$ ($\alpha = 1, \ldots, N_f$), and N_m Majorana modes, $m_{A,\beta}$ ($\beta = 1, \ldots, N_m$), Q^A is given by

$$Q^A = \left(\prod_\alpha (-1)^{f^\dagger_{A,\alpha} f_{A,\alpha}} \right) \left(\prod_\beta m_{A,\beta} \right), \qquad (10)$$

and similarly for Q^B. Note that Q^A is a bosonic operator if N_m is even (corresponding to $\mu = 0$) and a Majorana operator if N_m is odd ($\mu = \pi$).

When interactions are included, only the "total parities" Q^A, Q^B are well-defined. The separate modes f_α, m_β can "decay" into other combinations of modes, but their total is closely related to the symmetry Q, as we have just outlined.

B. Time Reversal Symmetry

We shall now examine the consequences of time-reversal symmetry on the degeneracies of the entanglement energies. It turns out that the combination of time-reversal and fermion parity conservation can have nontrivial effects.

Bosonic models. Let us introduce the approach by reiterating the results for a bosonic chain in the presence of time reversal symmetry with $[\mathcal{H}_S, T] = 0$. The eigenstates $|\phi_\gamma\rangle$ can be represented by $|\phi_\gamma\rangle = |\alpha\beta\rangle$, where α, β enumerate the low-"energy" states associated with the two edges. Entanglement eigenstates which differ in their α (β) index can be connected by a local operator close to the left (right) edge, respectively. The transformation of the eigenstates of $\mathcal{H}_S$ factors into parts referring to the two ends (see App. A).

It can therefore be represented as a product of two unitary transformations U^A, U^B acting on the ends of the segment, so that

$$T|\alpha, \beta\rangle = \sum_{\alpha', \beta'} U^A_{\alpha'\alpha} U^B_{\beta'\beta} |\alpha'\beta'\rangle \qquad (11)$$

and $[U^A, U^B] = 0$. For a discussion of subtleties related to the anti-unitarity of T, see App. B. Applying T to an eigenstate twice yields

$$T^2|\alpha\beta\rangle = \sum_{\alpha', \beta', \alpha'', \beta''} U^A_{\alpha''\alpha'} U^B_{\beta''\beta'} (U^A)^*_{\alpha'\alpha} (U^B)^*_{\beta'\beta} |\alpha''\beta''\rangle. \qquad (12)$$

Using $T^2 = \mathbb{1}$ and that the two ends of the segment are independent, it follows that

$$U^A (U^A)^* = U^B (U^B)^* = \exp(i\kappa)\mathbb{1}, \ \kappa = 0, \pi \qquad (13)$$

We can thus distinguish two different phases, corresponding to $\kappa = 0, \pi$. Let us now focus on the consequences for the entanglement spectrum. Assume that $|\phi_\gamma\rangle$ is an eigenstate of $\mathcal{H}_S$ with eigenvalue E_γ, then $U^A|\phi_\gamma\rangle$, $U^B|\phi_\gamma\rangle$, $U^A U^B|\phi_\gamma\rangle$ are also eigenstates with the same eigenvalue because $\mathcal{H}_S$ commutes with U_A and U_B. If $\kappa = \pi$, the unitaries U^A, U^B are anti-symmetric and thus the four states are mutually orthogonal, resulting in a four fold degeneracy of the entanglement spectrum. If $\kappa = 0$, the entanglement spectrum does not necessarily have any degeneracies. For example, in the Haldane phase of spin-1 chains, we find $\kappa = \pi$ and therefore the entire entanglement spectrum of a segment S is four fold degenerate.[18,44]

This method may be generalized to give a classification of phases with any given set of symmetries. For each relationship between the physical symmetries (e.g., $T^2 = \mathbb{1}$ in the case just described), there is a corresponding relationship between the factored symmetries of the entanglement spectrum[34], in which certain phases (e.g., κ) can appear. Certain combinations of these phases are "gauge invariant" (independent of how the phases of the factored symmetries are chosen). These combinations distinguish between topological phases. In fermionic models, an additional possibility is that symmetry operators at opposite ends may either commute or anticommute, as described in the previous section.

Fermionic models. We now consider a Hamiltonian which has both fermion parity conservation with $Q^2 = \mathbb{1}$ and time reversal symmetry with $T^2 = \mathbb{1}$. In the presence of both symmetries, we show that each of the two phases defined in the previous section ($\mu = 0, \pi$) can be subdivided into four different phases. Furthermore, we discuss the consequences for the entanglement spectrum in each case. As T simply takes the complex conjugate (spin degrees of freedom are not considered here), it does not change the total fermion number and thus $[T, Q] = 0$. We will now classify the phases by examining how the properties of the factored versions of Q and T depart

from the relations of the full transformations, $T^2 = \mathbb{1}$, $[T, Q] = 0$.

We first consider the case $\mu = 0$, *i.e.*, $\left[Q^A, Q^B\right] = 0$. Then, both Q^A and Q^B are bosonic operators, and can be diagonalized simultaneously. Then, we distinguish two cases: $Q^A T = e^{i\phi} T Q^A$ with $\phi = 0, \pi$, and similarly for Q^B. Note that ϕ has to be the same for Q^A and Q^B, since $Q = Q^A Q^B$ satisfies $[T, Q] = 0$. If $\phi = \pi$, time reversal changes the parity of the fermion number in either end. (A similar situation occurs at the vortex cores of time-reversal invariant topological superconductors[35].) We now examine the two cases $\phi = 0, \pi$ separately.

($\mu = 0$, $\phi = 0$, $\kappa = 0$ or π)– The case $\phi = 0$ is analogous to the bosonic case considered above, with two phases, one corresponding to $\kappa = \pi$, characterized by a four-fold degenerate entanglement spectrum of a segment, and one to $\kappa = 0$, in which there is no necessary degeneracy in the entanglement spectrum.

($\mu = 0$, $\phi = \pi$, $\kappa = 0$ or π)– If $\phi = \pi$, U^A and U^B (defined through $T = U^A U^B$; see App. B) are both fermionic operators, since they change the fermion parity. We know that $T^2 = U^A U^B \left(U^A U^B\right)^* = -U^A U^{A*} U^B U^{B*} = \mathbb{1}$. This can only be satisfied if $U^A (U^A)^* = \exp(i\kappa)\mathbb{1}$ and $U^B (U^B)^* = -\exp(i\kappa)\mathbb{1}$ with $\kappa = 0, \pi$. Note that $\left\{Q^A, U^A\right\} = 0$ and $\left\{Q^B, U^B\right\} = 0$, where both U^A and U^B commute with $\mathcal{H}_S$. ($\{\cdot, \cdot\}$ denotes an anti–commutator.) Therefore, each entanglement level is four–fold degenerate, where the degeneracy corresponds to states with all possible combinations of $Q^A = \pm 1$ and $Q^B = \pm 1$.

Next, we consider the case $\mu = \pi$. In this case, $\left\{Q^A, Q^B\right\} = 0$, so these two operators cannot be diagonalized simultaneously. Rather, every entanglement eigenstate can be labelled by the eigenvalue $q = \pm 1$ of the parity operator $Q = -iQ^A Q^B$, where $|\alpha\beta, q = \pm 1\rangle$ are degenerate. Since $[T, Q] = 0$, we must have either $\left[T, Q^A\right] = 0$ and $\{T, Q^B\} = 0$, or vice versa. Therefore, we define a parameter $\phi = 0, \pi$ such that $TQ^A = e^{i\phi} Q^A T$ and $TQ^B = e^{i(\phi+\pi)} Q^B T$.

In this case ($\mu = \pi$), phases with $\phi = 0$ and π behave very similarly. To see this, we just note that if $\phi = \pi$ then QT commutes with Q^A. Therefore, we will define a modification of time reversal that commutes with Q^A, $T' := QT$ if $\phi = \pi$ and $T' := T$ if $\phi = 0$. Let the factors of T' be $T' = U^{A'} U^{B'}$. One can check that $U^{A'}, U^{B'}$ are bosonic. The entanglement spectrum can be divided into two sectors with a fixed value of q. The operator $U^{A'}$, being bosonic, depends only on α, β and hence acts the same way on both sectors. Define κ by $e^{i\kappa}\mathbb{1} = U^{A'} U^{A'*} = U^{B'} U^{B'*}$. The possible values for κ are 0 and π:

($\mu = \pi$, $\phi = 0$ or π, $\kappa = 0$)– If $\kappa = 0$, each entanglement eigenstate in each q sector can be singly degenerate. Therefore, counting the $q = \pm 1$ degeneracy, each entanglement eigenstate has a minimal degeneracy of 2.

($\mu = \pi$, $\phi = 0$ or π, $\kappa = \pi$)– If $\kappa = \pi$, the spectrum in each of the $\pm q$ sectors is fourfold degenerate (for the same reasons as in the bosonic case above with $\kappa = \pi$, *i.e.*, there is a Kramer's doublet at each edge). Taking the $q = \pm 1$ degeneracy into account, every entanglement eigenstate is at least eight fold degenerate.

There are therefore eight different phases classified by triplets (μ, ϕ, κ), where each entry is 0 or π. $\mu = \pi$ if there are Majorana modes at the ends of the segment, $\phi = \pi$ if time reversal and Q^A anticommute, and $\kappa = \pi$ if $U^A (U^A)^* = -\mathbb{1}$ (leading to Kramers' doublets at the edges of the segment). Since, as long as the time-reversal and fermion parity symmetries are preserved, (μ, ϕ, κ) can only take the values 0 or π, they cannot change smoothly; the only way for them to change is through a non-analytic change of the ground state wave function, *i.e.* a quantum phase transition.

The eight different phases and their corresponding minimal degeneracies are summarized in Table I in Section VI. The degeneracies illustrate a distinction between interacting and noninteracting systems. As we will show in Sec. IV, the eight distinct phases can be realized by taking M copies of a single chain in the large t phase, where $M = 1, \ldots, 8$. Without interactions, the degeneracy of the Schmidt spectrum would be equal to 2^M. Interactions can partly lift this degeneracy but cannot connect the eight phases defined by (μ, ϕ, κ) adiabatically.

IV. ADDITION OF PHASES

The eight phases we have just obtained obey a group structure, which is defined by the rules of "adding" them together. This group turns out to be $\mathbb{Z}_8$. The addition rules reveal interesting distinctions between bosons and fermions. We will work out the addition table in some detail.

Two systems can be added together by placing them side by side: hence if one system is in phase P_1 and another is in P_2, then the combined "ladder" system is in phase $P_1 + P_2$. (The combined system then remains in $P_1 + P_2$ even when the two constituent systens are coupled, as long as the coupling Hamiltonian is symmetric under time-reversal and fermion parity, and the bulk gap does not collapse.) This rule creates a finite group for a given set of symmetries. In particular, every element in this group has an inverse and the trivial phase is the identity element.

The distinction between fermionic and bosonic systems is related to the inverse operation. If the system consists only of bosons, then the inverse element of any phase P is its complex conjugate (*i.e.*, its time reversal):

$$P + P^* = \text{trivial phase.} \tag{14}$$

For an example, consider a spin one Heisenberg chain. A single chain cannot be adiabatically connected to the trivial phase because its ends transform as spin 1/2 degrees of freedom. However, as shown in Ref. [36], two coupled chains can be connected continuously to the rung singlet phase (i.e., a product state of spin zeros on the

rungs). The two chains are no longer distinguished from the trivial phase by their ends because the two half-integer spins couple to form integer spin states.

In general, phases of bosonic chains are distinguished by the projective representation of the symmetry groups acting on the entanglement eigenstates (see Ref. 18). Each element in the symmetry group $\sum$ is represented in the entanglement eigenbasis as a left-hand unitary matrix $U^A(\Sigma)$ acting on the left index of the state, and a right-hand matrix $U^B(\Sigma)$ acting on the right index. Then the combined operation of two elements Σ_1 and Σ_2 is represented by $U^A(\Sigma_1\Sigma_2) = e^{i\rho_A(\Sigma_1,\Sigma_2)}U^A(\Sigma_1)U^A(\Sigma_2)$, and similarly for U^B. To see that $P+P^*$ is trivial, consider the eigenstates of the entanglement Hamiltonian for a segment in the combined system

$$|\alpha\beta\rangle_{\text{coupled}} = |\alpha_1\beta_1\rangle_P|\alpha_2\beta_2\rangle_{P^*}. \tag{15}$$

The left-hand matrix U^A representing a symmetry Σ for the coupled system is

$$U^{A,\text{coupled}}_{\alpha_1'\alpha_2';\alpha_1\alpha_2}(\Sigma) = U^A_{\alpha_1'\alpha_1}(\Sigma)U^{A*}_{\alpha_2'\alpha_2}(\Sigma) \tag{16}$$

where the second factor is complex-conjugated because the second chain is time-reversed. Hence, the phase factors cancel, $U^{A,\text{coupled}}(\Sigma_1)U^{A,\text{coupled}}(\Sigma_2) = U^{A,\text{coupled}}(\Sigma_1\Sigma_2)$, and the resulting system is in a trivial phase.

Now we can try to build more complicated phases out of simpler ones by placing them side-by-side. For bosonic systems, this procedure does not generate new phases in the presence of time reversal symmetry. Time reversal symmetry and Eq. (14) imply that $P+P=0$. Hence starting from one phase, it is not possible to get more than two phases (the original phase and the trivial phase). There may be additional phases that would have to be built up from independent starting points. The group is always a product of $\mathbb{Z}_2$'s, in other words.

However, for fermionic spin chains, P and P^* are not necessarily inverses. The p-wave superconducting state P_1 described by Eq. (1) with $t>\mu$ which is an order eight phase, as discovered by Fidkowski and Kitaev[1], is an illustration. Eq. (16) breaks down because operators on the two chains can anticommute with each other. In fact, starting from a single Majorana chain, we can generate all possible combinations of μ, ϕ, and κ. We now demonstrate this idea for a number of examples:

1. Consider the Majorana chain, the ground state of Eq. (1) with $t>u$ which is in the $(\mu,\phi,\kappa) = (\pi,0,0)$ phase. When two copies are combined together, the resulting phase has $(\mu,\phi,\kappa)_{\text{coupled}} = (0,\pi,0)$, i.e., the ends are not Majorana fermions any more, but time reversal changes the fermion parity of the ends. The fermion parity for the segment of the combined chain is given by $Q = Q_1Q_2 = (iQ_1^AQ_1^B)(iQ_2^AQ_2^B) = (Q_1^AQ_2^A)(Q_1^BQ_2^B)$, where the Q_n^A, Q_n^B are fermionic parity operators of a chain n with Majorana ends (see Sec. III A). One can measure the parities of the ends separately because $Q^A = -iQ_1^AQ_2^A$ and a similar operator on B are bosonic operators, so $\mu_{\text{coupled}}=0$. On the other hand $\phi_{\text{coupled}}=\pi$ because T anti-commutes with Q^A on account of the factor of i. Furthermore, one finds that $\kappa_{\text{coupled}}=\pi$. (See App. B.)

2. Consider two chains with $(\mu,\phi,\kappa)=(0,\pi,\pi)$. Q_1^A, Q_2^A are bosonic, therefore $Q^A_{\text{coupled}} = Q_1^AQ_2^A$ is bosonic as well, and $\mu_{\text{coupled}}=0$. Time reversal acting on the left edge is represented as $U^A_{\text{coupled}} = U_1^AU_2^A$. It anti-commutes with Q_1^A, Q_2^A, but commutes with their product, therefore $\phi_{\text{coupled}}=0$. Since both U_1^A, U_2^A change the fermion parity, they both have to be *fermionic*. Therefore $\{U_1^A, U_2^A\} = 0$ and we get that

$$\begin{aligned} U^A_{\text{coupled}}U^{A*}_{\text{coupled}} &= \left(U_1^AU_2^A\right)\left(U_1^AU_2^A\right)^* \\ &= -U_1^AU_1^{A*}U_2^AU_2^{A*} \\ &= -\mathbb{1}. \end{aligned} \tag{17}$$

   Hence $\kappa_{\text{coupled}}=\pi$, and the resulting phase is labelled by $(0,0,\pi)$.

3. Combining two chains with $(0,0,\pi)$ finally gives the trivial phase, because all the symmetries are represented by bosonic operators, therefore κ simply doubles to give $0 \bmod (2\pi)$. This conforms with the fact that the Majorana chain is an order eight element of the group.

Working out the addition rule in general gives the table of phases which are summarized in Table I. A concise way to describe the general addition rule is to define $\lambda \equiv \kappa+\phi (\text{mod } 2\pi)$. Then we represent a state by a 3-digit binary number $\left(\frac{\lambda}{\pi},\frac{\phi}{\pi},\frac{\mu}{\pi}\right)$. These numbers add modulo 8 when the phases are combined.

V. TRANSLATIONAL INVARIANCE AND θ

If, in addition to fermion parity conservation, translational invariance is also present, the number of distinct phases is doubled. Below, we derive the associated invariant, θ, which can take the values 0 or π, independent of the invariants (μ,ϕ,κ) described above. The degeneracy of the entanglement spectrum, however, is not modified in either the $\theta=0$ or π phases, and is given by Table I.

Let us consider a fermionic chain with translational invariance. According to the Sec. III A, the fermion parity of a segment $S=[1,L]$ extending from $j=1$ to $j=L$ (where L is much larger than the correlation length ξ) can be written as

$$Q(1,L) = e^{-i\frac{\mu}{2}}f(L)\,Q^A(1)\,Q^B(L). \tag{18}$$

Here, we have kept track explicitly of the position of the operators Q^A and Q^B, and of an overall constant sign

$f(L)$ (which was absorbed into the definition of Q^A and Q^B before). Translational invariance removes the necessity of choosing the sign of Q^B (and Q^A) separately for each segment. This symmetry also allows us to write the parity operator of the segment $[1, L]$ in terms of those of the two segments $S_1 = [1, L']$ and $S_2 = [L'+1, L]$ (where $L', L - L' \gg \xi$) as

$$\begin{aligned} Q(1,L) &= Q(1,L')\,Q(L'+1,L) \\ &= e^{-i\frac{\mu}{2}} f(L')\, f(L-L') \\ &\quad \times Q^A(1)\left[e^{-i\frac{\mu}{2}} Q^B(L')\,Q^A(L'+1)\right] Q^B(L). \end{aligned} \tag{19}$$

Equating Eq. 18 and 19 gives that, within the low–entanglement subspace, we must have

$$f(L')\, f(L-L')\left[e^{-i\frac{\mu}{2}} Q^B(L')\,Q^A(L'+1)\right] = f(L). \tag{20}$$

Eq. 20 can hold for every state in the low-entanglement subspace only if these states are all eigenstates of $e^{-i\frac{\mu}{2}} Q^B(L')\,Q^A(L'+1)$. Q^A and Q^B can be defined in such a way that the corresponding eigenvalue is 1. Then, we get that $f(L')\,f(L-L') = f(L)$, which is solved by

$$f(L) = e^{i\theta L}. \tag{21}$$

From the requirement that $[Q(1,L)]^2 = \mathbb{1}$, we get that θ can only take the values 0 or π. Thus, each of the eight phases found in the previous section is further split into two distinct phases, corresponding to the two allowed values of θ. For example, for the Majorana chain model (Eq. 1), $t = 0$, $u = +1$ and $t = 0$, $u = -1$ describe distinct phases, although both have $\mu = 0$. The ground state has all sites occupied or unoccupied, corresponding to $\theta = \pi$ or $\theta = 0$, respectively.

Note that both θ and μ have a concrete consequence not only for the entanglement spectrum but also for the parity of the ground state in periodic chains. If the length of the chain is much larger than the correlation length, the parity depends only on μ and θ and the chain length, $(-1)^{(\mu+\theta L)/\pi}$. Thus, a phase with $\mu = \pi$ has an odd number of fermions on a chain of an even length.[20] The phase θ determines whether the parity of the ground state alternates as a function of L or not. This is shown in appendix C.

VI. SUMMARY AND OUTLOOK

We have described a systematic procedure for classifying the phases of 1D interacting fermions. We focussed on spinless fermions with time-reversal symmetry and particle number parity conservation. In the non-interacting case, these models are classified by an integer number, i.e., by $\mathbb{Z}$.[9,10] We used concepts of entanglement to classify the phases in the presence of interactions. We derive an effective description of the dominant entanglement states which then allows us to recognize "topological" features based on projective representations of the symmetries. We found, in agreement with the results of Fidkowski *et. al.* (Ref. [1]), that in the presence of interactions there are only eight distinct phases. Each of these eight phases is characterized by a unique set of bulk invariants (μ, ϕ, κ), which can take the values 0 or π. These invariants are related to the transformation laws of the entanglement eigenstates under symmetry operations, and the phases have a characteristic degeneracy of entanglement levels. Furthermore, the phases obey a $\mathbb{Z}_8$ group structure and each of the eight phases can be generated by adding single chains together. All possible phases and the addition rules are summarized in Table I. If translational symmetry is also present, the number of distinct phases increases to 16.

The symmetries we have focused on describe only one of the 10 Altland-Zirnbauer classes[37] of topological insulators. The framework described here can also be used to show how the phases in these other classes are modified by interactions. To analyze each of the classes of topological insulators (and in both interpretations when particle-hole symmetry is present), one only has to determine the appropriate algebra of symmetries and then determine the possible projective representations of this algebra.

Interactions cause the meaning of the Altland-Zirnbauer classes to bifurcate, however. At the mean-field level, a superconductor has an emergent particle-hole symmetry in its band structure. Thus, the classes which have such a symmetry can be interpreted as describing either superconductors or systems that have a true particle-hole symmetry (such as the Hubbard model for fermions with spin on a bipartite lattice at half-filling). When interactions are included, these two interpretations are distinct. Thus, the class BDI, for example, has particle-hole and time reversal symmetry. This can be interpreted as describing superconductors. This means that one symmetry, particle conservation, breaks down, and only fermion parity Q is left. The only two symmetries are T and Q, giving the problem treated here. BDI has an alternative interpretation, according to which it describes systems with a true particle-hole symmetry C that reverses the sign of $\langle n_i \rangle - \bar{n}$. (Here, n_i is the occupation number of a site and $\bar{n}$ is the mean occupation number.) In this case, particle number N is conserved, and besides this there are two other symmetries T and C. These satisfy the algebra $T^2 = C^2 = \mathbb{1}$, $CN + NC = 2\bar{n}L$ where L is the length of the system. (Every other pair of these symmetries commute.) The set of phases is different for the two interpretations; particle number conservation rules out Majorana Fermions. A complete classification of systems in all Altland-Zirnbauer classes in one dimension, following either of the two interpretations of particle-hole symmetry mentioned above, would be an interesting project for future work.

A generalization of these results to higher dimensional systems is an interesting (and challenging) open problem. In some of the symmetry classes of topological insulators and superconductors, strong arguments have

Number of chains	(μ, ϕ, κ)	Degeneracy of segment
1	$(\pi, 0, 0)$	2
2	$(0, \pi, \pi)$	4
3	(π, π, π)	8
4	$(0, 0, \pi)$	4
5	$(\pi, 0, \pi)$	8
6	$(0, \pi, 0)$	4
7	$(\pi, \pi, 0)$	2
8	$(0, 0, 0)$	1

TABLE I: Degeneracies and addition table. All possible phases of fermions are realized by simply taking copies of some number of Majorana chains (see next section); the first column is the number of chains. The next column gives the parameters classifying a given state. The third column gives the degeneracy of the Schmidt spectrum, counting *both* ends.

been given that the non–interacting classification does not change when interactions are included. This is particularly clear when the topological invariant is related to a quantized physical response, *e.g.*, in the integer quantum Hall effect[12,13] and in 3D time–reversal invariant topological insulators[11]. However, for other classes, the situation is less clear. For example, the non–interacting classification of 3D chiral superconductors is $\mathbb{Z}$,[9,10] similar to the one–dimensional case considered here. It would be interesting to consider the effect of interactions on the phase diagrams of such systems.

Note As we were writing this article, we learned that a similar classification is being worked out by Fidkowski and Kitaev.[38] Our results are consistent with theirs.

Acknowledgment

We thank Lukasz Fidkowski and Shinsei Ryu for useful discussions. E. B. was supported by the NSF under grants DMR-0705472 and DMR-0757145. F. P. and A. M. T. acknowledge support from ARO grant W911NF-07-1-0576.

Appendix A: Low Energy Operators and the Ends of the Chain

An intuitive argument, given above, suggests that low-energy operators acting on the entanglement eigenstates may be represented approximately by operators located near the ends of the chain: in each Schmidt state $|\alpha\beta\rangle_S$ the expectation values of the spin and other operators have some particular spatial dependence near the ends of the chain depending on α and β, but this decays exponentially to the ground state away from the ends. Therefore it should be possible to transform between these states by using operators defined on just the ends. A special case is the effective representations of symmetries in terms of operators at the ends of the segments, which we used to define the topological phases.

To see that these effective operators exist, one can use a matrix product state representation[39] of the wave function. (We focus here just on bosonic systems. For systems including fermions a similar argument can be developed using bosonization but the discussion of this general result gets complicated by the classification of the phases by μ.)

A basis for the low-energy operators can be constructed as follows. For each fixed choice of α_1 and α_2 let $\mathcal{O}^A(\alpha_2, \alpha_1)$ be the operator that transforms α_1 into α_2.

We will now give an approximate representation for $\mathcal{O}^A(\alpha_1, \alpha_2)$ that gives the correct transformation of *low energy states* of $\mathcal{H}_S$. Let the matrices Γ_m, Λ define the bulk state (m varies over a basis for the physical Hilbert space). The ground state wavefunction of a ring of length N is given by

$$|\psi\rangle = \sum_{\{m_i\}} \mathrm{tr}(\Gamma_{m_1}\Lambda\Gamma_{m_2}\ldots\Lambda\Gamma_{m_N}\Lambda)|m_1 m_2 \ldots m_N\rangle \tag{A1}$$

Γ_m, Λ can be brought into a canonical form, satisfying $\sum_m \Gamma_m\Lambda\Gamma_m^\dagger = \sum_m \Gamma_m^\dagger\Lambda\Gamma_m = \mathbb{1}$, where Λ is a diagonal matrix with non–negative entries.[40,41] For a generic wave function, Γ_m, Λ are infinite-dimensional.

Let us define the states $|\alpha\beta\rangle_{1L}$ of a segment of the chain stretching from 1 to L, where $L < N$, as

$$|\alpha\beta\rangle_{1,L} = \sum_{\{m_i\}} (\Gamma_{m_1}\Lambda\Gamma_{m_2}\ldots\Lambda\Gamma_{m_L})_{\alpha\beta}|m_1 m_2 \ldots m_L\rangle \tag{A2}$$

When L is large these states are nearly orthonormal, that is $|\langle\alpha'\beta'|\alpha\beta\rangle_{1,L} - \delta_{\alpha'\alpha}\delta_{\beta'\beta}| \sim Ce^{-\frac{L}{\xi}}$, where ξ is the length scale for the decay and C is a constant dependending on the indices. In this limit, $|\alpha\beta\rangle_{1L}$ are the entanglement eigenstates of the segment. On the other hand, if the length of the chain is fixed and $\alpha, \beta, \alpha', \beta'$ increase, the orthonormality must eventually break down for high enough $\alpha, \beta, \alpha', \beta'$ (since the Hilbert space of a segment of length L is finite). Indeed, C grows as a function of $\alpha, \alpha', \beta, \beta'$.

The ground state wavefunction (A1) can now be written as

$$|\psi\rangle = \sum_{\alpha,\beta} \lambda_\alpha\lambda_\beta|\alpha\beta\rangle_{1,L}|\alpha\beta\rangle_{L+1,N}. \tag{A3}$$

This gives the Schmidt decomposition into states of the environment $|\alpha\beta\rangle_{L+1,N}$ and states of the chain $|\alpha\beta\rangle_{1,L}$, with a Schmidt eigenvalue $\lambda_\alpha\lambda_\beta$ [or, equivalently, an entanglement energy $E = E_\alpha + E_\beta = -2(\ln\lambda_\alpha + \ln\lambda_\beta)$]. The Schmidt eigenstates become orthogonal to each other in the limit $N \to \infty$ and $L \to \infty$.

We can now give an "effective" expression for $\mathcal{O}^A$ in terms of local operators acting on sites $1, \ldots, \ell$ (near the left edge of the segment $[1, L]$), valid for a low entanglement-energy subspace with $E_\alpha < E_{\mathrm{cut}}$. E_{cut} is a cutoff which depends on ℓ. The accuracy of our effective

expression improves as ℓ becomes larger (provided that $N \gg \ell$). Define

$$\mathcal{O}^{A,\text{eff}}_{\alpha_2\alpha_1} = \sum_{\gamma=1}^{\chi} |\alpha_2\gamma\rangle_{1,\ell}\langle\alpha_1\gamma|_{1,\ell} \tag{A4}$$

where χ is a cutoff of the entanglement spectrum that satisfies $E_{\gamma=1,\ldots,\chi} < E_{\text{cut}}$. $\mathcal{O}^{A,\text{eff}}_{\alpha_2\alpha_1}$ acts only on the ℓ first sites of the segment $[1, L]$.

We now apply $\mathcal{O}^{A,\text{eff}}_{\alpha_2\alpha_1}$ to the state Eq. (*A*2), with $\alpha \leq \chi$. This state can be expanded $\sum_{\gamma'} |\alpha\gamma'\rangle_{1\ell}\lambda_\gamma|\gamma'\beta\rangle_{\ell+1,N}$. Using the approximate orthornormality of the states on the segment from 1 to ℓ, we find that the operator in fact transforms α_1 into α_2. Intuitively, the sum over γ, the state of the internal end of the $\ell-$site segment ensures that this operator keeps the right type of entanglement between the left and right side of the "cut" at ℓ. The error of (A4) scales as $F(\chi)e^{-\frac{\ell}{\xi}}$, where $F(\chi)$ is a function of χ. ($F(\chi)$ grows with χ, hence, to deal with a larger range of "energies", a larger value of ℓ must be used. This is because, as the "energy" of a state increases, it penetrates further into the bulk.)

Note that $\mathcal{O}^{A,\text{eff}}_{\alpha_2\alpha_1}$ does not transform high–entanglement energy states correctly, that is states such as $|\alpha\beta\rangle_{1N}$ with $\alpha > \chi$. This is because $|\alpha_1\gamma\rangle$ and $|\alpha\gamma'\rangle$ are orthonormal only if they are both low energy states. This limitation is unavoidable: it is not possible to find a perfect representation for an operator, such as Q, in terms of just the ℓ sites near each end. One can add an extra particle somewhere outside of the reach of these sites, changing the value of Q but not of an observable on the ends of the chain. The *physical* energy of this state may not be much greater than the gap. However, not being able to describe states like this is not a problem when one is studying the ground state of the system: its *entanglement* energy is large, which means that it contributes negligibly to the value of any observable in the ground state.

Appendix B: Factoring antiunitary operators

In the analysis of time–reversal symmetry, we defined a parameter κ by factoring T into two operators U^A and U^B, acting near the two opposite edges of the segment. To determine how chains add to one another, it is necessary to know the commutation and anticommutation properties of these operators. We ignored a small detail, however: since T is antiunitary, it cannot be factored either as the product of two unitary or two antiunitary operators. One solution is just to explicitly write how T transforms the basis states as we did in Eqs. (11) and (12). This becomes cumbersome after a while however, and later in the text we have treated U^A and U^B as unitary operators in Hilbert space, without keeping explicit track of their indices. Here we will explain the meaning of this.

We will first discuss the bosonic case. Eq. (11) gives the action of T only on basis states. Taking a superposition gives a factorization of T that is correct for any state:

$$T = U_A U_B K, \tag{B1}$$

where U_A and U_B are *unitary* operators at the two ends and K is defined by

$$K\sum_{\alpha\beta} a_{\alpha\beta}|\alpha\beta\rangle = \sum_{\alpha\beta} a^*_{\alpha\beta}|\alpha\beta\rangle, \tag{B2}$$

where $a_{\alpha\beta}$ are arbitrary coefficients. We can now define κ by $(U^A K)^2 = e^{i\kappa}\mathbb{1}$. Thus, $U^A K$ is an antiunitary symmetry squaring to $-\mathbb{1}$ in the nontrivial phase as in Kramers' theorem. Note that this equation is equivalent to the definition given above, Eq. (13). This is because $KU^A K^{-1} = U^{A*}$ when the matrices are represented in the basis $|\alpha\beta\rangle$. (Note that the complex conjugate of a matrix depends on the basis being used, unlike the adjoint.)

We can argue physically that operators U^A and U^B satisfying Eq. (B1) can always be found. Consider the ratio TK^{-1} between T, which is represented by complex conjugation in terms of the microscopic degrees of freedom, and K which describes complex conjugation in the entanglement eigenstate basis. This operator is unitary. Furthermore it acts independently on the two ends: one may check that $K\mathcal{O}^{A,B}K$ is an operator acting on end A or B respectively, by expressing it in the basis of entanglement eigenstates.

The operator $U^A K$ used to define κ is non–local. It does not commute with operators at end B, because it takes complex conjugates of them. However, we can still argue that κ is well-defined: Square Eq. (B1): $\mathbb{1} = U^A U^B K U^A U^B K$. Since $K^2 = \mathbb{1}$, we can write this also as $\mathbb{1} = U^A U^B (KU^A K)(KU^B K)$. $KU^A K$ is an operator which acts on end A, therefore it commutes with U^B. Hence $\mathbb{1} = [U^A(KU^A K)][U^B(KU^B K)]$. Since the two factors are local, each must be a pure phase, hence $(U^A K)^2 = e^{i\kappa}\mathbb{1}$.

Now the operators U^A and U^B are not uniquely defined because *complex conjugation*, K, *is basis dependent*. Changing the basis of eigenstates in which Eq. (B2) is imposed (e.g., multiplying the entanglement states by phase factors) changes K. This does not change topological properties like the value of κ however: the unitary transformation that changes the basis can be carried out continuously, starting from the identity. In this process, κ cannot change because it can only be 0 or π.

For fermionic systems with $\mu = 0$, one can decompose T using Eq. (B1). When $\mu = \pi$, the situation is more complicated because the parity eigenvalue q cannot be associated with either one of the edges. We have to make sure that K still maps operators at each end of the system to other operators at that end. In particular, $KQ^A K$ must be a local operator at end A.

This condition is satisfied if K is defined to be complex conjugation in the basis $|\alpha\beta q\rangle$ *provided* that Q^A

and Q^B are represented by either purely real or purely imaginary matrices in that basis. One way to satisfy this requirement is to first choose a basis for $q = +1$ and then to construct the states in the sector from them, $|\alpha\beta, q = -1\rangle = Q^A|\alpha\beta, q = +1\rangle$. Then Q^A is represented by σ^x, acting in the $q = \pm 1$ basis. In this basis, each state is an eigenstate of $Q = \sigma^z$. Last, $Q^B = iQQ^A = \sigma^y$. Since Q^A is real and Q^B is imaginary, the two ends are not mixed by applying K. (If the relative phases of the basis states are changed, then simple complex conjugation would mix Q^A and Q^B into one another.) We have taken the convention that Q^A is real and Q^B is imaginary above.

Now let us show how to calculate κ when two $(\mu, \phi, \kappa) = (\pi, 0, 0)$ chains are combined. The factorization $T = U^A U^B K$ must be carried out in a basis of states of the form $|\alpha\beta\rangle$, according to our conventions. One basis for the states on the two chains together is given by $\{|\pm\rangle_1|\pm\rangle_2\}$ (where the sign represents the values of q_1, q_2. (We do not explicitly write the bosonic indices α, β.). These states map to themselves under time reversal. However, U^A and U^B cannot be the identity because we know they must be fermionic; this is the wrong basis for defining K by simple complex conjugation.

Let us transform the states to a basis in which there is no entanglement between the ends; we therefore use states that are eigenvectors of the local operators $Q^A = -iQ_1^A Q_2^A$ and $Q^B = iQ_1^B Q_2^B$, namely $|q_A q_B\rangle$. (The relative minus sign between Q^A and Q^B ensures that the total parity is $q_A q_B = q_1 q_2$.)

To construct the basis, first find an eigenfunction of Q^A and Q^B with eigenvalues $+1$, (choose the phase arbitrarily):

$$|+_A +_B\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1|+\rangle_2 - i|-\rangle_1|-\rangle_2) \tag{B3}$$

Now generate the other basis states from this by applying Q_1^A and Q_1^B:

$$\begin{aligned}
|-_A +_B\rangle &= Q_1^A|+_A +_B\rangle = \frac{1}{\sqrt{2}}(|-\rangle_1|+\rangle_2 - i|+\rangle_1|-\rangle_2) \\
|+_A -_B\rangle &= -iQ_1^B|+_A +_B\rangle = \frac{1}{\sqrt{2}}(|-\rangle_1|+\rangle_2 + i|+\rangle_1|-\rangle_2) \\
|-_A -_B\rangle &= Q_1^A|+_A -_B\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1|+\rangle_2 + i|-\rangle_1|-\rangle_2)
\end{aligned} \tag{B4}$$

The phases are just conventions in the first two definitions, and the phase in the third equation follows from the independence of the ends: Q_1^A has to act on q_A the same way no matter what the value of q_B is. Now it is clear that T switches the fermion parity of each end in this basis, since changing the sign of i exchanges the states $|q_A q_B\rangle$ and $|-q_A, -q_B\rangle$.

Now we can define K to map each of *these* basis states to itself. Clearly, $T = Q_2$ because Q_2 also exchanges the same pairs of wave functions, or more precisely $T = Q_2 K$. Hence $U^A = -iQ^A$ and $U^B = Q^B$. One can check that $(U^A K)^2 = -\mathbb{1}$, so $\kappa = \pi$.

Note that, in spite of all this trouble, the value of κ in a phase with $\mu = 0, \phi = \pi$ does not have any physical significance–the four-fold degeneracy of the spectrum is already explained by the fact that $\phi = \pi$. The reason T changes fermion parity at the ends is that the two ends can only be disentangled by a change of basis that includes complex phases.

Appendix C: Parity of the ground state on a periodic chain

The parity of the ground state on a periodic chain is given by $e^{i(\theta L + \mu)}$. This follows from a fact in Sec. V. Consider two subsegments of the chain, one ending at X and the other starting at $X + 1$. The ground state wave function is an eigenfunction of the following:

$$e^{-i\frac{\mu}{2}} Q^B(X) Q^A(X+1)|\psi\rangle = |\psi\rangle. \tag{C1}$$

When $\mu = \pi$, this relation describes the correlations between the Majorana degrees of freedom in adjacent segments of the chain.

We now suppose the periodic chain has length L and break it at two places, between L' and $L' + 1$ and between L and 1. The total fermion parity of the ground state is the product of the parity on the two segments, $(e^{i(\theta L' - \frac{\mu}{2})} Q^A(1) Q^B(L'))(e^{i(\theta(L-L') - \frac{\mu}{2})} Q^A(L' + 1) Q^B(L))$. Rearranging and using Eq. (C1), the ground state parity comes out as $e^{i(\theta L + \mu)}$. The extra minus sign when $\mu = \pi$ comes from anticommuting the Q operators.

[1] L. Fidkowski and A. Kitaev, Phys. Rev. B **81**, 134509 (2010).
[2] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[3] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Science **314**, 1757 (2006).
[4] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
[5] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
[6] M. Konig, S. Wiedmann, C. Brune, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[7] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008), 0902.1356.
[8] M. M. Salomaa and G. E. Volovik, Phys. Rev. B **37**, 9298 (1988).
[9] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Phys. Rev. B **78**, 195125 (2008).

[10] A. Kitaev, arXiv:0901.2686 (2009).
[11] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, Phys. Rev. B **78**, 195424 (2008).
[12] Q. Niu and D. J. Thouless, Journal of Physics A: Mathematical and General **17**, 2453 (1984).
[13] J. E. Avron and R. Seiler, Phys. Rev. Lett. **54**, 259 (1985).
[14] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).
[15] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).
[16] H. Li and F. D. M. Haldane, Phys. Rev. Lett. **101**, 010504 (2008).
[17] Z. Nussinov and G. Ortiz, Proceedings of the National Academy of Sciences **106**, 16944 (2009).
[18] F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, Phys. Rev. B **81**, 064439 (2010).
[19] X. Chen, Z. Gu, and X. Wen, arXiv:1008.3745 (2010).
[20] A. Kitaev, Physics-Uspekhi **44**, 131 (2001).
[21] R. Thomale, D. P. Arovas, and B. A. Bernevig, arXiv:0912.0028 (2009).
[22] L. Fidkowski, Phys. Rev. Lett. **104**, 130502 (2010).
[23] A. M. Läuchli, E. J. Bergholtz, J. Suorsa, and M. Haque, Phys. Rev. Lett. **104**, 156404 (2010).
[24] T. J. Osborne, J. Eisert, and F. Verstraete (2010).
[25] M. Hastings and T. Koma, Communications in Mathematical Physics **265**, 781 (2006).
[26] D. Pérez-García, M. Wolf, M. Sanz, F. Verstraete, and J. Cirac, Phys. Rev. Lett. **100**, 167202 (2008).
[27] F. D. M. Haldane, Phys. Lett. **93A**, 464 (1983).
[28] T. Hirano, H. Katsura, and Y. Hatsugai, Phys. Rev. B **77**, 094431 (pages 5) (2008).
[29] E. Berg, E. G. Dalla Torre, T. Giamarchi, and E. Altman, Phys. Rev. B **77**, 245119 (2008).
[30] Z.-C. Gu and X.-G. Wen, Phys. Rev. B **80**, 155131 (2009).
[31] F. Pollmann, E. Berg, A. M. Turner, and M. Oshikawa, arXiv:0909.4059 (2009).
[32] I. Peschel, J. Stat. Mech. p. P06004 (2004).
[33] S.-A. Cheong and C. L. Henley, Phys. Rev. B **69** (2004).
[34] A. M. Turner, F. Pollman, and E. Berg, Work in in progress (2010).
[35] X.-L. Qi, T. L. Hughes, S. Raghu, and S.-C. Zhang, Phys. Rev. Lett **102**, 187001 (2009).
[36] S. Todo, M. Matsumoto, C. Yasuda, and H. Takayama, Phys. Rev. B **64**, 224412 (2001).
[37] A. Altland and M. R. Zirnbauer, Phys. Rev. B **55**, 1142 (1997).
[38] L. Fidkowski and A. Kitaev, arXiv:1008.4138 (2010).
[39] M. Fannes, B. Nachtergaele, and R. W. Werner, Commun. Math. Phys. **144**, 443 (1992).
[40] G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003).
[41] R. Orús and G. Vidal, Phys. Rev. B **78**, 155117 (pages 11) (2008).
[42] Note that it is not possible to make a fermionic operator out of an even number of fermionic factors. Why then is it possible to make half-integer spins out of integer spins in a spin chain? The point is that operators cannot transform in a "fractional way." In fact, the spin operators continue to have the ordinary vector symmetry: they *change* the spin by integer values, for example. It is the eigen*states* and not the operators that have half-integer values. their symmetry properties; only states transform strangely.
[43] In a chain with $\mu = \pi$, operators on both sides of the chain can change q_S. The distinction between operators acting on the left and those acting on the right end is that they are proportional to σ_x, σ_y respectively. Though these act on the same variable q_S, they anticommute.
[44] Note that Ref. 18 discusses the entanglement spectrum at a single cut of an infinite system, and therefore the degeneracy is two–fold. Here, we are discussing a finite but large segment. There is a double degeneracy associated with each edge of the segment, so overall, the entanglement spectrum is *four*–fold degenerate.

Area laws for the entanglement entropy – a review

J. Eisert,[1,2,3] M. Cramer,[3,4] and M.B. Plenio[3,4]

[1] *Institute of Physics and Astronomy, University of Potsdam, 14469 Potsdam, Germany*

[2] *Institute for Advanced Study Berlin, 14193 Berlin, Germany*

[3] *Institute for Mathematical Sciences, Imperial College London, London SW7 2PG, UK*

[4] *Institut für Theoretische Physik, University of Ulm, 89069 Ulm, Germany*

Physical interactions in quantum many-body systems are typically local: Individual constituents interact mainly with their few nearest neighbors. This locality of interactions is inherited by a decay of correlation functions, but also reflected by scaling laws of a quite profound quantity: The entanglement entropy of ground states. This entropy of the reduced state of a subregion often merely grows like the boundary area of the subregion, and not like its volume, in sharp contrast with an expected extensive behavior. Such "area laws" for the entanglement entropy and related quantities have received considerable attention in recent years. They emerge in several seemingly unrelated fields, in the context of black hole physics, quantum information science, and quantum many-body physics where they have important implications on the numerical simulation of lattice models.

In this Colloquium we review the current status of area laws in these fields. Center stage is taken by rigorous results on lattice models in one and higher spatial dimensions. The differences and similarities between bosonic and fermionic models are stressed, area laws are related to the velocity of information propagation in quantum lattice models, and disordered systems, non-equilibrium situations, and topological entanglement entropies are discussed. These questions are considered in classical and quantum systems, in their ground and thermal states, for a variety of correlation measures. A significant proportion of the article is devoted to the clear and quantitative connection between the entanglement content of states and the possibility of their efficient numerical simulation. We discuss matrix-product states, higher-dimensional analogues, and variational sets from entanglement renormalization and conclude by highlighting the implications of area laws on quantifying the effective degrees of freedom that need to be considered in simulations of quantum states.

Contents

I. INTRODUCTION

In classical physics concepts of entropy quantify the extent to which we are uncertain about the exact state of a physical system at hand or, in other words, the amount of information that is lacking to identify the microstate of a system from all possibilities compatible with the macrostate of the system. If we are not quite sure what microstate of a system to expect, notions of entropy will reflect this lack of knowledge. Randomness, after all, is always and necessarily related to ignorance about the state.

In quantum mechanics positive entropies may arise even without an objective lack of information. To see this, let us consider a quantum lattice systems (see e.g., Fig. 1) as an example for a quantum many-body system where each of the vertices i of the lattice L is associated with an individual quantum system. This quantum many-body system is thought to be in its non-degenerate pure ground state $\rho = |\psi\rangle\langle\psi|$ at zero temperature which has vanishing *von-Neumann entropy*

$$S(\rho) = -\mathrm{tr}[\rho \log_2 \rho].$$

Let us now distinguish a region of this quantum lattice system, denoting its sites with the set I and all other sites with $O = L\backslash I$. If we consider the reduced state $\rho_I = \mathrm{tr}_O[\rho]$ of the sites

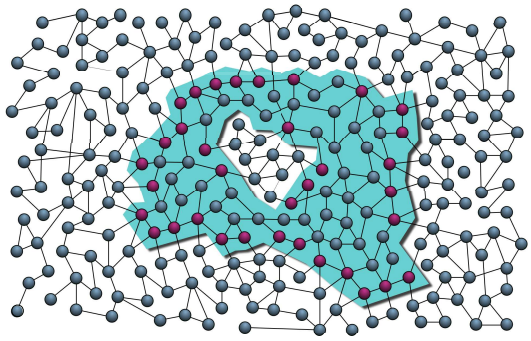


FIG. 1 A lattice L with a distinguished set $I \subset L$ (shaded area). Vertices depict the boundary ∂I of I with surface area $s(I) = |\partial I|$.

of the region I, the state will not be pure in general and will have a non-vanishing von-Neumann entropy $S(\rho_I)$. [1]

In contrast to thermal states this entropy does not originate from a lack of knowledge about the microstate of the system. Even at zero temperature we will encounter a non-zero entropy! This entropy arises because of a very fundamental property of quantum mechanics: Entanglement. This quite intriguing trait of quantum mechanics gives rise to correlations even in situations where the randomness cannot be traced back to a mere lack of knowledge. The mentioned quantity, the entropy of a subregion is called *entanglement entropy* or *geometric entropy* and, in quantum information, *entropy of entanglement*, which represents an operationally defined entanglement measure for pure states (for recent reviews see refs.[125,186]).

In the context of quantum field theory, questions of scaling of entanglement entropies in the size of I have some tradition. Seminal work on the geometric entropy of the free Klein-Gordon field[23,207] and subsequent work on conformal field theories[36,43,118,122,226] was driven in part by the intriguing suggested connection to the Bekenstein-Hawking black hole entropy[17,18,117].

In recent years, studies of properties of the entanglement entropy in this sense have enjoyed a revival initiated in refs.[7,171,172,223]. Importantly, this renewed activity is benefitting from the new perspectives and ideas of quantum information theory, and from the realisation of their significance for the understanding of numerical methods and especially their efficiency for describing quantum many-body physics. Quantum information theory also provides novel conceptual and mathematical techniques for determining properties of the geometric entropy analytically.

At the heart of these studies are questions like: What role do genuine quantum correlations—entanglement—play in quantum many-body systems? Typically, in such investigations, one abstracts to a large extent from the microscopic specifics of the system: Quite in the spirit of studies of *critical phenomena*, one thinks less of very detailed properties, but is rather interested in the *scaling* of the entanglement entropy when the distinguished region grows in size. In fact, for quantum chains, this scaling of entanglement as genuine quantum correlations—a priori very different from the scaling of two-point correlation functions—reflects to a large extent the critical behavior of the quantum many-body system, and shares some relationship to conformal charges.

At first sight one might be tempted to think that the entropy of a distinguished region I, will always possess an extensive character. Such a behavior is referred to as a *volume scaling* and is observed for thermal states. Intriguingly, for typical ground states, however, this is not at all what one encounters: Instead, one typically finds an *area law*, or an area law with a small (often logarithmic) correction: This means that if one distinguishes a region, the scaling of the entropy is merely linear in the *boundary area* of the region. The entanglement entropy is then said to fulfill an *area law*. It is the purpose of this article to review studies on area laws and the scaling of the entanglement entropy in a non-technical manner.

The main four motivations to approach this question (known to the authors) are as follows:

- **The holographic principle and black hole entropy:** The historical motivation to study the entanglement or geometric entropy stems from considerations of black hole physics: It has been suggested in the seminal work of refs.[23,207] that the area law of the geometric entropy for a discrete version of a massless free scalar field—then numerically found for an imaginary sphere in a radial symmetry—could be related to the physics of *black holes*,[118] in particular the Bekenstein-Hawking entropy of a black hole which is proportional to its boundary surface. It has been muted that the *holographic principle*[29]—the conjecture that the information contained in a volume of space can be represented by a theory which lives in the boundary of that region—could be related to the area law behavior of the entanglement entropy in microscopic theories.

- **Distribution of quantum correlations in quantum many-body systems:** Area laws also say something quite profound on how quantum correlations are distributed in ground states of local quantum many-body systems. Interactions in quantum many-body systems are typically local, which means that systems interact only over a short distance with a finite number of neighbors. The emergence of an area law then provides support for the intuition that short ranged interactions require that quantum correlations between a distinguished region and its exterior are established via its boundary surface. That a strict area law emerges is by no means obvious from the decay of two-point correlators, as we will see. Quantum phase transitions are governed by quantum fluctuations at zero temperature, so it is more than plausible to observe signatures of *criticality* on the level of entanglement and *quantum correlations*. This situation is now particularly clear in one-dimensional

[1] Of interest are also other entropies, such as the *Renyi entropies*, $S_\alpha(\rho) = (1-\alpha)^{-1} \log_2 \mathrm{tr}[\rho^\alpha]$ with $\alpha \geq 0$. For $\alpha \searrow 1$ the usual von-Neumann entropy is recovered. In particular in the context of simulatability, Renyi entropies for arbitrary α play an important role.

systems,[4,7,13,36,38,44,46,69,72,73,83,85,92,111,132,134,138,146,147,223,225] but progress has also been made in higher-dimensional systems,[31,55,57,86,90,120,140,184,191,219,238] with rigorous area laws specifically for quasi-free bosonic[55,57,184] and fermionic[56,86,96,238] systems, as well as in disordered systems[190].

- **Complexity of quantum many-body systems and their simulation:** One of the key motivations for studying area laws stems from a quite practical context: The numerical simulation of quantum many-body systems. In fact, if there is little entanglement in a ground state of a many-body systems, one might suspect on intuitive grounds that one can describe this ground state with relatively few parameters. More specifically, for one-dimensional systems, one would expect numerical algorithms like the powerful *density-matrix renormalization group method*[199,232] (DMRG) to perform well if the ground state contains a small amount of entanglement. This suspicion can in fact be made rigorous[111,176,203,215] as it turns out that the scaling of entanglement specifies how well a given state can be approximated by a matrix-product state[84,199] as generated in DMRG. It is hence not the decay behavior of correlation functions as such that matters here, but in fact the scaling of entanglement.

- **Topological entanglement entropy:** The topological entanglement entropy is an indicator of *topological order*,[165,228,236] a new kind of order in quantum many-body systems that cannot be described by local order parameters[108,140,148,165,175]. Lattice models having a non-vanishing topological entanglement entropy may be seen as lattice instances of topological quantum field theories. Here a global feature is detected by means of the scaling of geometric entropies.

In this Colloquium we do not have sufficient space to give an account of all known derivations of area laws for the entanglement entropy. However, we will try not to merely remain at a superficial level and only state results, but will explain a number of key techniques and arguments. When we label main statements as "theorems" this is done to highlight their special role, to make it easier to follow the line of reasoning. For details of arguments and proofs, often technically involved, we refer the reader to the original work. The reason for the technicality of proofs originates from the type of question that is posed: To distinguish a region of a lattice breaks the translational symmetry of the problem – even in a translationally invariant setting. While numerical studies are sometimes easier to come by, analytical argument can be technically involved, even for quasi-free models. In this article, we discuss the study of entanglement entropy primarily (i) from the viewpoint of quantum information theory, (ii) with an emphasis on rigorous and analytical results, and (iii) the implications on the efficiency of numerical simulation.

II. LOCAL HAMILTONIANS AND AREA LAWS

Throughout this article, we will consider quantum many-body systems on a lattice. Such quantum lattice systems are ubiquitous in the condensed matter context[227] where they play a key role in obtaining an understanding of material properties from a microscopic basis. Lattices systems are also of considerable importance in the study of quantum field theories where a lattice provides a natural ultra-violet cut-off and facilitates numerical simulations of quantum fields[159]. One could think, e.g., of systems of strongly correlated electron systems or lattice vibrations of a crystal lattice. With the advent of research on *cold atoms* in optical lattices, quantum lattice systems can also be prepared in laboratory conditions with an unprecedented degree of control[22].

We will consider—at least in parts of this article—general lattices. Each vertex of the *lattice* is associated with a quantum system, such as a spin, a bosonic or a fermionic system. It is convenient to think of this lattice as a simple graph $G = (L, E)$ with vertices L, and the edge set E labeling neighborhood relations. G could be the graph representing a one-dimensional chain with periodic boundary conditions, and in fact a good proportion of this article will deal with such quantum chains. For later purposes, it will be convenient to think in terms of such a slightly more general picture, however. The Hilbert space of the total many-body system is then the tensor product

$$\mathcal{H} = \bigotimes_{j \in L} \mathcal{H}_j$$

where $\mathcal{H}_j$ is the Hilbert space associated with the physical system on lattice site j. On such a lattice, one has $\mathrm{dist}(j,k)$ for $j, k \in L$ as the natural *graph theoretical distance* which is the length of the shortest path connecting j and k. For a *cubic lattice* of dimension $\mathcal{D}$ with periodic boundary conditions, in turn, $\mathrm{dist}(j,k) = \sum_{d=1}^{\mathcal{D}} |j_d - k_d|$, where the components of $j, k \in L$ are taken modulo the base length of the cubic lattice.

We will be concerned largely with *local* Hamiltonians on lattices. This means that the physical system associated with a specific lattice site will interact only with its neighbors and *not* with all sites of the lattice. The total Hamiltonian can hence be written as

$$H = \sum_{X \subset L} H_X,$$

where H_X has a compact support X, independent of the system size, that is the number of lattice sites denoted by $|L|$.

The *boundary surface area* $s(I)$ of a distinguished region I of the lattice L can be defined in a very natural fashion on such a graph as the cardinality of the set of boundary points

$$\partial I = \{i \in I : \text{ there is a } j \in L \backslash I \text{ with } dist(i,j) = 1\}, \quad (1)$$

so $s(I) = |\partial I|$, see Fig. 1. Throughout the article, unless defined specifically otherwise, we will say that the entanglement entropy satisfies an *area law* if

$$S(\rho_I) = O(s(I)).$$

This means that the entropy of the reduced state ρ_I scales at most as the boundary area of the region I.

Before we dive into the details of known results on area laws in quantum many-body systems, let us appreciate how unusual it is for a quantum state to satisfy an area law. In fact, a quantum state picked at random will exhibit a very different scaling behavior. If one has a lattice system with d-dimensional constituents and divides it into a subsystem $I \subset L$ and the complement $O = L\backslash I$, then one may consider the expected entanglement entropy of I for the natural choice, the unitarily invariant Haar measure. One finds[89,174,206]

$$\mathbb{E}[S(\rho_I)] > |I| \log_2(d) - \frac{d^{|I|-|O|}}{2\log_2(2)}.$$

That is, asymptotically, the typical entropy of a subsystem is almost maximal, and hence linear in the number of constituents $|I|$. Hence a "typical" quantum state will asymptotically satisfy a volume law, and not an area law. As we will see that area laws are common for ground states of quantum many-body systems, we find that in this sense, ground states are very non-generic. This fact is heavily exploited in numerical approaches to study ground states of strongly correlated many-body systems: One does not have to vary over all quantum states in variational approaches, but merely over a much smaller set of states that are good candidates of approximating ground states of local Hamiltonians well, that is states that satisfy an area law.

III. ONE-DIMENSIONAL SYSTEMS

Most known results on area laws refer to one-dimensional chains such as *harmonic* or *spin chains*. This emphasis is no surprise: After all, a number of physical ideas—like the Jordan-Wigner transformation—as well as mathematical methods—such as the theory of Toeplitz determinants and Fisher-Hartwig techniques—are specifically tailored to one-dimensional translationally invariant systems.

If we distinguish a contiguous set of quantum systems of a chain, a *block* $I = \{1, \ldots, n\}$ the boundary of the block consists of only one (two) site(s) for open (periodic) boundary conditions. An area law then clearly means that the entropy is upper bounded by a constant independent of the block size n and the lattice size $|L|$, i.e.,

$$S(\rho_I) = O(1). \tag{2}$$

We will see that in quantum chains, a very clear picture emerges concerning the scaling of the entanglement entropy. Whether an area law holds or not, will largely depend on whether the system is at a *quantum critical point* or not. We will summarize what is known in one-dimensional systems at the end of the detailed discussion of quantum chains, starting with bosonic harmonic chains.

A. Bosonic harmonic chain

Bosonic harmonic quantum systems, as well as fermionic models and their quantum spin chain counterparts like the XY model, play a seminal role in the study of quantum many-body systems. Harmonic lattice systems model discrete versions of Klein-Gordon fields, vibrational modes of crystal lattices or of trapped ions and serve generally as lowest order approximations to anharmonic systems. The fact that they are integrable renders even sophisticated questions like the scaling of the geometric entropy in instances amenable to fully analytical study, even in higher spatial dimensions. In fact, in the latter case these so-called quasi-free models are the only settings that allow for rigorous analytical results so far. Hence, they do form the central object of consideration to explore what should be expected concerning general scaling laws.

The Hamiltonian for a harmonic lattice L is given by

$$H = \frac{1}{2} \sum_{i,j \in L} \big(p_i P_{i,j} p_j + x_i X_{i,j} x_j \big), \tag{3}$$

where $X, P \in \mathbb{R}^{|L|\times|L|}$ are real, symmetric and positive matrices determining the coupling structure of the systems. The canonical operators x_i, p_i satisfy the canonical commutation relations $[x_j, p_k] = \mathrm{i}\delta_{j,k}$. In terms of the bosonic annihilation operators $b_j = (x_j + \mathrm{i}p_j)/\sqrt{2}$ the Hamiltonian eq. (3) reads

$$H = \frac{1}{2} \sum_{i,j} \big(b_i^\dagger A_{i,j} b_j + b_i A_{i,j} b_j^\dagger + b_i B_{i,j} b_j + b_i^\dagger B_{i,j} b_j^\dagger \big), \tag{4}$$

where $A = (X + P)/2$, $B = (X - P)/2$. Ground and thermal states of the above Hamiltonian are fully characterized by the second moments of the canonical operators, while first moments vanish[75] (entanglement properties of the state are invariant under changes of first moments anyway). The second moments define the *covariance matrix*

$$\Gamma_{i,j} = \langle \{ r_i, r_j \} \rangle = \langle r_i r_j \rangle + \langle r_j r_i \rangle, \tag{5}$$

where $r = (x_1, \ldots, x_{|L|}, p_1, \ldots, p_{|L|})$ is the vector of canonical operators. The covariance matrix of the ground state of eq. (3) is given by $\Gamma = \Gamma_x \oplus \Gamma_p$, where

$$\Gamma_p = X^{1/2} \big(X^{1/2} P X^{1/2} \big)^{-1/2} X^{1/2}$$

and $\Gamma_x = \Gamma_p^{-1}$, see refs.[55,201]. On the level of covariance matrices unitary operations express themselves as symplectic transformations S that preserve the commutation relations $\sigma_{k,l} = \mathrm{i}[r_k, r_l]$, i.e., $S\sigma S^T = \sigma$. Importantly, *Williamson's Theorem* states that for any strictly positive matrix $A \in \mathbb{R}^{2N\times 2N}$ there exist a symplectic transformation S such that $SAS^T = D$, where D is a diagonal matrix with the same spectrum as the positive square roots of $(\mathrm{i}\sigma A)^2$. The eigenvalues d_i of D are called the *symplectic eigenvalues* of A.

Now, what is the entanglement content of the ground state? To answer this we need to define entanglement measures and compute them in terms of the properties of the covariance matrix. The first of these is of course the entropy of entanglement. Williamson's theorem shows that any function of a state that is unitarily invariant is fully determined by the symplectic eigenvalues. Notably, the entropy of a Gaussian state ρ with

symplectic eigenvalues $d_1, \ldots, d_N$ of the covariance matrix of ρ is given by

$$S(\rho) = \sum_{j=1}^{N} \left(\frac{d_j + 1}{2} \log_2 \frac{d_j + 1}{2} - \frac{d_j - 1}{2} \log_2 \frac{d_j - 1}{2} \right).$$

A key ingredient in the analytical work is another full entanglement measure that was defined in quantum information theory, the *logarithmic negativity*[70,74,182,186,224,246]. It is defined as

$$E_N(\rho, I) = \log_2 \|\rho^{\Gamma_I}\|_1,$$

where $\|A\|_1 = \mathrm{tr}[(A^\dagger A)^{1/2}]$ is the trace norm and ρ^{Γ_I} is the *partial transpose* of ρ with respect to the interior I. The partial transpose w.r.t. the second subsystem is defined as $(|i\rangle\langle k| \otimes |j\rangle\langle l|)^{\Gamma_2} = |i\rangle\langle k| \otimes |l\rangle\langle j|$ and extended by linearity. On the level of covariance matrices the partial transpose is partial time reversal, i.e. $p_i \mapsto -p_i$ if $i \in I$ while x_i remains invariant. Then for ρ with covariance matrix $\Gamma = \Gamma_x \oplus \Gamma_p$ we find that ρ^{Γ_I} has covariance matrix $\Gamma' = \Gamma_x \oplus (F\Gamma_p F)$, where the diagonal matrix F has entries $F_{i,j} = \pm\delta_{i,j}$, depending on whether a coordinate is in I or O: Then one finds for a state with covariance matrix $\Gamma = \Gamma_x \oplus \Gamma_p$ the logarithmic negativity[7,53]

$$E_N(\rho, I) = \frac{1}{2} \sum_{k=1}^{|L|} \log_2 \max\{1, \lambda_k(\Gamma_p^{-1} F \Gamma_x^{-1} F)\},$$

where the $\{\lambda_k\}$ denote the eigenvalues. The logarithmic negativity has two key features. Mathematically, the importance of $E_N(\rho, I)$ is due to

$$E_N(\rho, I) \geq S(\rho_I) \tag{6}$$

which holds for all pure states ρ. This *upper bound* for the entanglement entropy is simpler to compute as one does not have to look at spectra of reductions ρ_I but of the full system. This renders a study of area laws possible even in higher dimensional systems. Secondly, in contrast to the entropy of entanglement, the negativity is also an entanglement measure for mixed states, such as thermal states and provides an upper bound on other important measures of mixed state entanglement[20,49,186,212].

All of the above holds for general lattices L but for the moment we will focus on the one-dimensional setting, that is $L = \{1, \ldots, N\}$ where N is even to allow us to consider the *symmetrically bisected chain* $I = \{1, \ldots, N/2\}$ with periodic boundary conditions and $P = \mathbb{1}$. We concentrate on the ground state and discuss thermal states later.[2] It is worth noting that in higher spatial dimension the natural analog of this setting, the half-space, is of some importance as it allows for a reduction of the problem in question to the 1-D case discussed here[56]. Furthermore, the scaling behavior of the entanglement of the half-chain has direct consequences on the availability of efficient representations of the state by means of matrix-product states as will be discussed in some detail later on in this article. For a general nearest-neighbor coupling this means that X is the circulant matrix,

$$X = \mathrm{circ}(a, b, 0, \ldots, 0, b), \tag{7}$$

as a consequence of translational invariance. b specifies the coupling strength, a defines the on-site term, $\lambda_{\min}(X) = a - 2|b|$, i.e., positivity demands $a > 2|b|$, and the energy gap above the ground state is given by $\Delta E = \lambda_{\min}^{1/2}(XP) = (a - 2|b|)^{1/2}$. For the logarithmic negativity of the symmetrically bisected half-chain we find[7]:

Theorem 1 (Exact negativity of the half-chain) *Consider a Hamiltonian of a harmonic chain on $L = \{1, \ldots, N\}$ with periodic boundary conditions, $P = \mathbb{1}$, and nearest-neighbor interactions as in eq. (7). Then the entanglement entropy of the symmetrically bisected chain and the logarithmic negativity satisfy*

$$S(\rho_I) \leq E_N(\rho, I) = \frac{1}{4} \log_2 \left(\frac{a + 2|b|}{a - 2|b|} \right) = \frac{1}{2} \log_2 \left(\frac{\|X\|^{1/2}}{\Delta E} \right) \tag{8}$$

where $\|\cdot\|$ is the operator norm and $\Delta E = \lambda_{min}^{1/2}(X)$.

The quantity $\|X\|$ will later be related to the speed of sound in the system. This expression for the block entanglement quantified with respect to the negativity is exact and no approximation. This was to the knowledge of the authors a first rigorous area law for a lattice system, complementing earlier seminal work for fields[43]. Remarkably, this expression is entirely independent of N, the system size. The most important observation here is that an area law holds, which can be expressed in terms of the spectral gap in the system: Whenever the system is non-critical in the sense that the energy gap ΔE satisfies $\Delta E \geq c > 0$ with a system size independent constant c, a one-dimensional area law will hold. The above link of entanglement entropy and spectral gap in the system can be established in much more generality and we will delay this discussion to a later subsection.

The argument leading to Theorem 1 is involved, and for details we refer to ref.[7]. The interesting aspect of this proof is that the spectrum of the half chain can not be obtained analytically, thus not allowing for a direct computation of the entanglement content. Instead, it is the particular combination of spectral values of the partial transpose entering $E_N(\rho, I)$ itself that can be explicitly computed. The proof makes heavy use of the symmetry of the problem, namely the invariance under a flip of the two half chains.

This result suggested that the locality of the interaction in the gapped model is inherited by the locality of entanglement, a picture that was also later confirmed in more generality. Note that the above bound is a particularly tight one, and that it may well suggest what prefactor in terms of the energy gap and speed of sound one might expect in general area laws, as we will discuss later.

[2] We do not discuss the entanglement properties in excited states here as this area has not been explored in detail so far[60,208].

Let us now consider an important model for which the energy gap vanishes in the thermodynamical limit $N \to \infty$: Taking $a = m^2 + 2N^2$, $b = -N^2$, identifying lattice sites by $i = xN$, and the canonical operators by $x_i = N^{-1/2}\phi(x)$, $p_i = N^{-1/2}\pi(x)$, one obtains the Klein-Gordon field Hamiltonian

$$H = \frac{1}{2}\int_0^1 \mathrm{d}x \left(\pi^2(x) + \left(\frac{\partial}{\partial x}\phi(x)\right)^2 + m^2\phi^2(x)\right), \quad (9)$$

in the field limit $N \to \infty$. (For a detailed discussion of the continuum limit for the Klein-Gordon field, see also ref.[25].) From the expression (8) for the entanglement, we immediately obtain

$$E_N(\rho, I) = \frac{1}{4}\log_2\left(1 + \frac{4N^2}{m^2}\right) \to_{N\to\infty} \frac{1}{2}\log_2\left(\frac{2N}{m}\right). \quad (10)$$

This is a striking difference to the area laws that we have observed earlier, now the entanglement does not saturate but diverges with the length of the half-chain.[3] The behavior observed here will be mirrored by a similar logarithmic divergence in critical quantum spin chains and fermionic systems. This will be discussed in the following section.

B. Fermionic chain and the XY model

Following the initial work on bosonic models of ref.[7] similar questions were explored in fermionic systems and the associated spin models. The numerical studies in refs.[147,223] presented a significant first step in this direction. Their key observation, later confirmed rigorously[132,134,138] using techniques that we will sketch in this section, is that the scaling of the entanglement entropy as a function of the block size appears to be related to the system being *quantum critical or not*. Again, for a gapped system, away from a quantum critical point, the entanglement entropy would saturate, i.e., an area law holds. In turn in all cases when the system was critical, the numerical study indicated that the entanglement entropy grows beyond all bounds. More specifically, it grows logarithmically with the block size. This behavior is also consistent with the behavior of geometric entropies in conformal field theory[43,122] which applies to the critical points of the models discussed in refs.[93,143,147,223]. The intriguing aspect here is that being critical or not is not only reflected by the scaling of expectation values of two-point correlators, but in fact by the ground state entanglement, so genuine quantum correlations.

This section defines the setting, introduces the basic concepts required and outlines the rigorous results in more detail. Fermionic quasi-free models, that is Hamiltonians that are quadratic in fermionic operators f_i and $f_i^\dagger$,

$$H = \frac{1}{2}\sum_{i,j\in L}\left(f_i^\dagger A_{i,j} f_j - f_i A_{i,j} f_j^\dagger + f_i B_{i,j} f_j - f_i^\dagger B_{i,j} f_j^\dagger\right) \quad (11)$$

may be treated by similar analytical techniques and follow similar intuition to the bosonic case. In eq. (11), to ensure Hermiticity of the Hamiltonian, $A^T = A$ and $B^T = -B$ must hold for the matrices A and B defining the coupling. The role of the canonical coordinates is taken by the *Majorana operators* $x_j = (f_j^\dagger + f_j)/\sqrt{2}$ and $p_j = \mathrm{i}(f_j^\dagger - f_j)/\sqrt{2}$, while the role of symplectic transformations is taken by orthogonal transformations. The energy gap above the ground state is given by the smallest non-zero singular value of $A + B$.

Note that, in contrast to the bosonic case, the ground state is $2^{|L|-\mathrm{rank}(A+B)} =: q$-fold degenerate. We define the ground state expectation $\langle\cdot\rangle = \mathrm{tr}[\cdot P_0]/q$, where P_0 projects onto the ground state sector. Then, as in the bosonic case, the ground state is fully characterized by two-point correlations embodied in the covariance matrix with entries

$$-\mathrm{i}\Gamma_{i,j} = \langle [r_i, r_j]\rangle = \langle r_i r_j\rangle - \langle r_j r_i\rangle,$$

where now $r = (x_1, \ldots, x_{|L|}, p_1, \ldots, p_{|L|})$ collects Majorana operators. One then finds

$$\Gamma = \begin{pmatrix} 0 & -V \\ V^T & 0 \end{pmatrix}, \quad V = |A+B|^+(A+B), \quad (12)$$

where $\cdot^+$ indicates the Moore-Penrose generalized inverse of a matrix[124], i.e., for a unique ground state one simply has $V = |A+B|^{-1}(A+B)$. The entropy of a contiguous block I of fermions in the ground state can be expressed in terms of the singular values σ_k of the principle submatrix V_I of V. One finds $S(\rho_I) = \sum_k f(\sigma_k)$, where

$$f(x) = -\frac{1-x}{2}\log_2\left(\frac{1-x}{2}\right) - \frac{1+x}{2}\log_2\left(\frac{1+x}{2}\right). \quad (13)$$

All the above holds for general lattices but for the moment we will turn to a discussion of $L = \{1, \ldots, N\}$.

We have started the discussion on the level of fermionic operators to highlight the similarity to the bosonic case. It is important to note however, that these fermionic models share a close relationship to natural spin models in the 1-D setting. This is revealed by the *Jordan Wigner transformation* which relates fermionic operators with spin operators according to

$$\sigma_i^z = 1 - 2f_i^\dagger f_i, \quad \frac{\sigma_i^x + \mathrm{i}\sigma_i^y}{2} = \prod_{k=1}^{i-1}(1 - 2f_k^\dagger f_k) f_i, \quad (14)$$

where $\sigma_i^x, \sigma_i^y, \sigma_i^z$ denote the Pauli operators associated with site $i \in L$. The fermionic model eq. (11) is hence equivalent to a spin model with short or long-range interactions.

The most important model of this kind is the XY model with a transverse magnetic field, with nearest-neighbor interaction, $A_{i,i} = \lambda$, $A_{i,j} = -1/2$ if $\mathrm{dist}(i,j) = 1$, and

[3] Compare also the divergence of the entanglement entropy in collectively interacting chains[211].

$B_{i,j} = -B_{j,i} = \gamma/2$ for $\mathrm{dist}(i,j) = 1$. This gives rise to

$$H = -\frac{1}{2}\sum_{\langle i,j\rangle}\left(\frac{1+\gamma}{4}\sigma_i^x\sigma_j^x + \frac{1-\gamma}{4}\sigma_i^y\sigma_j^y\right) - \frac{\lambda}{2}\sum_{i\in L}\sigma_i^z, \tag{15}$$

where $\langle i,j\rangle$ denotes summation over nearest neighbors, γ is the anisotropy parameter, and λ an external magnetic field.[4] Once again, translational invariance of the model means that the spectrum can be readily computed by means of a discrete Fourier transform. One obtains

$$E_k = \left((\lambda - \cos(2\pi k/N))^2 + \gamma^2\sin^2(2\pi k/N)\right)^{1/2},$$

for $k = 1,\ldots,N$. This is a well-known integrable model[11,154].

In the plane defined by (γ,λ) several critical lines can be identified: Along the lines $|\lambda| = 1$ and on the line segment $\gamma = 0$, $|\lambda| \leq 1$, the system is critical, $\lim_{N\to\infty}\Delta E(N) = 0$. For all other points in the (γ,λ) plane the exists a $c > 0$ independent of N such that $\Delta E \geq c$. The class of models with $\gamma = 1$ are called *Ising model*. The most important case subsequently is the isotropic case of the XY model, then often referred to as *XX model* or isotropic XY model. This is the case when $\gamma = 0$. The XX model is critical whenever $|\lambda| \leq 1$. The XX model is equivalent to the *Bose-Hubbard model* in the limit of hard-core bosons, so the Bose-Hubbard model with the additional constraint that each site can be occupied by at most a single boson.

Let us assume that we have a non-degenerate ground state, such that the entropy of entanglement $S(\rho_I)$ really quantifies the entanglement content. For the translation-invariant system at hand, the entries $V_{i,j} = V_{i-j}$ of V are given by

$$V_l = \frac{1}{|L|}\sum_{k=1}^{|L|} g_k e^{2\pi \mathrm{i} l k/|L|}, \quad g_k = \lambda_k\left(\frac{A+B}{|A+B|}\right). \tag{16}$$

The entanglement properties of the model are encoded in the numbers g_k. For $N = |L| \to \infty$, we can write

$$V_l = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi\, g(\phi) e^{\mathrm{i}l\phi},$$

for $|l| \leq N/2$, where $g : [0,2\pi) \to \mathbb{C}$ is called the *symbol* of V. Note that the *Fermi surface* is defined by the discontinuities of the symbol. In order to evaluate the entropy of a reduction $S(\rho_I)$, we merely have to know the singular values of $n \times n$-submatrices $V_I = V_n$ of V, see eq. (13). For isotropic models, i.e., for $B = 0$, V then being symmetric, the singular values are the absolute values of the eigenvalues. In other words, in order to understand the correlation and entanglement structure of sub-blocks of such systems, one has to understand properties of matrices the entries of which are of the form $T_{i,j} = T_{i-j}$. Such matrices are called *Toeplitz matrices*. An $n \times n$ Toeplitz matrix is entirely defined by the $2n-1$ numbers T_l, $l = 1-n,\ldots,n-1$.

The spectral values $\lambda_1(T_n),\ldots,\lambda_n(T_n)$ of T_n are just the zeros of the characteristic polynomial

$$\det(T_n - \lambda\mathbb{1}) = \prod_{k=1}^{n}(\lambda_k(T_n) - \lambda),$$

so in order to grasp the asymptotic behavior of the spectrum of T_n, it is sufficient to know the asymptotic behavior of this determinant expression.[5] The mathematical theory of determinants of such *Toeplitz matrices* is very much developed. The Fisher-Hartwig Theorem provides exactly the tools to study the asymptotic behavior of Toeplitz determinants in terms of the symbol. Crudely speaking, what matters are the zeros of g and the jumps: Once g is written in what can be called a normal form, one can "read off" the asymptotic behavior of the sequence of Toeplitz determinants defined by this symbol. Note that the matrices $V_n - z\mathbb{1}$ take here the role of T_n. The exact formulation of the Fisher-Hartwig Theorem is presented in the appendix.

This machinery was used in refs.[72,132,134,138] to evaluate the asymptotic behavior of the block entropy for the critical XX model and other isotropic models. In the first paper introducing this idea[134], in fact, there is a single jump from 1 to -1 in the symbol defining the Toeplitz matrices (and no zeros), which gives rise to the prefactor of $1/3$ of $\log_2(n)$ in the formula for the entanglement entropy in the XX model. This prefactor—which emerges here rather as a consequence of mathematical properties of the symbol—is related to the conformal charge of the underlying conformal field theory. In more general isotropic models, as has been pointed out in ref.[138], the number of jumps determines the prefactor in the entanglement scaling. Hence in such quasi-free isotropic fermionic models, the connection between criticality and a logarithmic divergence is very transparent and clear: If there is no Fermi surface at all, and hence no jump in the symbol,

[4] Note that the boundary conditions give rise to a (sometimes overlooked) subtlety here. For open boundary conditions in the fermionic model, the Jordan Wigner transformation relates the above fermionic model to the spin model in eq. (15) with open boundary conditions. For periodic boundary conditions, the term $f_N^\dagger f_1$ is replaced by the operator $(\prod_{j=1}^N(2f_j^\dagger f_j - 1))f_N^\dagger f_1$. Hence, strictly speaking, the periodic fermionic model does not truly correspond to the periodic XY model[65]. Importantly, the degeneracy of ground states is affected by this. For a degenerate ground state, the *entanglement of formation*[20], the *relative entropy of entanglement*[212], the *distillable entanglement*[20] or the logarithmic negativity of the ground state sector are the appropriate entanglement quantifiers[125,186], and no longer the entropy of entanglement. Only in the case that for large subsystems n one can almost certainly locally distinguish the finitely many different ground states, the entropy of entanglement for each of the degenerate ground states still gives the correct value for the entanglement of a subsystem.

[5] Once this quantity is known, one can evaluate the entropy by means of

$$S(\rho_I) = \sum_{k=1}^{n} f(|\lambda_k|) = \lim_{\varepsilon\to 0}\oint \frac{\mathrm{d}z}{2\pi\mathrm{i}} f_\varepsilon(z)\frac{\partial}{\partial z}\log\det(V_n - z\mathbb{1}). \tag{17}$$

Here, the integration path in the complex plane has been chosen to contain all eigenvalues $\lambda_k(V_n)$. The function f_ε is a continuation of f: We require that $\lim_{\varepsilon\to 0} f_\varepsilon(z) = f(|z|)$, including the parameter ε such that f_ε is analytic within the contour of integration.

the system will be gapped and hence non-critical. Then, the entropy will saturate to a constant.

In contrast, in case there is a Fermi surface, this will lead to jumps in the symbol, and the system is critical. In any such case one will find a logarithmically divergent entanglement entropy. The prefactor is determined by the number of jumps. So more physically speaking, what matters is the number of boundary points of the Fermi surface in the interval $[0, 2\pi)$. So—if one can say so in a simple one-dimensional system—the "topology of the Fermi surface determines the prefactor". This aspect will later be discussed in more detail. Refs.[132,134,138] find the following:

Theorem 2 (Critical quasi-free fermionic chains) *Consider a family of quasi-free isotropic fermionic Hamiltonians with periodic boundary conditions as in eq. (11) with $B = 0$. Then, the entanglement entropy of a block of $I = \{1, \ldots, n\}$ continuous spins scales as*

$$S(\rho_I) = \xi \log_2(n) + O(1).$$

$\xi > 0$ is a constant that can be related to the number of jumps in the symbol (defined above). This applies, e.g., to the scaling of the entanglement entropy in the XX spin model, for which

$$S(\rho_I) = \frac{1}{3} \log_2(n) + O(1).$$

The constant ξ is not to be mistaken for the conformal charge which will be discussed later. These arguments correspond to the isotropic model with $B = 0$, where the Fisher-Hartwig machinery can be conveniently applied. In contrast, the anisotropic case, albeit innocent-looking, is overburdened with technicalities. Then, in order to compute the singular values of submatrices of V as in eq. (12), it is no longer sufficient to consider Toeplitz matrices, but *block Toeplitz* matrices where the entries are conceived as 2×2 matrices. This setting has been studied in detail in ref.[132] in case of a non-critical anisotropic system, finding again a saturation of the entanglement entropy and in ref.[133] where the prefactor of the area law for the entanglement entropy in the gapped XX model was computed rigorously. Ref.[93] discusses also other Renyi entropies in this model.

Using an idea that originates from the concept of single-copy entanglement all these technicalities may be avoided and we can prove that the entanglement entropy diverges at least logarithmically in case of a critical (anisotropic) Ising model. The Ω notation just means that there is asymptotically a lower bound with this behavior.[6]

Theorem 3 (Divergence for the critical Ising model) *The entanglement entropy in the critical Ising model scales as*

$$S(\rho_I) = \Omega(\log_2 n). \tag{18}$$

[6] $f(n) = \Omega(g(n))$ if $\exists C > 0, n_0 : \forall n > n_0 : |Cg(n)| \leq |f(n)|$.

The starting point leading to this result from ref.[72] is a lower bound in the operator norm of ρ_I leading to

$$\begin{aligned} -\log_2 ||\rho_I||_\infty &= -\log_2 \det((\mathbb{1} + V_n)/2) \\ &\geq -\frac{1}{2} \log_2 |\det(V_n)|. \end{aligned} \tag{19}$$

This makes a big difference: We now no longer need the singular values of V_n (which would lead to an enormously complicated block Toeplitz expression, for a case for which the Fisher-Hartwig conjecture has not yet been proven). Instead—as the absolute value of the determinant is just as well the product of the absolute values of the eigenvalues as of the singular values—we can use the ordinary Fisher-Hartwig machinery to get an asymptotic handle on eigenvalues. For the critical Ising model, we can again find an explicit factorization of the Fisher-Hartwig-symbol, in terms of a function reflecting a single discontinuity and an analytical function. Using again a proven instance of the Fisher-Hartwig conjecture[152]—albeit a different one than used in the case of an isotropic model—one finds the bound as in eq. (18); for details see ref.[72]. The entanglement in two blocks of the critical Ising model has been studied in ref.[81].

Another useful starting point to obtain bounds to entanglement entropies in fermionic systems is to make use of quadratic bounds to the entropy function: Such quadratic bounds immediately translate to a bound to the function f in the expression of the entropy of a fermionic state in terms of the covariance matrix as in eq. (13), as

$$(1 - x^2)^{1/2} \geq f(x) \geq 1 - x^2. \tag{20}$$

This immediately translates to a bound of the form $\mathrm{tr}[(\mathbb{1} - V_I V_I^T)^{1/2}] \geq S(\rho_I) \geq \mathrm{tr}[\mathbb{1} - V_I V_I^T]$, where V_I is the submatrix of V associated with the interior I[83]. These bounds have also been exploited in the higher-dimensional analysis in ref.[238].

A method to obtain area laws in particular for symmetrically bisected quantum chains is the so-called method of *corner transfer matrices*. This method has first been used in ref.[36] for the computation of the entanglement entropy, using ideas going back to ref.[177]. The infinite sum of ref.[36] could be performed in ref.[176], giving also rise to a formula for the entanglement entropy in the XX model. This idea has also applied to further models in ref.[231].

To conclude the discussion of critical quasi-free fermionic models let us note that the correspondence of being critical (gapped) and having a logarithmically divergent (saturating) entanglement entropy holds true for *local* systems only. If one allows for long-ranged interactions, then one can indeed find gapped, non-critical models that exhibit a logarithmically divergent entanglement entropy[73]:

Theorem 4 (Gapped model with long-range interactions) *There exist models with long-range interactions, the coupling strength being bounded by $r/dist(j,k)$ for some constant $r > 0$, such that for some constant $\xi > 0$*

$$S(\rho_I) = \xi \log_2(n) + O(1).$$

Hence, being gapped—albeit having power-law correlations—does not necessarily imply an area law. If one allows for long-range interactions (and a *fractal structure* of the Fermi-surface), one can show that one can even approach arbitrarily well a *volume law* for the entanglement entropy[83,85]. Interestingly, states that are defined by *quantum expanders* can have exponentially decaying correlations and still have large entanglement, as has been proven in refs.[19,112]. These models give again rise to long-range Hamiltonians, but they still very clearly demonstrate a strong distinction between correlations and entanglement.

C. General gapped local spin models

We will now turn to the discussion of general 1-D gapped spin models with local interactions, where each site supports a d-dimensional quantum system. As it is stated rigorously in the theorem below for such models *an area law always holds*[111]. The proof is deeply rooted in the existence of *Lieb-Robinson bounds* which have also been essential in the proof of the exponential decay of correlation functions in gapped local models[115,161].[7]

As we allow for arbitrary d, it is sufficient to consider Hamiltonians on the chain $L = \{1, \ldots, N\}$ that have interactions only to nearest neighbors. Then

$$H = \sum_{j \in L} H_{j,j+1} \tag{21}$$

where $H_{j,j+1}$ is supported on sites j and $j+1$. We also impose a constraint of *finite-interaction strength* in that the operator norm $\|H_{i,i+1}\| \leq J$ for some $J > 0$. Then ref.[111] finds:

Theorem 5 (Area law for gapped spin chains) *Consider a local Hamiltonian H as in eq. (21) with finite interaction strength. Suppose H has a unique ground state with a spectral gap ΔE to the first excited state. Let us as before consider the block $I = \{1, \ldots, n\}$. Then,*

$$S(\rho_I) \leq S_{max} = c_0 \xi \ln(6\xi) \ln(d) 2^{6\xi \ln(d)}, \tag{22}$$

for some numerical constant $c_0 > 0$ of order unity, and where $\xi = \max(2v/\Delta E, \xi_C)$, v is the velocity of sound and $\xi_C > 0$ is of the order of unity.

The proof of this statement is quite intricate[111] and well beyond the scope in this article. At its heart is the way locality enters by virtue of the Lieb-Robinson Theorem. It is a statement on the existence of a *speed of sound* in local Hamiltonian systems with finite-dimensional constituents: Let us imagine we single out two disjoint sets X, Y from a lattice, and consider observables A and B that have support only on X and Y, respectively. Then $[A, B] = 0$. If we evolve A with time under a local Hamiltonian H it is no longer exactly true that $A(t)$ and B commute: $A(t)$ will be significantly supported on more and more sites, "melting away", and developing a long tail in support. For short times or large distances between sets X and Y, the commutator of $A(t)$ and B will be very small. How small exactly is governed by the *Lieb-Robinson Theorem*[111,115,153,161]:

Theorem 6 (Lieb-Robinson-Theorem) *Let H be as in eq. (21) a local Hamiltonian with a finite interaction strength. Then there exists a velocity of sound $v > 0$ and $\mu, c > 0$ such that for any two operators A and B with support on disjoint sets X and Y we have that*

$$\|[A(t), B]\| \leq c \|A\| \, \|B\| \exp\left(-\mu(dist(X, Y) - v|t|)\right), \tag{23}$$

where the distance between sets is taken to be dist$(X, Y) = \min_{i \in X, j \in Y}(|i - j|)$, and where

$$A(t) = e^{\mathrm{i}Ht} A e^{-\mathrm{i}Ht}.$$

The velocity v is of order J.

This statement, natural as it may seem when viewed with a reasonable physical intuition, is a rigorous, and profound statement on how locality manifests itself in quantum lattice systems. From this bound, the *decay of correlation functions* in gapped models can be proven[115,161], an area law as above,[111], the quantization of the Hall conducance for interacting electrons[116], as well as statements concerning *propagation of quantum information and correlations* through local dynamics[31].[8] Lieb-Robinson bounds also feature in the proof of a higher-dimensional Lieb-Schultz-Mattis theorem[110,162].

We will later, in Subsection IV.G, encounter another consequence of the Lieb-Robinson theorem, namely that quenched non-equilibrium systems generically satisfy area laws when starting from a product state and undergoing time evolution under a local Hamiltonian. This perspective receives a lot of attention in the context of non-equilibrium dynamics of quantum many-body systems. Here, the Lieb-Robinson is also the basis for the functioning of numerical *light cone methods* to study time evolution of quantum many-body systems LightCone,MPS,NonMPS,NonMPS2, in which effectively, only the essential part inside the causal cone is simulated.

D. Results from conformal field theory

In critical models the correlation length diverges and the models become scale invariant and allow for a description in terms of conformal field theories. According to the universality hypothesis, the microscopic details become irrelevant for a

[7] This result is compatible with an earlier result of an area law in *1-D gapped quantum field theories*, based on the c-theorem presented in ref.[36]. This work also connected the role of the boundary points between regions I and O with the cluster decomposition in quantum field theory. In gapless system with open boundaries, the entropy is then half of the one in the situation of having periodic boundary conditions.

[8] The assumption that we have a spin system, meaning finite-dimensional local constituents, is crucial here.

number of key properties. These *universal* quantities then depend only on basic properties like the symmetry of the system, or the spatial dimension. Models from the same universality class are characterized by the same fixed point Hamiltonian under renormalization transformations, which is invariant under general rotations. Conformal field theory then describes such continuum models, which have the symmetry of the *conformal group* (including translations, rotations, and scalings). The universality class is characterized by the *central charge* c, a quantity that roughly quantifies the "degrees of freedom of the theory". For free bosons $c = 1$, whereas the Ising universality class has $c = 1/2$.

Once a model is known to be described by a conformal field theory, powerful methods are available to compute universal properties, and entanglement entropies (or even the full reduced spectra) of subsystems.[9] This approach applies for $1+1$-dimensional systems, that is with one spatial dimension. In the seminal work ref.[122] the entanglement entropy in $1+1$-dimensions has been calculated, see also refs.[43,226]. The work refs.[36,44] put this into a more general context, and also allow for non-contiguous regions I. The local spectra of the reductions are discussed in refs.[42,166,168]. Block-block entanglement is also discussed in ref.[91,156]. For a short non-technical review, see ref.[38].

Starting point of the computations is the observation that powers of the reduced density matrix ρ_I^n for any positive integer n can be computed. The series $\mathrm{tr}[\rho_I^n] = \sum_j \lambda_j(\rho_I)^n$ is absolute convergent and analytic for all $\mathrm{Re}(n) > 1$. The derivative exists, and hence one can make use of

$$S(\rho_I) = \lim_{n \searrow 1} -\frac{1}{\ln 2} \frac{\partial}{\partial n} \mathrm{tr}[\rho_I^n]$$

to compute the entanglement entropy. This procedure is typically referred to as *"replica trick"*. This leads in $1+1$-dimensions to the expression[122]

$$S(\rho_I) = \frac{c}{3} \log_2(l/a) + O(1), \tag{24}$$

where c is as above the central charge, l is the length of a single interval forming region I, a is an *ultraviolet cutoff*, corresponding to a lattice spacing, to avoid an ultraviolet divergence, cp. eq. (10). The above constant C is hence nothing but $c/3$. This divergence is also removed by using the mutual information[46], see Section V.B. The offset constant in eq. (24) is non-universal. So the logarithmic divergence of the entanglement entropy in the length of the interval is recovered here. From the expression given in ref.[36] for ρ_I^n, one also finds for the Renyi entropies for $\alpha > 1$

$$S_\alpha(\rho_I) = \frac{c}{6} (1 + 1/\alpha) \log_2(l/a) + O(1).$$

[9] Conformal field theory provides—in this context specifically in $1+1$-dimensions—a powerful repertory of methods to compute quantities that are otherwise inaccessible especially for non-integrable models. From a mathematical physics perspective, it is the lack of a rigorous proof of the relationship between the lattice model and the conformal field theory that makes such a treatment, pedantically speaking, non-rigorous.

If one is close to the critical point, where the correlation length $\xi > 0$ is large but finite, one can often still effectively describe the system by a conformal field theory. One then obtains for the entanglement entropy[36] (compare also ref.[105])

$$S(\rho_I) \to \frac{c}{3} \log_2(\xi/a).$$

E. Disordered spin chains

Natural systems will generally exhibit a certain amount of *quenched disorder* which means that the model parameters are drawn randomly and the resulting correlation functions or entanglement entropies $\mathbb{E}[S(\rho_I)]$ have to be considered as being averaged over the a priori distributions, with average $\mathbb{E}$. The critical behavior of quantum spin chains with "quenched" disorder is remarkably different from its counterpart in the corresponding pure case, in several respects. Hence, it is only natural to ask whether the scaling of the entanglement entropy is influenced by having some disorder in the model. This question has first been posed in ref.[189] for the spin-$1/2$ *random anti-ferromagnetic Heisenberg model*,

$$H = \sum_{j \in L} J_j \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \sigma_j^z \sigma_{j+1}^z \right),$$

with $\{J_j\}$ drawn from a suitable continuous distribution. The low energy properties of this model, along with the random XX model, are described by what is called a *random-singlet phase*[145]. Using a real-space renormalization group approach[189], the intuition can be developed that in this phase, singlets form in a random fashion, distributed over all length scales. The entanglement entropy of a sub-block is hence obtained by effectively counting the singlets that cross the boundary of the sub-block. This intuition is further developed in ref.[190]. Within the framework of a real-space renormalization group approach—it is shown that the averaged entanglement entropy for a large class of disordered models scales like

$$\mathbb{E}[S(\rho_I)] = \frac{\gamma}{3} \log_2(n) + O(1).$$

In this class one hence observes universal behavior in the scaling of the averaged entanglement entropy. The intuition elaborated on above is further corroborated by work on the *random anti-ferromagnetic XXZ chain*[126], described by a Hamiltonian

$$H = \sum_{j \in L} J_j \left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \Delta_j \sigma_j^z \sigma_{j+1}^z \right),$$

where $\{J_j\}$ are positive, uncorrelated random variables drawn from some probability distribution and the uncorrelated anisotropy parameters $\{\Delta_j\}$ are also taken from a probability distribution. In this work, the observation is further explored that the scaling of the averaged entanglement entropy can be universal, even if correlation amplitudes are not, in that they would manifest themselves only in non-leading order terms in the entanglement entropy. This intuition is also further corroborated in refs.[155,240], on the entanglement entropy in a 2-D situation.

From a fully rigorous perspective, the entanglement entropy in the random Ising model—for which ref.[189] finds a scaling with the effective central charge of $\ln(\sqrt{2})$— has recently been revisited with methods and ideas of *percolation theory*[100]. This approach is more limited than the Fisher-Hartwig techniques in terms of the class of models that can be considered—the Ising model only—but is more powerful in that also disordered systems with no translational invariance can be considered.

Theorem 7 (Non-translationally invariant Ising model) *Consider the Ising model*

$$H = -\frac{1}{2}\sum_{j,k\in L, dist(j,k)=1} \lambda_{j,k} X_j X_k - \sum_{j\in L} \delta_j Z_j,$$

where $\lambda_{j,k} \geq 0$ and $\delta_j \geq 0$ are the spin-coupling and external field intensities, respectively, which may depend on the lattice site in a non-translationally invariant system. The total number of sites is $N = 2m + n + 1$, with $\{1, \ldots, n\}$ being the distinguished region. Then there exist γ, α, C with properties as in the subsequent footnote [10]. If $\gamma > 4\ln 2$, then there exist constants $c_1, c_2 > 0$ depending only on γ such that

$$S(\rho_I) \leq c_1 \log_2(n) + c_2, \tag{26}$$

for $m \geq 0$, so the entanglement entropy is at most logarithmically divergent.

The general picture that emerges is that the entanglement entropy scales as in the non-random case, but with a different prefactor in the logarithmic divergence. This seems natural, as the disorder tends to "localize" excitations, and hence, with faster decaying correlations one would expect less entanglement to be present in the system. Yet, there are exceptions: Cases in which one does find a logarithmic divergence, but with a larger prefactor compared to the non-random case. This includes the random quantum Potts model with spin dimension d: Here, for the very large dimension of $d > 41$ one finds a larger factor[197]. The exploration and complete classification of the role of disorder to entanglement properties of ground states—including non-critical and higher-dimensional models—remains an interesting challenge.

F. Matrix-product states

Matrix-product states (MPS) play a very central role in the context of area laws for the entanglement entropy. They form the class of states that is at the root of the workhorse of simulating strongly correlated quantum chains — DMRG. This link will be elaborated upon in detail in Section VI. Here, we focus on the entanglement and correlation properties of MPS. In the original sense, MPS are states defined on quantum chains consisting of N sites, each constituent being a d-level system. There are several ways of defining and introducing MPS, the relationship of which may not be entirely obvious. This is also the reason that it was left unnoticed for some time that MPS— as being generated in DMRG—and *finitely correlated states*[84]—as being considered in the mathematical physics literature—are up to translational invariance essentially the same objects.

One way of looking at MPS is via a *valence-bond picture*: For each of the constituents one introduces a virtual substructure consisting of two particles. Per site with Hilbert space $\mathbb{C}^d$, one associates a Hilbert space $\mathbb{C}^D \otimes \mathbb{C}^D$ for some D. This D is sometimes referred to as the dimension of the *correlation space*, or D called the *auxiliary or virtual dimension*.

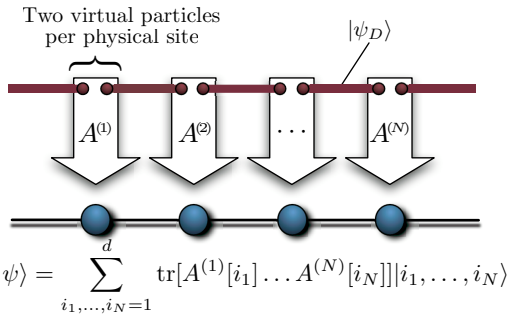


FIG. 2 The valence-bond picture underlying matrix-product states as explained in the main text.

These D-dimensional virtual systems are thought to be prepared to a maximally entangled state with each one particle of each of the neighbors, arranged on a ring (see Fig. 2). In other words, one starts from a pure state defined by the state vector $|\psi_D\rangle^{\otimes n}$, where we have defined the maximally entangled state vector as

$$|\psi_D\rangle = \frac{1}{\sqrt{D}}\sum_{k=1}^{D} |k,k\rangle. \tag{27}$$

Then, one applies a local linear map to each of the pairs of systems associated with every physical constituent in the center of the chain,

$$A^{(k)} = \sum_{j=1}^{d}\sum_{a,b=1}^{D} A^{(k)}_{a,b}[j]|j\rangle\langle a,b|, \tag{28}$$

where $k = 1, \ldots, N$. This procedure will prepare a certain class of states: Indeed the MPS. We may conceive for each site $k \in L$ the collection of complex numbers $A^{(k)}_{a,b}[j]$ as the elements of d matrices $A^{(k)}[1], \ldots, A^{(k)}[d]$. For a quantum spin chain with $d = 2$, we hence simply have two matrices $A^{(k)}[1], A^{(k)}[2]$ per site. This procedure of locally projecting

[10] Let $\lambda, \delta \in (0,\infty)$ and write $\theta = \lambda/\delta$. There exist constants $\alpha, C \in (0,\infty)$—depending on θ only— and a constant $\gamma = \gamma(\theta)$ satisfying $0 < \gamma < \infty$ if $\theta < 1$, such that, for all $n \geq 1$,

$$\|\rho_N^n - \rho_M^n\| \leq \min\{2, Cn^\alpha e^{-\gamma N}\}, \qquad 2 \leq N \leq M. \tag{25}$$

Here ρ_N^n denotes the reduced state of n sites in a system of total size N. One may find such a γ satisfying $\gamma \to \infty$ as $\theta \searrow 0$.

to the physical dimension d gives rise to state vectors of the form

$$|\psi\rangle = \sum_{i_1,\ldots,i_N=1}^{d} \mathrm{tr}[A^{(1)}[i_1]\ldots A^{(N)}[i_N]]|i_1,\ldots,i_N\rangle. \quad (29)$$

This is the most frequently used form of representing *matrix product states*. For open boundary conditions in a chain $L = \{1,\ldots,N\}$, $A^{(1)}[i_1]$ and $A^{(N)}[i_N]$ are row and column vectors, respectively. MPS are described by only a number of parameters polynomial in the system size, $O(dND^2)$, in contrast to the scaling of the dimension of the full Hilbert space $(\mathbb{C}^d)^{\otimes N}$, which is exponential in N.[11]

The particular construction of MPS immediately shows that MPS satisfy an area law. In fact, it follows trivially from their definition (see also fig. 2) that

$$S(\rho_I) \leq 2\log_2(D),$$

so the entanglement entropy is always bounded from above by a constant in N. MPS have hence an in-built area law property. As we have shown earlier the ground states of a variety of Hamiltonians exhibit exactly such an area-scaling when the system is non-critical but a logarithmic divergence when the system is critical. This already suggests that MPS may be a good description for ground states of non-critical systems but that this description may become less efficient in critical systems. Indeed, it will be discussed and highlighted later in this article that systems satisfying an area law can be economically represented as MPS so that MPS with a small auxiliary dimension D can indeed typically approximate ground states of local Hamiltonians.

G. Single-copy entanglement

The entanglement entropy—occupying center stage in this article—quantifies entanglement in a very precise sense: For pure states it is the *distillable entanglement*[125,186], so the rate with which one can locally extract maximally entangled pairs from a supply of identically prepared system. Specifically, local refers here to a subsystem I of the system, but to a collective operation on many identically prepared states. In a quantum many-body system, needless to say, this means that one performs operations that are local to all constituents in I collectively in all specimens at hand.

When having the entanglement content in mind, one can equally reasonably ask how much entanglement is contained in a *single* spin chain. The concept of single-copy entanglement grasps this notion of distilling entanglement from a single specimen of a quantum spin chain with certainty.

If D is the largest integer such that one can deterministically transform a state into the maximally entangled state $|\psi_D\rangle\langle\psi_D|$ (see eq. (27)) by local operations and classical communication (LOCC), i.e.

$$\rho \mapsto |\psi_D\rangle\langle\psi_D|, \quad (30)$$

one assigns the value $E_1 = \log_2(D)$ to the state as its *single-copy entanglement*. For pure states, such transformations on the level of specimens are perfectly well understood[135,163,220] and is linked to the well-established theory of majorization in linear algebra[124]. For our present purposes, for a pure state $\rho = |\psi\rangle\langle\psi|$, we find that eq. (30) holds if and only if $\|\rho_I\| \leq 1/D$. Hence,

$$E_1(\rho_I) = \log_2(\lfloor \|\rho_I\|^{-1} \rfloor).$$

This, in turn, means that the single-copy entanglement can be derived from the α-Renyi entropy of the reduction in the limit of large α. A quite surprising insight is that in critical systems, we do not only find a local spectrum leading to the logarithmic divergence of the entanglement entropy. But that there is more structure to the spectrum, governing all of its Renyi entropies. For example, for quasi-free models, we find that once the entanglement entropy diverges, so does the single-copy entanglement, with a prefactor that is asymptotically exactly half the value of the entanglement entropy[72,168].

Theorem 8 (Single-copy entanglement) *Consider a family of quasi-free fermionic Hamiltonians as in Theorem 2. Then, whenever the entanglement entropy scales as*

$$S(\rho_I) = \xi\log_2(n) + O(1),$$

for some constant $\xi > 0$, then the scaling of the single-copy entanglement is found to be

$$E_1(\rho_I) = \frac{\xi}{2}\log_2(n) + O(1).$$

This means that exactly half the entanglement can be distilled from a single critical chain than what is available as a rate in the asymptotic setting[168,179]). This finding has also been corroborated by the behavior of all critical models for which the local spectra can be described by their conformal field theory in quite some generality[168]. Ref.[241] studies Renyi entropies in *boundary critical phenomena*, and hence also arrives at a relationship between the entanglement entropy and

[11] If one allows D to (exponentially) grow with the system size, one can easily show that actually every state vector from $(\mathbb{C}^d)^{\otimes N}$ can be represented as a MPS of the form as in eq. (29). It is important to note that MPS can not only be described with linearly many parameters in the system size: One can also efficiently compute local properties from them, which is a property not merely following from the small number of parameters to define them. For expectation values of observables having a non-trivial support on sites $k,\ldots,k+l \in L$, we find $\langle S_k\ldots S_{k+l}\rangle = \mathrm{tr}[E_{\mathbb{1}}^{(1)}\ldots E_{\mathbb{1}}^{(k-1)}E_{S_k}^{(k)}\ldots E_{S_{k+l}}^{(k+l)}E_{\mathbb{1}}^{(k+l+1)}\ldots E_{\mathbb{1}}^{(N)}]$, where the *transfer operators* are defined as

$$E_S^{(l)} = \sum_{j,k=1}^{d}\langle j|S|k\rangle\left(A^{(l)}[k]\otimes(A^{(l)}[j])^*\right),$$

the star denoting complex conjugation. The decay of correlation functions can also be studied: If all matrices are the same per site, $A^{(k)}[j] =: A[j]$ for all $j \in L$, and similarly define $E_{\mathbb{1}}$, then one finds

$$\langle S_kS_{k+l}\rangle - \langle S_k\rangle\langle S_{k+l}\rangle = O(|\lambda_2(E_{\mathbb{1}})|^{l-1}),$$

where $\lambda_2(E_{\mathbb{1}})$ denotes the second to largest eigenvalue of the transfer operator of the identity $E_{\mathbb{1}}$.

the single-copy entanglement. Ref.[191] considers the entropy loss along the renormalization group trajectory driven by the mass term in free massive theories, and discusses also the single-copy entanglement in such situations. Ref.[139] studies the situation of single-copy entanglement in the situation of bipartite systems between blocks when there is a gap of a finite number of sites between the two blocks. Interestingly, there are critical models in which the single-copy entanglement still diverges in this sense.[12]

H. Summary of one-dimensional systems

In a nutshell, the situation in one-dimensional translation invariant models is quite clear: If a system is local and gapped, an area law always holds rigorously. In many specific models, prefactors can be computed. In contrast, if the interactions may be long-ranged area laws may be violated. For critical lattice models for which one can directly evaluate the entanglement entropy, a logarithmic divergence is encountered. This picture is supported by the findings of conformal field theory. The situation will be less transparent and more intricate in higher-dimensional models. In any case, in the light of the previous findings, one may be tempted to formulate the following conjecture on the numerical bound on the right hand side of the previously discussed area law:

Conjecture 1 (Area bound in one-dimension) *There exists a function $f : \mathbb{R}^+ \to \mathbb{R}^+$—equipped with further suitable properties—such that in any gapped one-dimensional model, we have*

$$S(\rho_I) \leq f(v/\Delta E),$$

where ΔE is the spectral gap and v is the speed of sound as used in the Lieb-Robinson bound.

Indeed, while most explicit studies do indicate a behavior linear in $\ln(1/\Delta E)$ of the entanglement entropy—the above mentioned quasi-free models—one can construct models[99,130] for which one finds a dependence which is polynomial in $1/\Delta E$.

IV. AREA LAWS IN HIGHER DIMENSIONS

For a chain, to satisfy an "area law" for the entanglement entropy means simply that it saturates with increasing block sizes. Needless to say, the notion of having entropic quantities scaling like the boundary area of a subregion becomes specifically relevant in case of higher dimensions: Then the boundary of the region I is a non-trivial object. Now we are in a position to approach the question: Given a ground state of a quantum many-body system, does the entanglement entropy of a subregion I fulfill an area law? This question has been initiated in refs.[23,207], where also a numerical answer has been found.

The answer to this question for ground states is very much developed in case of—once again—quasi-free bosonic or fermionic models. Even in such systems, the rigorous answer to this question will turn out to be technically quite involved. The reason for these technicalities is essentially rooted in the very fact that one distinguishes a subregion, thus, e.g., breaking translational symmetry of translationally invariant systems, and analytical methods are hard to come by. The first rigorous higher-dimensional area law has been proven—to the knowledge of the authors—in refs.[184], with refinements for arbitrary harmonic interactions in ref.[55,57] so that for such bosonic free models the problem can be considered solved in all generality forming a "laboratory" of what one should expect in general systems.

For critical fermionic models[12,56,86,96,151,238], quite intriguingly, one can find small violations of area laws: The area law is then only satisfied up to a logarithmic correction. In this section, we will discuss quasi-free models in great detail. Beyond such quasi-free models, no rigorous results are known for states at zero temperature, with the exception of classes of states that satisfy an area law by their very construction, and a subsection will specifically be devoted to those.

The models discussed here, however, do provide a clear intuition: Whenever one has a gapped and local model, and hence a length scale provided by the correlation length, one should reasonably expect an area law to hold. In cases where the number of eigenstates with vanishing energy-density is not exponential in the volume - a technical condition in its own right - one can even prove an area law with at most a logarithmic correction from a sufficient decay of correlations[158]. The converse is not true, as we will see later, and one can have area laws even for critical systems in which the correlation length does not provide a length scale. For systems at non-zero temperatures, by contrast, the entropy of entanglement neither forms a meaningful measure of entanglement nor for quantum correlations. For appropriately defined measures of correlations, however, one can restore an area law which holds in generality for a large class of systems.

A. Quasi-free bosonic and fermionic models: Sufficient conditions for an area law

We will follow the general description of ref.[57], where we think of the model being defined on a general lattice L specified by a general simple graph. We consider quadratic bosonic Hamiltonians as in eq. (4) and quadratic fermionic Hamiltonians as in eq. (11). The key step is to relate correlation functions to entropic quantities. As before in the case of a harmonic chain, it is very involved to think of the entropy of entanglement itself. What comes to our help, however, is again that we can use the logarithmic negativity as an upper bound to

[12] Note that the single copy entanglement still grasps bipartite entanglement in a quantum many-body system. Identifying scaling laws for genuine multi-particle entanglement is an interesting enterprise in its own right. Notably, refs.[26,167] consider the *geometric entanglement* (the logarithm of the largest Hilbert-Schmidt scalar product with a pure product state) and relate it to the conformal charge of the underlying model.

the entanglement entropy (see eq. (6)). The logarithmic negativity is easier to treat analytically, as we can at all times refer to the full system, and not to subsystems I. In fact, we find that the logarithmic negativity can be bounded from above by the L_1-norm of a submatrix of the covariance matrix[184]. For fermions in turn, the entropy may be bounded directly using the bound in eq. (20).

Theorem 9 (Entropic bounds from matrix norms) *The entanglement entropy of ground states of quadratic bosonic Hamiltonians as in eq. (4) satisfies*

$$S(\rho_I) \leq E_N(\rho, I) \leq 8\|\Gamma_x\| \sum_{i\in I, j\in O} |\langle p_i p_j\rangle| \,.$$

The entanglement entropy of unique ground states of quadratic fermionic Hamiltonians as in eq. (11) satisfies

$$S(\rho_I) \leq 2 \sum_{i\in I, j\in O} |\langle f_i^\dagger f_j\rangle + \langle f_i f_j\rangle|.$$

This is a key tool towards proving the main theorem: We can reduce the evaluation of an entropic quantity to a counting argument over terms that can be evaluated from two-point correlators. Note that the use of the logarithmic negativity results in an important simplification of the problem. This shows that ideas from quantum information theory indeed help in finding proofs of statements of the scaling of the entanglement entropy.

We are now in the position to state the bound of the scaling of the ground state entanglement in the boundary area $s(I)$, eq. (1), of the distinguished region I[53,55,57]. It is remarkable that merely the decay of two-point correlations matter here, and that even some critical models will give rise to an area law, as long as the algebraic decay of correlations is sufficiently strong.

Theorem 10 (Quadratic Hamiltonians on general lattices) *Let $\eta = \mathcal{D} + 1 + 2\varepsilon$, $\varepsilon > 0$, and assume that the ground state is unique and fulfills for $i, j \in L$, $i \neq j$, and some constant K_0*

$$\frac{K_0}{dist^\eta(i,j)} \geq \begin{cases} |\langle p_i p_j\rangle| & \text{for bosons,} \\ |\langle f_i^\dagger f_j\rangle + \langle f_i f_j\rangle| & \text{for fermions.} \end{cases}$$

Then

$$S(\rho_I) \leq K_0 c_{\mathcal{D}} \zeta(1+\varepsilon) s(I) \times \begin{cases} \|\Gamma_x\| & \text{for bosons,} \\ 1 & \text{for fermions,} \end{cases}$$

where ζ is the Riemann zeta function and the constant $c_{\mathcal{D}}$ depends only on the dimension of the lattice.

A general version of what one should expect to be true provides the connection to the spectral gap: For gapped models the correlation functions decay exponentially with the graph theoretical distance. One cannot apply the Lieb-Robinson Theorem to prove this, unfortunately, as the involved operators are unbounded. Hence, a technique that is applicable to describe clustering of correlations in such models had to be developed. The ideas of the proof go back to ref.[21], generalized to arbitrary lattices in ref.[53,57]. Key ideas of the proof are polynomial approximations in the sense of Bernstein's Theorem. For a thorough discussion of clustering of correlations in translation-invariant harmonic systems, see ref.[201]. For general lattices and gapped quadratic bosonic and fermionic Hamiltonians, one finds that two-point correlation functions decay exponentially. Together with the above theorem this leads to an area law whenever the model is gapped:

Corollary 1 (Area law for gapped quasi-free models) *The entanglement entropy of ground states of local gapped models of the type of eq. (4) for bosons and of eq. (11) for fermions for arbitrary lattices $G = (L, E)$ and arbitrary regions I satisfies for a suitable constant $\xi > 0$*

$$S(\rho_I) \leq \xi s(I).$$

B. Logarithmic correction to an area law: Critical fermions

What can we say about situations in which the previous sufficient conditions are not satisfied? Specifically, how is the scaling of the entanglement entropy modified in case of critical fermionic models? This is the question that will feature in this subsection. Following the bosonic result in refs.[57,184], the entanglement entropy in fermionic models was first studied in ref.[238] for cubic lattices. Here, the quadratic bound in eq. (20) plays an important role, to relate bounds to the entropy to feasible expressions of the covariance matrix of the ground state. Here, not quite an area law, but only one up to a logarithmic correction is found. The results can be summarized as follows:

Theorem 11 (Violation of area laws for critical fermions) *For a cubic sublattice $I = \{1, \ldots, n\}^{\times \mathcal{D}}$ and an isotropic quasi-free model as in eq. (11) with a Fermi sea of non-zero measure and a finite non-zero surface there exist constants $c_0, c_1 > 0$ such that the ground state fulfils*

$$c_0 n^{\mathcal{D}-1} \ln(n) \leq S(\rho_I) \leq c_1 n^{\mathcal{D}-1} \ln^2(n).$$

The stated lower bound makes use of the assumption that the Fermi surface is finite (and of a technical assumption that the sets representing the states cannot have nontrivial irrelevant directions); assumptions both of which can be removed[86].

This fermionic quasi-free case already exhibits a quite complex phase diagram[12,151]. At the same time, ref.[96] formulated a similar result, based on a conjecture on the validity of Fisher-Hartig-type scaling result for higher dimensional equivalents of Toeplitz matrices, as was further numerically corroborated in ref.[12]. A logarithmic divergence is not directly inconsistent with the picture suggested in a conformal field theory setting, as relativistic conformal field theories do not have a Fermi surface[195]. It is still intriguing that critical fermions do not satisfy an area law, but have logarithmic corrections. In this sense, critical fermionic models could be said to be "more strongly entangled" than critical bosonic models.

C. Difference between critical fermions and bosons: Half spaces

The scaling of block entropies for bosons and fermions in higher spatial dimensions hence exhibit remarkable differences. Let us consider the case of a cubic lattice of $n^{\mathcal{D}}$ sites with periodic boundary conditions and $I = \{1,\dots,m\} \times \{1,\dots,n\}^{\times \mathcal{D}-1}$ (w.l.o.g. we distinguish the first spatial dimension). Then one may transform the Hamiltonian to a system of mutually uncoupled one-dimensional chains using a unitary discrete Fourier transform. After this decoupling procedure the entanglement between I and O is given by a sum of the entanglement between the sites $I' = \{1,\dots,m\}$ and $O' = \{m+1,\dots,n\}$ of the $n^{\mathcal{D}-1}$ individual chains

$$\frac{S(\rho_I)}{s(I)} = \frac{1}{n^{\mathcal{D}-1}} \sum_{i=1}^{n^{\mathcal{D}-1}} S(\rho^i_{I'}).$$

We start with a discussion of fermions and focus on the isotropic setting ($B = 0$ in eq. (11)). After taking the limit $n \to \infty$, the asymptotic behavior in m of the entanglement $S(\rho^i_{I'})$ can then be read off Theorem 2 to yield the following statement (for technical details see ref.[56]):

Theorem 12 (Prefactor for fermionic half spaces) *Asymptotically, the entanglement entropy of fermionic isotropic models of half spaces satisfies*

$$\lim_{n\to\infty} \frac{S(\rho_I)}{s(I)} = \frac{\log_2(m)}{6} \sum_{s=1}^{\infty} s v_s + O(1).$$

Here,

$$v_s = \frac{\left|\left\{\phi \in [0,2\pi)^{\mathcal{D}-1} \,\middle|\, \sigma(\phi) = s\right\}\right|}{(2\pi)^{D-1}}$$

is the integral over individual chains ϕ with s discontinuities $\sigma(\phi)$ in their symbol.

Hence, one encounters a logarithmic divergence in m of the entanglement entropy and the pre-factor depends on the topology of the Fermi-surface: The symbols exhibit discontinuities on the Fermi-surface. If the Fermi surface is of measure zero (i.e., the set of solutions to $\lambda_\phi = 0$, $\phi \in [0,2\pi)^{\mathcal{D}}$, is countable, as, e.g., in the critical bosonic case discussed below), we have $v_s = 0$ and the system obeys the area law.

For the bosons, we discuss the important case of $m = n/2$ for the Klein-Gordon Hamiltonian as in eq. (9). After the transformation to uncoupled chains, one finds Hamiltonians for the individual chains that correspond to a nearest-neighbor coupling matrix X of the form as in Theorem 1, which yields

$$\frac{E_N(\rho, I)}{s(I)} = \int \frac{\mathrm{d}\varphi}{4(2\pi)^{D-1}} \log_2 \left(\frac{\mathcal{D} - \sum_{d=1}^{\mathcal{D}-1} \cos(\varphi_d) + 1}{\mathcal{D} - \sum_{d=1}^{\mathcal{D}-1} \cos(\varphi_d) - 1} \right)$$

in the limit $n \to \infty$. This expression is independent of the mass m and finite: For $\mathcal{D} = 2$, it evaluates to $\log_2(3+2\sqrt{2})/4$ and similarly for $\mathcal{D} > 2$. Hence, despite being critical, the system obeys an area law, in contrast to the fermionic case (for $m = n/2$ the entanglement for a critical fermionic system would diverge in n).

Hence, in quasi-free critical models, it matters whether a system is bosonic or fermionic when it comes to the question whether or not an area law holds. The above results confirm the numerical analysis of ref.[207] for critical bosonic theories, and of ref.[151] for two-dimensional fermionic systems. Motivated by these findings, ref.[66] numerically studies the non-leading order terms of an area law in nodal fermionic systems: It is found that in non-critical regimes, the leading subarea term is a negative constant, whereas in critical models one encounters a logarithmic additive term. A lesson from these higher-dimensional considerations is that the simple relationship between criticality and a violation of an area law is hence no longer valid for local lattice models in $\mathcal{D} > 1$.

D. Entanglement in bosonic thermal states

In this subsection, we briefly discuss area laws for notions of entanglement in Gibbs states,

$$\rho_\beta = \frac{\exp(-\beta H)}{\mathrm{tr}[\exp(-\beta H)]}$$

for some inverse temperature $\beta > 0$. The second moments matrix, the covariance matrix, is then found to be $\Gamma = \Gamma_x \oplus \Gamma_p$[55]

$$\begin{aligned}
\Gamma_x &= X^{-1/2}\big(X^{1/2} P X^{1/2}\big)^{1/2} (\mathbb{1} + G) X^{-1/2}, \\
\Gamma_p &= X^{1/2}\big(X^{1/2} P X^{1/2}\big)^{-1/2} (\mathbb{1} + G) X^{1/2}, \\
G &= 2\Big(\exp\big[\beta\big(X^{1/2} P X^{1/2}\big)^{1/2}\big] - \mathbb{1}\Big)^{-1}.
\end{aligned}$$

Using the the methods of refs.[21,55] one again finds the suitable decay of correlations, which can be translated into an area law for the entanglement content. Here, the result—taken from refs.[55,57]—is stated in terms of the logarithmic negativity.

Theorem 13 (Entanglement in thermal bosonic states) *The logarithmic negativity of thermal states of quadratic finite-ranged bosonic Hamiltonians as in eq. (4) for $[X, P] = 0$ satisfies $E_N(\rho, I) \leq \xi s(I)$ for a suitable constant $\xi > 0$.*

Since the logarithmic negativity is an upper bound to the *entanglement of formation* and hence the *distillable entanglement*[125,186] this implies an area law for these quantities as well. It is important to stress that the entropy of a subregion as such no longer reasonably quantifies entanglement between that subregion and the rest of the lattice: Even classically correlated separable states will in general have a positive entropy of the reduced state. The latter quantity is then indeed extensive and fulfills a volume law, unlike the entanglement content. Area laws in thermal states have further been studied in detail in ref.[5], where an emphasis has been put on identifying regions where the states become separable. Refs.[48,87] investigate thermal bound entanglement—entanglement that is not distillable—in bosonic quadratic and spin systems.

E. Results from conformal field theory

In systems with more than one spatial dimension, the situation is more intricate, and there is no general expression known for entanglement entropies in $d+1$-dimensional conformal field theories. For interesting steps into a description of systems with two spatial dimensions in the framework of conformal field theory see refs.[90,195]. For a class of critical models in two spatial dimensions (including the quantum dimer model), it is found that $S(\rho_I) = 2f_s(L/a) + cg\ln(L/a) + O(1)$, where L is the length of the boundary area, f_s is an area law prefactor that is interpreted as a *boundary free energy*, and g is a coefficient that depends on the geometric properties of the partition into I and O. That is, in addition to a non-universal area law, one finds a universal logarithmically divergent correction. For a further discussion of steps towards a full theory of entanglement entropies in $d+1$-dimensional conformal field theories, see refs.[90,195].

F. States satisfying area laws by construction: Projected entangled pair states, graph states, and entanglement renormalization

In this section, we will discuss classes of states that have the area law already built into their very construction. In this sense, they grasp the entanglement structure of local higher-dimensional models. These are *projected entangled pair states*, so matrix-product states in higher dimensions, and states from *entanglement renormalization*. They are designed to be variational states well-approximating true ground states of local many-body systems: As was already true for matrix product states, they form a complete set of variational states. Yet, typically, for a much smaller, polynomial or constant, number of variational parameters, they often deliver a very good approximation. In projected entangled pair states, locality is respected in just the same way as for MPS. Entanglement renormalization, in turn, is based on a scale-invariant tree structure, intercepted by disentangling steps, which in higher dimensions nevertheless leads to an area law for the entanglement entropy.

Projected entangled pair states (PEPS) can be thought of as being prepared as MPS in higher dimensional cubic lattices $L = \{1,\dots,N\}^{\times\mathcal{D}}$, or in fact to any lattice defined by any undirected simple graph $G = (L,E)$. In this valence bond construction, one again associates a physical space with Hilbert space $\mathbb{C}^d$ with each of the vertices L of G. Then, one places a maximally entangled pair of dimension $D\times D$ (see eq. (27)) for some positive integer D between any two vertices that are connected by an edge $e\in E$. For a cubic lattice, one hence starts from a cubic grid of maximally entangled state vectors. Then, one applies a linear map $P^{(k)}:\mathbb{C}^D\otimes\cdots\otimes\mathbb{C}^D\to\mathbb{C}^d$ to each physical site, as

$$P^{(k)} = \sum_{j=1}^{d}\sum_{i_1,\dots,i_{|S_1(k)|}}^{D} A^{(k)}_{j,i_1,\dots,i_{e_k}}|j\rangle\langle i_1,\dots,i_{|S_1(k)|}|.$$

Here, $|S_1(k)|$ is the vertex degree of the vertex $k\in L$. The resulting state vector as such hence becomes

$$|\psi\rangle = \sum_{i_1,i_2,\dots,i_{|L|}=1}^{d} \mathcal{C}[\{A^{(k)}_{i_l}\}_l]|i_1,i_2,\dots,i_{|L|}\rangle,$$

where $\mathcal{C}$ denotes a contraction of all higher-order tensors with respect to the edges E of the graph. This amounts to a summation over all indices associated with connected vertices. The objects $A^{(k)}$ are hence tensors of an order that corresponds to the vertex degree of the lattice (a second order tensor—a matrix—for a one-dimensional chain, a three order tensor in hexagonal lattices, a fourth order tensor in cubic lattices with $\mathcal{D}=2$, and so on). This construction is the natural equivalent of the valence bond construction for matrix-product states as explained in eq. (28). This ansatz as such is the one of *tensor product states* that is due to ref.[157] which in turn is generalizing earlier work on AKLT-type valence bond states in two dimensions in refs.[121,164]. The generated class of states is referred to as *projected entangled pair states*[213], reflecting the preparation procedure. PEPS states are sometimes also in higher dimensions simply referred to as *matrix product states*[113]. This ansatz has proven to provide a powerful and rich class of states. Importantly, ref.[213] provides a first simulation method based on PEPS.[13]

This class of states is complete, in that any state of a given finite lattice can be arbitrarily well approximated by such a state if D is sufficiently large. Clearly, to compute local observables in such an ansatz, one has to contract this instance of a tensor network which in 2-D is actually computationally hard.[14] It is however possible to provide approximation techniques, related to the DMRG approach, that allow for the contraction of the tensor network and then for the computation of the expectation values of local observables[131,160,216,219].

A particularly simple yet important subset of the projected entangled pair states is constituted by the so-called *graph states*[119,120,198]: They are instances of *stabilizer states*[8,98] which can be thought of as being prepared in the following fashion: On any graph $G=(L,E)$, one associates each vertex with a $\mathbb{C}^2$-spin. This spin is prepared in $|+\rangle = (|0\rangle+|1\rangle)/2^{1/2}$. Then, one applies a *phase gate*

$$U = |0,0\rangle\langle 0,0| + |0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0| - |1,1\rangle\langle 1,1|$$

to each pair of vertices that are connected by an edge. This phase gate corresponds to an Ising interaction. Clearly, this construction makes sense for any simple graph, and this is a subset of the above projected entangled pair states. Graph

---

[13] Note that 1-D MPS based on a suitable order of the constituents do not form a good approximation for 2-D models. This is essentially rooted in the observation that one should expect an area law for the entanglement entropy in gapped 2-D models. For an important early discussion of spectra of subsystems in 2-D integrable models see ref.[50] and, for a more recent discussion of the implication on DMRG, ref.[213].

[14] In fact, it is known that the exact contraction of such a tensor network is contained in the complexity class #P-complete[202]. Clearly, this means that no algorithm is known with polynomial running time.

states readily satisfy an area law by construction[106,120] as one merely needs to count the edges over the boundary of a distinguished region to obtains the entanglement entropy, then obviously linear in the boundary area.[15]

Graph states may be generalized to *weighted graph states*[6,119,183] where the edges may carry a different weight, and in turn generalize to the ansatz of a *renormalization algorithm with graph enhancement* (RAGE),[127] being a strict superset of matrix-product states and weighted graph states, one that can nevertheless be efficiently contracted. As the graph defining the (weighted)-graph state does not need to have the same structure as the graph of the physical system whose quantum state we would like to describe, (weighted)-graph states may describe volume scaling on the level of the physical system. This makes them particularly suitable for simulation of time evolution, where no area law can be expected to hold.

Yet a different class of many-body states with applications in the simulation of quantum spin systems is given by the states generated by *entanglement renormalization* (MERA)[222]. This is a class of states the construction of which is inspired by a renormalization scheme. Consider a tree tensor network with the physical sites at the end. This can be efficiently contracted. Yet, when decimating, say, two spins of one layer to a single "superspin" in the next layer in a single step of a renormalization procedure , one loses information information about the state. The idea of a MERA ansatz is to allow for *disentangling unitaries*, effectively removing entanglement from a state, before doing a renormalization step.

More specifically, consider a cubic lattice $L = \{1,\ldots,N\}^{\times\mathcal{D}}$ in some dimension $\mathcal{D}$, embodying $N^{\mathcal{D}}$ sites. Each site $j \in L$ is associated with a physical system with Hilbert space $\mathbb{C}^d$. The MERA is essentially a unitary tensor network of depth $O(\ln(N))$, preparing $|\psi\rangle$ from $|0\rangle^{\otimes N}$. It consists of layers of isometries —performing the renormalization step—and disentanglers, which minimize the entanglement in each step before the next renormalization step. This renormalization step may be labeled with a fictitous time parameter. Each of the unitary disentanglers $U \in U(d^m)$ in the disentangling layer has a finite support on m sites. In the simplest possible realization of a MERA this would be $m = 2$. The unitaries can be taken to be different in each layer, and also different from each other within the layer. Unlike PEPS, they do not give rise to strictly translationally invariant states, even if all unitaries are taken to be identical in each layer.

Such a procedure can be defined for cubic lattices of any dimension $\mathcal{D}$. In $\mathcal{D} = 1$, one does in fact not observe an area law, but typically a logarithmic divergence of the entanglement entropy, quite like in critical spin systems. Indeed, the MERA ansatz as a scale invariant ansatz is expected to be suitable to approximate critical systems well, and numerical simulations based on the MERA ansatz corroborate this intuition[63,79,80,192]. A precise connection between homogeneous instances of a MERA ansatz and conformal field theory is established in refs.[97,180]. In more than one dimension, $\mathcal{D} > 1$, MERA satisfy again an area law, as a moment of thought reveals: On encounters asymptotically linearly many unitaries over a boundary that have entangling power, rendering the computation of an upper bound to the entanglement entropy a combinatorial problem. Despite this observation, first numerical work on fermionic instances of a MERA ansatz appear to deliver promising results for strongly correlated systems[14,51,52,181].

Theorem 14 (Area laws for PEPS, graph states, and MERA) *For any finite dimension D of the virtual systems, the entanglement entropy of a projected entangled pair state satisfies $S(\rho_I) \leq s(I)D$, where as before $s(I)$ denotes the surface area of I on a graph. Hence, also graph states with a fixed vertex degree satisfy area laws. A family of states from entanglement renormalization will also satisfies an area law for cubic lattices with $\mathcal{D} \geq 2$, and a logarithmic divergence in $\mathcal{D} = 1$.*

Interestingly, based on a PEPS description, one can construct critical models that still satisfy an area law in $\mathcal{D} = 2$[219], resembling the situation for critical quasi-free bosonic systems. The validity of an area law follows trivially from construction, so the technical part in the argument amounts to showing that a model is critical. In ref.[219] this is shown by employing a quantum-classical-correspondence: Take a classical two-body spin Hamiltonian of the form $H(\sigma_1,\ldots,\sigma_N) = \sum_{\text{dist}(i,j)=1} h(\sigma_i,\sigma_j)$, $\sigma_i = 1,\ldots,d$. This Hamiltonian will have at some inverse temperature $\beta > 0$ a partition function $Z = \sum_\sigma e^{-\beta H(\sigma)}$. From this classical partition function, a quantum state can be constructed by using the Boltzmann weights as superposition coefficients,

$$|\psi_{H,\beta}\rangle = \frac{1}{Z^{1/2}} \sum_{\sigma_1,\ldots,\sigma_N} e^{-\beta H(\sigma_1,\ldots,\sigma_N)/2} |\sigma_1,\ldots,\sigma_N\rangle.$$

This state vector has the properties that for diagonal observables, it gives rise to the same expectation values and correlation functions as the corresponding classical model does, it has a simple representation as a PEPS for $D = d$, and it is—as any PEPS—the ground state of a local Hamiltonian. The classical model can then be chosen such that the appropriate decay of correlation functions follows. This construction delivers critical spin models that nevertheless satisfy an area law.

[15] The other side of the coin of the difficulty of actually contracting tensor networks, even if they correspond to states that approximate ground states satisfying area laws well, is that such states can have computational power for quantum computing. Indeed, certain graph states or *cluster states*—as they are called for a cubic lattice—are universal resources for quantum computing: Quantum computing can be done by merely applying local measurements onto single sites of such a cluster states, without the need of additional unitary control. This computational model—known as *one-way computing*[188]—can also be understood as a teleportation scheme in virtual qubits[214]. The tensor networks that occur when performing Pauli measurements can still be efficiently contracted, but not under arbitrary measurements, leading to universal computation. The program of using general projected entangled pair states in quantum computing based on measurements only has been pursued in refs.[101,102], giving rise to a wealth of new *measurement-based quantum computational models*. This also highlights how the disadvantage of having no classical efficient description can be made an advantage: One can at each instance of the computation not efficiently compute the outcome, but on a physical system realizing this model one could efficiently simulate any quantum computer.

G. Quenches and non-equilibrium dynamics

A physical setting that receives a lot of attention in the recent literature is the one of non-equilibrium dynamics of quantum many-body systems. A specifically interesting setting is the one of a sudden quench[15,37,39–41,54,58,64,73,76–78,82,114,142,185,194,205,245]: Here, the initial condition is the non-degenerate ground state of some local Hamiltonian H, with state vector $|\psi\rangle$. Then, one suddenly (locally) alters the system parameters to a new Hamiltonian V. Since $|\psi\rangle$ will typically no longer be an eigenvalue of H, one arrives at a non-equilibrium situation: The state vector's time evolution is simply given by

$$|\psi(t)\rangle = e^{-\mathrm{i}tV}|\psi\rangle.$$

Studies of instances of such complex *non-equilibrium many-body dynamics* and questions of the *dynamics of quantum phase transitions* are enjoying a renaissance recently, not the least due to the advent of the high degree of control over quantum lattice systems with cold atoms in optical lattices.[16]

For finite times, infinite *quenched systems* satisfy an area law in the entanglement entropy[31,37,73] (strictly speaking whenever one considers time evolution under local finite-dimensional Hamiltonians starting from product states). For finite systems this holds true for times that are sufficiently small compared to the system size over the speed of sound. The intuition is that when suddenly switching to a new Hamiltonian, local excitations will be created. These excitations will propagate through the lattice, but—except from an exponentially suppressed tail—at most with the Lieb-Robinson velocity of Theorem 6[31,73,114,115]. This is yet again a consequence for the approximate locality in quantum lattice systems, reminding of the situation in relativity and implies that correlations can only slowly build up, resulting in an area theorem. In turn, such a quench does in general give rise to a linear increase in the entanglement entropy, a statement that is provably correct, and has been encountered in numerous numerical studies on quenched non-equilibrium systems[15,31,37,54,73,205]. In fact, finite subsystems can locally relax in time, to appear as if they were in a thermal state[54]. These results may be summarized in the following statement.

Theorem 15 (Area laws in non-equilibrium systems) *Let $|\psi\rangle$ be a product initial state vector, and H a local Hamiltonian. Then, for any time $t > 0$ there exist constants $c_0, c_1 > 0$ such that for any subset I the entanglement entropy of the time evolved reduction ρ_I of $\rho(t) = e^{-\mathrm{i}tH}|\psi\rangle\langle\psi|e^{\mathrm{i}tH}$ satisfies*

$$S(\rho_I(t)) \leq c_0 s(I) + c_1. \tag{31}$$

Specifically, this is true for any local Hamiltonian on a cubic lattice in dimension $\mathcal{D}$. This means that for any constant time, the entanglement entropy satisfies what is called an area law. In turn, there are product initial state vectors $|\psi\rangle$ of one-dimensional spin chains, local Hamiltonians H, and constants $c_2, c_3, c_4, L_0, s_0, t_0 > 0$ such that

$$S(\rho_I(t)) \geq c_2 t + c_3,$$

for $L \geq L_0$ and $s \geq s_0$ and $t_0 \leq t \leq c_4 s$, for $I = \{1, \ldots, s\}$.

That is, for any fixed time t, one encounters an area law for the entanglement entropy, but the prefactor can grow linearly in time. In fact, by a suitable choice of blocks, one can show that a lower bound grows linearly in time! This fact is responsible for the hardness of simulating time evolution of quantum many-body systems using instances of the DMRG approach: to represent such states faithfully, exponential resources are then required. Similar bounds give rise to statements on the minimal time needed in order to prepare states with topological order using local Hamiltonians[31].

There is an interesting *localization effect* of entanglement under quenched disorder, linking to the previous discussion on ground state entanglement in disordered systems. Whereas one obtains from Lieb-Robinson bound the estimate in time

$$S(\rho_I(t)) \leq c_0|t| + c_1$$

for suitable constants $c_0, c_1 > 0$, in the disordered one-dimensional XY spin chain this bound is replaced by the tighter bound

$$S(\rho_I(t)) \leq c_0 \ln(N|t|) + c_1,$$

again for appropriate constants[35]. This means that due to quenched disorder, the growth of entanglement is merely logarithmic in time, not linear. There is an intuitive explanation for this: The linear sound cone provided by the Lieb-Robinson bounds is replaced by a logarithmically growing or even a constant one, leading to a suppressed entanglement propagation. A similar behavior is observed under time-dependent fluctuating disorder[34].

H. Topological entanglement entropy

The topological entanglement entropy is a quantity that is constructed in a fashion that enables it to characterize quantum many-body states that exhibit *topological order*, a concept introduced in refs.[228,229] (see also refs.[230,236]). On both sides of a critical point in a system undergoing a quantum phase transition, the quantum many-body system may have a different kind of quantum order; but this order is not necessarily one that is characterized by a local order parameter: In systems of, say, two spatial dimensions, topological order may occur. Topological order manifests itself in a degeneracy of the ground state manifold that depends on the topology of the entire system and the quasi-particle excitations then show an exotic type of anyonic quasi-particle statistics. These are features that make topologically ordered systems interesting for quantum computation, when exactly this degeneracy can be exploited in order to achieve a quantum memory robust

[16] The interesting situation of locally perturbing the *state* and hence generating a non-equilibrium situation has also been considered in refs.[40,77,78], where an area law is always expected to hold.

against local fluctuations. They even allow in theory for robust instances of quantum computation, then referred to as *topological quantum computation*[94,141].

The topological entanglement entropy is now designed as an instrument to detect such topological order. Introduced in refs.[140,148], it received significant attention recently[1,95,107,108,137,149,175]. The details of the relationship between positive topological entanglement entropy and topological quantum order are discussed in ref.[165].

In ref.[140] a disc in the plane I is considered with boundary length L. This disk is thought to be much larger than the correlation length, and it is hence assumed that an "area law" in the above sense holds. The entanglement entropy of ρ_I will then have the form

$$S(\rho_I) = \alpha L - \gamma + O(1), \tag{32}$$

where the last term vanishes in the limit $L \to \infty$. The prefactor α is non-universal and ultraviolet divergent. However, $\gamma > 0$ is an additive constant which is universal and characterizes a global feature of the entanglement in the ground state. This quantity is referred to as *topological entanglement entropy* in ref.[140]. To avoid ambiguities when distinguishing the constant term from the linear one in eq. (32), ref.[140] makes use of the following construction: The plane is divided into four regions, each of them being large compared to the correlation length. A, B and C are arranged as neighboring each other in three identical subparts of a disk. D is the exterior of the disk. The respective reductions to the parts are denoted as ρ_A and ρ_{AB} to regions A and jointly A and B, respectively. The *topological entropy* S_{Topo} is then defined as

$$\begin{aligned} S_{\text{Topo}} &= S(\rho_A) + S(\rho_B) + S(\rho_C) \\ &- S(\rho_{AB}) - S(\rho_{BC}) - S(\rho_{AC}) + S(\rho_{ABC}). \end{aligned} \tag{33}$$

This is a linear combination of entropies of reductions, constructed specifically in a way such that the dependencies on the length of the respective boundaries of regions cancel. It is not directly meant as an information theoretical quantity, although the differences of entropies resembling a mutual information expression. Also, slightly different definitions with similar properties are conceivable, and indeed, the independent proposal of ref.[148] makes use of an alternative combination of entropies. The important aspect here is the above mentioned cancellation of the boundary term. Taking the behavior as in eq. (32) for granted, one indeed finds

$$S_{\text{Topo}} = -\gamma.$$

From the way S_{Topo} is constructed it is a topological invariant, and depends only on a universal quantity (unaltered under smooth deformations, as long as one stays away from critical points), and on the fashion how the regions are located with respect to each other, but not on their specific geometry (again assuming that the correlation length is much smaller the regions and does not matter). Interestingly, topological order is hence a global property that is detected by the entanglement entropy. This construction can also readily be used in numerical studies. The explicit computation how the entanglement entropy detects the presence of topological order in an actually time-dependent model undergoing a quantum phase transition from a spin-polarized to a topologically ordered phase has been systematically explored in ref.[107], further strengthening the findings of ref.[140].

Since its proposal, this and related quantities have been considered in a number of contexts. A natural candidate to explore this concept is the *toric code state* of ref.[141]: Consider for this a square lattice $I = \{1, \ldots, n\}^{\times 2}$ with periodic boundary conditions, and place the physical two-dimensional quantum spins on the vertices of this lattice.[17] This lattice is tiled into two sublattices of different color, red and white. Every white p and red plaquette s is then associated with one of the commuting operators

$$A_s = \prod_{j \in \partial s} \sigma_j^z, \; B_p = \prod_{j \in \partial p} \sigma_j^x, \tag{34}$$

respectively, with non-trivial support on four spins each, where as before $\sigma_i^x, \sigma_i^y, \sigma_i^z$ denote the Pauli operators supported on i. The Hamiltonian of the system—a local Hamiltonian—is then taken to be

$$H = -\sum_s A_s - \sum_p B_p.$$

This is a gapped and frustration-free Hamiltonian. It is also straightforward to verify that for any closed path g the operator $\prod_{j \in g} \sigma_j^z$ commutes with all operators in eq. (34). The ground state manifold depends on the topology of the lattice and is in the chosen case four-fold degenerate. The topological entanglement entropy, evaluated for this toric code state, gives $\gamma = \ln(2)$. The ground states can readily be cast into a PEPS language, as has been done in ref.[219]. An analysis how topological order can be grasped in a language of entanglement renormalization or MERA has been performed in ref.[1]: Indeed, the topological degrees of freedom can then be distilled to the top of the tensor network.

An equally important explicit and closely related model is the *loop model on a honeycomb lattice* of ref.[141]. Ground states of more general string-net lattice models can also often be expressed in terms of remarkably simple tensor networks[33,103,104]. Entanglement entropies of *topological color codes*[24] have been studied in ref.[137]. Equivalents of the topological entanglement entropy for finite temperature—where the very robustness can be probed—have been considered and introduced in refs.[47,129]: Notably, for Gibbs states it still makes sense to consider quantities of the type as in eq. (33), only with the respective entropies being replaced by mutual informations grasping correlations instead of entanglement, as discussed in detail in Section V. It is found that the interplay between thermal effects, topological order and the size of the lattice indeed give rise to well-defined scaling relations.

---

[17] Equivalently, one can place the physical spins on the edges and formulate the operators $\{A_s\}$ and $\{B_p\}$ as being non-trivially supported on the respective four spins associated with vertices and plaquettes.

The study of entanglement entropies in fractional quantum Hall states in a spherical geometry has been initiated in ref.[108]; in ref.[243] Abelian Laughlin states as well as *Moore-Read states* have been considered, where also rigorous upper bounds for *particle entanglement entropies* have been derived. Particle partitioning entanglement in itinerant many-particle systems has been studied in ref.[244]. The MPS representation of the Laughlin wave function has been derived in ref.[128]. The tolopogical entanglement of integer quantum Hall states has been computed in ref.[193]. Topological entanglement Renyi entropies have been considered in ref.[88]. Similar quantities in *Chern-Simons theories*—the best understood topological field theories—have been identified in ref.[67]. The suggestion that the full spectrum of H in $\rho_I = \mathrm{e}^{-H}$ should be considered to detect topological order has been proposed in ref.[149]. As being certified by this list of recent developments, studies of entanglement entropies as indicators of topological order are still under rapid development.

I. Relationship to black hole entropy

As mentioned before, one of the particularly intriguing motivations for the study of area laws of the entanglement entropy is the suspected relationship to the area-dependence of the black hole entropy. The *Bekenstein-Hawking area law*[10,17,117] suggests that black hole carries an entropy that is proportional to its *horizon area* $\mathcal{A}$,

$$S_{\mathrm{BH}} = \frac{kc^3\mathcal{A}}{4G\hbar}.$$

Hence, according to this relationship, the (thermodynamical) entropy of a black hole is just a quarter of its area measured in Planck units[29], i.e., when $k = c = G = \hbar = 1$. For the sum of this black hole entropy and the matter entropy S_{Matter} a second law of thermodynamics is proposed to hold. Such a generalized second law of thermodynamics led to the suggestion that one would have a "spherical entropy bound" for matter: In asymptotically flat spacetime, any weakly gravitating matter system would satisfy $S_{\mathrm{Matter}} \leq 2\pi kEr/(\hbar c)$, interestingly not containing G. E denotes the total mass energy of the system, whereas r stands for the smallest radius of a sphere that contains the matter system at hand. The range in which one can expect the validity of such a law is discussed in ref.[29].

The linear relationship between the boundary area and the (thermodynamical) entropy—formally, the two equations look identical —suggests that one may expect a close relationship between these area laws: On the on hand, for the (von-Neumann) entanglement entropy of a subregion of a free quantum field in flat space time, on the other hand for the black hole entropy. This intriguing connection was first suggested and explored in refs.[23,207] and extended in refs.[43,122,226]. Indeed, there are physical arguments that make the reduction of the situation of having a scalar field in a static spherically symmetric space-time to a scalar field in flat space time plausible[61]. The exact status of the relationship between these quantities (or to what extent they are related by originating from a common cause—the general locality of interactions) is still subject to debate.[18]

This relationship has even been employed to take steps in computing the entanglement entropy in higher-dimensional conformal field theories: The AdS/CFT correspondence—relating a $d+2$-dimensional anti de Sitter (AdS) space to a $d+1$-dimensional conformal field theory (CFT)[3,237]—has been made use of to study the Bekenstein formula in the AdS context[195,196], see also ref.[46]. In this way, the above formula is used as a tool to compute the geometric entropy in a plausible fashion in situations where the exact computation is not known to be possible using the tools of conformal field theory.

The *holographic principle*—dating back to work in refs.[123,209]—goes even further, and suggests that generally, all information that is contained in a volume of space can be represented by information that resides on the boundary of that region. For an extensive review, see ref.[29].

V. AREA LAWS FOR CLASSICAL SYSTEMS AND FOR TOTAL CORRELATIONS

A. Classical harmonic systems

Throughout this article, we have been concerned with quantum systems on a lattice. What if we have classical systems on a lattice, could one still expect an area law to hold? Obviously, the concept of entanglement is no longer meaningful. Also, the Shannon entropy of, say, the marginal distribution of a distinguished region I would not quantify correlations in a reasonable fashion. What is worse, in case of harmonic classical systems on a lattice, when thinking in terms of phase space cells, this quantity is burdened with the usual Gibbs paradox. However, it does make perfect sense to talk about *classical correlations in classical systems*, the appropriate quantity grasping such correlations being the mutual information:

Given a probability distribution p on the lattice L, one can quantify the correlations between the marginals with respect to a distinguished region I and its complement O by means of the *mutual information*. It tells us how much information can be obtained on O from measurements in I, and equally on I by measurements in O. This quantity enjoyes a number of very natural properties. The mutual information is always positive—there can be no negative correlations—and will vanish exactly if the probability distribution factorizes, in which case one can not learn anything about O from I. Given the marginals p_I and p_O of the probability distribution p on I and O, respectively, the mutual information is defined as

$$I(I:O) = S(p_I) + S(p_O) - S(p), \tag{35}$$

[18] For a short general review on this connection see e.g., ref.[61], for a calculation of the one-loop correction to the Bekenstein-Hawking entropy in the presence of matter fields and its relationship to the geometric entropy see ref.[210], and for an entanglement-based view of the Bekenstein-Hawking entropy law see refs.[32,195].

where here $S(p) = -\sum_j p_j \log_2(p_j)$ is the standard information theoretical *Shannon entropy*. It is noteworthy that the mutual information does not suffer from the Gibbs paradoxon as will be shown below. How does the mutual information scale with the size of a region in case of a harmonic coupled classical system? The subsequent statement clarifies this situation: Consider a *classical harmonic lattice system*, with Hamiltonian

$$H = \frac{1}{2}\left(\sum_{j\in L} p_j^2 + \sum_{j,k\in L} x_j V_{j,k} x_k\right), \tag{36}$$

where now $x = (x_1, \ldots, x_N)$ and $p = (p_1, \ldots, p_N)$ are the vectors of classical position and momentum variables of classical oscillators arranged on a cubic lattice $L = \{1, \ldots, N\}^{\times \mathcal{D}}$. The phase space coordinates are then $\xi = (x, p)$. The matrix $V \in \mathbb{R}^{|L|\times|L|}$ with a finite-ranged interaction defines the interaction.

The state of the system is defined by the phase space density, so a *classical distribution* $\rho : \mathbb{R}^{N^{\mathcal{D}}} \to \mathbb{R}^+$. For any non-zero inverse temperature $\beta > 0$, this phase space distribution is nothing but

$$\rho_\beta(\xi) = \frac{1}{Z} e^{-\beta H(\xi)}, \;\; Z = \int d\xi e^{-\beta H(\xi)}.$$

To define the mutual information, following the standard procedure, we split the phase space into cubic cells each with a volume $h^{2N^{\mathcal{D}}}$, with $h > 0$ being some constant. From the phase space space density, we can then identify a discrete probability distribution, from an average of the phase space density over these cells, $p_j = \int_{\text{Cell}} d\xi \rho(\xi)$ for $j \in L$. The discrete classical entropy is then defined as the Shannon entropy of this probability distribution as

$$S_C(h) = -\sum_{j\in L} p_j \log_2(p_j).$$

We now come back to the situation of having a lattice system with an interior I and an exterior O. The respective discrete classical entropies are defined as $S_I(h)$ and $S_O(h)$. Obviously, the values of these entropies will depend on the choice of h, and in the limit $h \to 0$, they will diverge, logarithmically in h. This is a familiar observation in classical statistical physics, the divergence being resolved in the third law of thermodynamics. Here, we are, however, interested in classical correlations, as being quantified in terms of the mutual information which in the limit of $h \to 0$ is well-defined. Hence we can define the *classical mutual information of a harmonic lattice system* as $I(I : O) = \lim_{h\to 0} (S_I(h) + S_O(h) - S_C(h))$. We are now in the position to state the area theorem for classical harmonic systems[57]:

Theorem 16 (Correlations in classical harmonic systems) *Consider a harmonic lattice system with Hamiltonian as in eq. (36) on a general lattice $G = (L, E)$. Then the classical mutual information $I(I : O)$ of the Gibbs state at some inverse temperature $\beta > 0$ satisfies an area law,*

$$I(I : O) = O(s(I)).$$

The interesting aspect of this proof[57] is that it relates this question of the *classical* mutual information to a quantity that arises in the quantum case in case where the coupling matrix V_x is replaced by V_x^2, and is hence a simple corollary of earlier results on *quantum systems*, now with a coupling that is replaced by the squared coupling matrix. Hence, a "*quantum proof*" can be applied to establish a statement on classical lattice systems. The lesson to learn is that whenever one has local interactions—even in classical systems—one should not be too surprised if this manifests itself in an area law in the correlations.

B. Classical correlations quantum spin models

The situation is even simpler for finite-dimensional constituents. Indeed, quite in contrast to the overburdening technicalities that render the question of area laws in higher-dimensional quantum systems at zero temperature so difficult, the situation can here be clarified with hardly any mathematics at all: An elegant, but simple argument shows that total correlations in quantum (and classical) systems at non-zero temperatures always satisfy an area law. This is a statement on correlations—not entanglement, in contrast to the discussion of Subsection IV.D— in thermal states $\rho_\beta = e^{-\beta H}/Z$ for some $\beta > 0$ for classical or quantum systems[239]. The relevant quantity grasping correlations is again the mutual information

$$I(I : O) = S(\rho_I) + S(\rho_O) - S(\rho), \tag{37}$$

where S stands either for the von-Neumann quantum entropy, or for the Shannon entropy of the probability distribution. The classical variant was first discussed in ref.[57], the quantum version in refs.[45,46]. Ref.[45] introduces this quantity to avoid divergencies of the entanglement entropy in quantum field theory: In a similar fashion as above, regulators will in fact cancel each other, and the familiar *ultraviolet divergence in the quantum field limit* disappears.

Interestingly, a general statement on the scaling of correlations at non-zero temperature in terms of eq. (37) can be derived which holds for any spin model with local dimension d (see page 295 of ref.[30] and the subsequent ref.[239]):

Theorem 17 (Classical correlations at non-zero temperature) *Consider a classical or a quantum system with finite local dimension d defined on a translation-invariant lattice $G = (L, E)$. Consider the Gibbs state at some inverse temperature $\beta > 0$ of a local Hamiltonian H with two-site interactions. In the classical case, where each of the lattice sites corresponds to a spin with configuration space $\mathbb{Z}_d$,*

$$I(I : O) \leq |s(I)| \log(d). \tag{38}$$

For a quantum system with local Hilbert spaces $\mathbb{C}^d$, the mutual information satisfies the area law

$$I(I : O) \leq \beta \|h\| \, |s(I)|, \tag{39}$$

where $\|h\|$ is the largest eigenvalue of all Hamiltonians across the boundary of I and O.

This statement is valid in remarkable generality, given the simplicity of the argument. We will focus on quantum systems in the following. One can write the Hamiltonian H having two-site interactions as $H = H_I + H_\partial + H_O$, where H_I and H_O collect all interaction terms within the regions, whereas H_∂ stands for terms connecting the two regions. The Gibbs state ρ_β for some inverse temperature $\beta > 0$ minimizes the free energy $F(\rho) = \mathrm{tr}[H\rho] - S(\rho)/\beta$. Clearly, therefore,

$$F(\rho_\beta) \leq F(\rho_I \otimes \rho_O),$$

from which $I(I:O) \leq \beta \mathrm{tr}[H_\partial(\rho_I \otimes \rho_O - \rho_\beta)]$ is obtained. As the right hand side depends only on terms coupling the inside to the outside, i.e surface terms, eq. (39) follows straightforwardly. A naive limit $\beta \to \infty$ will not yield an area law for zero temperature, as the right hand side of eq. (39) then clearly diverges, but for any finite temperature, one obtains a bound.

VI. CONNECTION TO SIMULATABILITY

There is an intimate connection between area laws for the entanglement entropy and questions of the simulatability of quantum many-body systems. The fact that there is "little entanglement" in a system that satisfies an area law is at the core of the functioning of so powerful numerical techniques as the *density-matrix renormalization group* (DMRG) methods. To describe the large research field of numerical simulation using DMRG-type methods would be beyond the scope of the present review. Instead, we will concentrate on the direct relationship between the "effective degrees of freedom" that must be considered when classically describing quantum systems.

A. Numerical simulations with the density-matrix renormalization group method

This connection is particularly clear in one-dimensional systems, that is for quantum spin chains. Indeed, one can say that the fact that ground states of gapped systems satisfy an area law—and to a lesser extent that critical systems merely have a logarithmic divergence of the entanglement entropy—is responsible for the success of the density-matrix renormalization approach. Matrix-product states also satisfy a one-dimensional area law. As MPS are underlying the DMRG approach this suggest that the entanglement content of a state and the best possible performance of a DMRG approach can be intimately linked.

Historically, DMRG was born out of an idea of renormalization, where one iteratively identifies the relevant degrees of freedom, grasping the essential physics of the problem, when going from one step of the procedure to the next one. This general idea goes back to the *real-space renormalization group approach*, presented in ref.[235] in the mid 1970ies. This approach was particularly successful in the numerical assessment of the Kondo problem, whereas for other problems, results were not quite what was hoped for. The birth of the DMRG approach as such was related to a clear analysis of the strengths and weaknesses of the real-space renormalization group approach to study the low-energy properties of quantum many-body systems[234]. Ref.[232] is seen as the manuscript in which the DMRG method has actually been introduced. Since then, this method has seen a standard method in the numerical study of strongly correlated quantum many-body systems. For a recent review, see ref.[199].

Initially, the formulation of DMRG was based on the above renormalization idea. However, in the following years it became clear that DMRG generates matrix-product states, an insight that has been reported in ref.[173] for the thermodynamical limit of DMRG, and in ref.[68] for finite-size DMRG methods with the latter placing a particular emphasis on exploiting a rotational symmetry in variational approaches. Ref.[178] gives a relatively early exhaustive overview over variational ansatzes with matrix-product states and the relationship with the DMRG idea. Ref.[173] already hinted at the possibility for treating period boundary conditions in the MPS picture but chose translation invariant matrices. Ref.[218] relaxed this constraint to demonstrate that a suitable formulation significantly outperforms standard DMRG for periodic boundary conditions in terms of memory requirements.

Hence, DMRG—in its several variants—can be seen as a variational method, where the optimization problem

$$\begin{aligned} &\text{minimize} && \langle\psi|H|\psi\rangle, \\ &\text{subject to} && |\psi\rangle \in (\mathbb{C}^d)^{\otimes N}, \end{aligned} \tag{40}$$

impractical already because of its exponentially large feasible set, is replaced by a variant of an optimization problem over a polynomially large set

$$\begin{aligned} &\text{minimize} && \langle\psi|H|\psi\rangle, \\ &\text{subject to} && |\psi\rangle \in (\mathbb{C}^d)^{\otimes N} \text{ is an MPS vector of dimension } D. \end{aligned} \tag{41}$$

In this variant, or—more accurately—in each of these variants one does not attempt in one go to identify the global optimum, but rather effectively iteratively solves for the local matrices involved. Such an iteration will then certainly converge (albeit strictly speaking not necessarily to the global minimum).[19]

B. Approximation of states with matrix-product states

Any such method, can then only be as good as the best possible MPS can approximate the true ground state at hand. This, in fact, is related to the entanglement content, in that it matters whether or not the true ground state satisfies an area law or not. In the light of previous discussions, this connection is not that surprising any more: After all MPS satisfy an area law for the entanglement entropy. Hence, one aims at approximating ground states with states that have in this sense little entanglement, and those states can be well approximated by MPS that satisfy an area law in the first place.

This connection has been hinted at already in the first work on DMRG[232], where the spectrum of the half chain has been

[19] For mixed state simulations, see refs.[62,217].

considered and put into relationship with the *"truncation error"* in DMRG. This is a key figure of merit of the quality of an approximation in a step, so unity minus the weight of those terms being kept in a step of the iteration.

This connection between the decay of spectral values of half chains, the more rapid the decay the better can DMRG perform, has been made more precise and fleshed out in ref.[177]. In ref.[147] the relationship to criticality in this context has been emphasized. Ref.[191] is a short review on this question. In more recent quantitative approaches, the optimal approximation that can possibly be obtained by a MPS of a given D is considered. Let us denote with $\mathcal{H}_N = (\mathbb{C}^d)^{\otimes N}$ the Hilbert space of a quantum chain of length N. MPS are considered as defined in eq. (29) for open boundary conditions. Given a family $\{|\psi_N\rangle\}_N$ of state vectors, it is said that it *can be approximated efficiently by MPS* if for every $\delta > 0$ there exists a sequence $|\psi_{N,D(N)}\rangle$ of MPS with $D(N) = O(\text{poly}_\delta(N))$ such that

$$\| \, |\psi_N\rangle\langle\psi_N| - |\psi_{N,D(N)}\rangle\langle\psi_{N,D(N)}| \, \|_1 \leq \delta,$$

where $\|.\|_1$ denotes the usual trace-norm. In contrast, it is said that this sequence *cannot be approximated efficiently by MPS* if there exists some $\delta > 0$ such that no sequence of MPS with $D(N)$ growing polynomially can approximate $|\psi\rangle\langle\psi|$ up to a small error δ in trace-norm[203]:

Theorem 18 (Approximatability with MPS) *Consider sequences of state vectors $\{|\psi_N\rangle\}_N \in \mathcal{H}_N$ of a quantum chain of length N, and denote as before the reduced state of a block $I = \{1, \ldots, n\}$ of length n with ρ_I. If the sequence of ρ_I satisfies an area law for a Renyi entropy S_α for $\alpha < 1$,*

$$S_\alpha(\rho_I) = O(1),$$

then the sequence $\{|\psi_N\rangle\}_N$ is efficiently approximable by MPS. In contrast, if the von-Neumann entropy $S_1(\rho_I) = \Omega(n)$, so grows at least linearly with the block size, then it cannot be approximated efficiently by MPS. This means that states satisfying a volume law cannot be approximated. The same holds true if any Renyi entropy S_α for some $\alpha > 1$ grows at least as $S_\alpha(\rho_I) = \Omega(n^\kappa)$ for some $\kappa < 1$. Otherwise, the connection is undetermined, in that examples both for approximable and inapproximable states can be found.

This statement clarifies the connection between the entanglement content and the possibility of describing states with matrix-product states. The validity of an area law implies that there is sufficiently little entanglement in the state such that an economical description in terms of matrix-product states is possible. The enormous success of DMRG is related to the fact that gapped systems satisfy an area law. Even if the system is critical, the logarithmic divergence still allows for a relatively economical description in terms of matrix-product states. The fact that Renyi-entropies for α smaller than or larger than unity feature here may be seen rather as a technical detail. The general message is clear: The area-like entanglement scaling, with or without small corrections, allows for an efficient approximation in D for matrix-product states.

To reiterate the point made in Subsection IV.G: Quenched, non-equilibrium systems can indeed fall exactly into the category of having an effectively linearly growing block entropy, so are characterized by a volume law for the entanglement entropy. More precisely, we face the interesting situation that while for each time, we have an area law in n, the constant in the upper bound grows in time such that for a suitable choice for the sub-block, one arrives effectively at a volume law, as made precise in Theorem 15. This has severe practical implications: For small times, *t-DMRG*[58,59,142,199,221,233], the *time-dependent version of DMRG*, can very accurately keep track of the dynamics of the system. This is a variant in which one essentially makes a Lie-Trotter approximation of the time evolution operator, and then approximates in each time step the resulting state vector by an MPS, going back to ref.[221]. The functioning of this algorithm can essentially be traced back to the observation that an arbitrarily good approximation to the propagator can be established with polynomial computational resources in the system size[169]. In time, however, one will eventually encounter typically an exponential increase in the number of degrees of freedom to be kept in order to faithfully describe the state. This eventually limits the time up to which one can numerically simulate time evolution using a variant of DMRG. The increase in the entanglement content also eventually limits classical simulations of *quantum adiabatic algorithms* based on MPS, which nevertheless perform often impressively well (for a careful numerical analysis, see, e.g., ref.[9]). It is interesting to note, however, that this complexity does not necessarily translate in the difficulty of following the time-evolution of specific observables when evolving them in the *Heisenberg picture* using t-DMRG. Then, in some cases the Heisenberg time evolution can be carried out exactly for finite bond dimension and arbitrary long times[109,187,242].

There are numerical simulation methods that allow for the simulation of certain quantum states that do not satisfy an area law. MERA already allows for a logarithmic divergence of the entanglement entropy in one-dimensional systems. Weighted graphs state based approaches[6] and its 1-D variant, the *renormalization algorithm with graph enhancement*[127] can cope with instances of volume laws for the entanglement entropy, the latter in 1-D the former in arbitrary spatial dimensions. Early work on the simulation of a particular kind of discrete time evolution, namely the application of random unitary circuits, suggests that this may be a promising approach for the "efficient simulation of quantum many-body systems beyond area laws".

We end this subsection with a note rather from the computer science than from the physics perspective: The fact that a true ground state is well-approximated by an MPS does, strictly speaking, not necessarily mean that DMRG will also efficiently *find* this best approximation. In practice, DMRG works well, and it typically produces good and reasonable results. It is remarkable how well this approximation is found in the iterative scheme as being pursued by any DMRG algorithm: After all, the full problem eq. (41) is a non-convex polynomial global optimization problem of very high order ($\langle\psi|H|\psi\rangle$ is of degree N^2 in D). Still, by local variations and sweeping one achieves very good results. The ultimate

reason for this impressive performance is yet to be ultimately understood.

Having said that, the worst case complexity of the problem of finding the best approximation can be computationally difficult in the sense of computer science. In fact, the class of problem of keeping some matrices fixed and varying over a finite subset has in worst case instances that are *NP-hard*[71]. In non-translation invariant settings, one even finds that if one could efficiently identify the best possible MPS approximation, one could solve efficiently NP-hard problems[200]. Even more strongly put, the problem of approximating the ground state energy of a system composed of a chain of quantum systems is *QMA-complete*[2].

This should be seen as a warning sign: The functioning of variational algorithms such as DMRG is essentially based on heuristics, and in worst case one can encounter hard problems. The energy landscape is then so rugged that one gets stuck in local optima. Still, while it is important to acknowledge that DMRG is strictly speaking not certifiable, it is still true that it works very well in practice and is one of the pillars of the numerical assessment of strongly correlated systems in 1-D.

C. Implications on higher-dimensional simulations

For higher dimensional systems, tensor-product states or PEPS, as well as those of MERA, satisfy area laws, as has been discussed in Subsection IV.F. This fact suggests that when minimizing $\langle\psi|H|\psi\rangle$ for an $N \times N$-lattice subject to $|\psi\rangle \in (\mathbb{C}^d)^{\otimes N^2}$ being a PEPS or MERA described by polynomially many real parameters, one encounters a good approximation whenever the system at hand already satisfies an area law. In the light of the fact that even critical two-dimensional systems can satisfy an area law, this would mean that they can be well-described by PEPS or MERA described by relatively few parameters. Numerical work in case of PEPS indicates that this is indeed the case[131,160,213,216], even in the thermodynamical limit of an infinite system[136] or for fermions[14,144,150].

A rigorous result similar to Theorem 18, yet, is still lacking for PEPS or MERA. The intuition developed so far, however, is in one way or the other quite certainly right: Whenever an area law is satisfied, PEPS with small bond dimension should give rise to a reasonably good approximation. Here, subtle aspects are rather connected to the fact that the exact contraction of the tensor networks of PEPS, and hence the computation of expectation values, is inefficient, and that approximate contractions have to be employed. Suitable subsets, such as the class of *string states*, can always be efficiently contracted, giving rise to very promising variational sets in higher-dimensional systems[204]. The method in ref.[170] also gives rise to certifiable approximations of 2-D ground states for a class of models, exploiting quasi-adiabatic evolutions.

As before, one has to distinguish the variational set as such from the computational method of varying over this set. Usually, one has to find practical and heuristically suitable methods of solving a global optimization problem over many variables. Several strategies may be followed when varying over suitable sets to simulate higher-dimensional strongly correlated systems: One may use *local variations such as in DMRG*, *imaginary time evolution*, or *flow methods*[63], making use of gradient flow and optimal control ideas to vary over the manifold of unitary gates that describe the variational set of states at hand. For MERA, the same intuition should hold true. Here, first approaches implemented have been focused on one-dimensional systems[63,79,192], but the ideas are also applicable in higher dimensions[14,52,80].

VII. PERSPECTIVES

In this Colloquium, we have presented the state of affairs in the study of area laws for entanglement entropies. As has been pointed out above, this research field is presently enjoying a lot of attention, for a number of reasons and motivations. Yet, needless to say, there are numerous open questions that are to be studied, of which we mention a few to highlight further perspectives:

- Can one prove that gapped higher-dimensional general local lattice models always satisfy an area law?
- In higher dimensional systems, critical systems can both satisfy and violate an area law. What are further conditions to ensure that critical systems satisfy an area law? What is the exact role of the Fermi surface in the study of area laws in fermionic critical models?
- Can one compute scaling laws for the mutual information for quasi-free systems?
- For what 1-D models beyond quasi-free and conformal settings can one find rigorous expressions for the entanglement entropy?
- Under what precise conditions do quenched disordered local models lead to having "less entanglement"?
- What are the further perspectives of using conformal methods for systems with more than one spatial dimension?
- Can the link between the Bekenstein formula in the AdS context and the scaling of geometric entropies in conformal field theories be sharpened?
- To what extent is having a positive topological entropy and encountering topological order one to one?
- How can the relationship between satisfying an area law and the efficient approximation of ground states with PEPS be rigorously established?
- What efficiently describable states satisfy an area law, such that one can still efficiently compute local properties?
- Are there further instances for 1-D systems satisfying an area law that allow for certifiable approximations in terms of matrix-product states?

These questions only touch upon the various perspectives that open up in this context. The quantitative study of a research area that could be called "Hamiltonian complexity"[20] is just beginning to emerge. The puzzle of how complex quantum many-body systems are, and how many effective degrees of freedom are exploited by nature, is still one of the intriguing topics in the study of interacting quantum systems.

VIII. ACKNOWLEDGEMENTS

We acknowledge very fruitful and stimulating discussions on the subject of this paper with a number of friends and colleagues over the last years. Among them, we specifically mention S. Anders, K. Audenaert, H.J. Briegel, P. Calabrese, H. Casini, J.I. Cirac, C.M. Dawson, M. Fannes, S. Farkas, R. Fazio, A. Fring, D. Gross, M.J. Hartmann, M.B. Hastings, J. Keating, V. Korepin, J.I. Latorre, R. Orus, T.J. Osborne, J. Pachos, I. Peschel, J. Preskill, U. Schollwöck, N. Schuch, F. Verstraete, G. Vidal, J. Vidal, R.F. Werner, M.M. Wolf, and P. Zanardi. We specifically thank T. Barthel, P. Calabrese, D. Gottesman, M.B. Hastings, J.I. Latorre, T.J. Osborne, I. Peschel, and X.-G. Wen for valuable feedback on a draft version of this review. This work has been supported by the DFG (SPP 1116), the EU (QAP, COMPAS, CORNER, HIP), the Royal Society, the EPSRC QIP-IRC, Microsoft Research, and the EURYI Award Scheme.

IX. APPENDIX: FISHER-HARTWIG THEOREM

In this appendix we briefly present an important technical result concerning the asymptotic behavior of Toeplitz matrices[28].

Lemma 1 (Fisher-Hartwig) *Consider a sequence of $n \times n$ Toeplitz matrices $\{T_n\}_n$ with entries $(T_n)_{i,j} = (T_n)_{i-j}$,*

$$(T_n)_l = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi\, g(\phi)\mathrm{e}^{-\mathrm{i}l\phi},$$

generated by $g : [0, 2\pi) \to \mathbb{C}$. Let g be of the form

$$g(\phi) = b(\phi)\prod_{r=1}^{R} t_{\beta_r}(\phi-\phi_r)u_{\alpha_r}(\phi-\phi_r),$$

with $t_\beta(\phi) = \mathrm{e}^{-\mathrm{i}\beta(\pi-\phi)}$, $u_\alpha = (2-2\cos(\phi))^\alpha$, $Re(\alpha) > -1/2$, and $b : [0, 2\pi) \to \mathbb{C}$ a smooth non-vanishing function with winding number zero. Then[16,27,152], for $|Re(\alpha_r)| < 1/2$ and $|Re(\beta_r)| < 1/2$ or $R = 1$, $\alpha = 0$ $|Re(\beta)| < 5/2$, the asymptotic behavior of the determinant of T_n is given by

$$\lim_{n\to\infty} \frac{det(T_n)}{EG^n n^{\sum_{r=1}^R(\alpha_r^2-\beta_r^2)}} = 1$$

where $E = O(1)$ in n and

$$G = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi\, \ln(b(\phi))\right).$$

[20] This term has been coined by B.M. Terhal.

References

1. Aguado, M., and G. Vidal, 2008, Phys. Rev. Lett. **100**, 070404.
2. Aharonov, D., D. Gottesman, S. Irani, and J. Kempe, 2009, Commun. Math. Phys. **287**, 41.
3. Aharony, O., S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz, 2000, Phys. Rep. **323**, 193.
4. Amico, L., R. Fazio, A. Osterloh, and V. Vedral, 2008, Rev. Mod. Phys. **80**, 517.
5. Anders, J., and A. Winter, 2008, Quant. Inf. Comp. **8**, 0245.
6. Anders, S., M. B. Plenio, W. Dür, F. Verstraete, and H. J. Briegel, 2006, Phys. Rev. Lett. **97**, 107206.
7. Audenaert, K. M. R., J. Eisert, M. B. Plenio, and R. F. Werner, 2002, Phys. Rev. A **66**, 042327.
8. Audenaert, K. M. R., and M. B. Plenio, 2005, New J. Phys. **7**, 170.
9. Banuls, M. C., R. Orus, J. I. Latorre, A. Perez, and Ruiz-Femenia, 2006, Phys. Rev. A **73**, 022344.
10. Bardeen, J. M., B. Carter, and S. W. Hawking, 1973, Commun. Math. Phys. **31**, 161.
11. Barouch, E., and B. M. McCoy, 1971, Phys. Rev. A **3**, 786.
12. Barthel, T., M.-C. Chung, and U. Schollwöck, 2006, Phys. Rev. A **74**, 022329.
13. Barthel, T., S. Dusuel, and J. Vidal, 2006, Phys. Rev. Lett. **97**, 220402.
14. Barthel, T., C. Pineda, and J. Eisert, 2009, Phys. Rev. A **80**, 042333.
15. Barthel, T., and U. Schollwöck, 2008, Phys. Rev. Lett. **100**, 100601.
16. Basor, E. L., 1978, Trans. Amer. Math. Soc. **239**, 33.
17. Bekenstein, J. D., 1973, Phys. Rev. D **7**, 233.
18. Bekenstein, J. D., 2004, Contemp. Phys. **45**, 31.
19. Ben-Aroya, A., and A. Ta-Shma, 2007, quant-ph/0702129.
20. Bennett, C. H., D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, 1996, Phys. Rev. A **54**, 3824.
21. Benzi, M., and G. H. Golub, 1999, BIT **39**, 417.
22. Bloch, I., J. Dalibard, and W. Zwerger, 2008, Rev. Mod. Phys. **80**, 885.
23. Bombelli, L., R. K. Koul, J. Lee, and R. D. Sorkin, 1986, Phys. Rev. D **34**, 373.
24. Bombin, H., and M. A. Martin-Delgado, 2008, Phys. Rev. A **77**, 042322.
25. Botero, A., and B. Reznik, 2004, Phys. Rev. A **70**, 052329.
26. Botero, A., and B. Reznik, 2007, arXiv:0708.3391.
27. Böttcher, A., and B. Silbermann, 1985, J. Func. Anal. **63**, 178.
28. Böttcher, A., and B. Silbermann, 2006, *Analysis of Toeplitz operators* (Springer, Berlin, Heidelberg, New York), 2nd edition.
29. Bousso, R., 2002, Rev. Mod. Phys. **74**, 825.
30. Bratteli, O., and D. W. Robinson, 1979, *Operator algebras and quantum statistical mechanics*, volume II of *Texts and Monographs in Physics* (Springer).
31. Bravyi, S., M. B. Hastings, and F. Verstraete, 2006, Phys. Rev. Lett. **97**, 050401.
32. Brustein, R., M. B. Einhorn, and A. Yarom, 2006, JHEP **0601**, 098.
33. Buerschaper, O., M. Aguado, and G. Vidal, 2008, arXiv:0809.2393.

34. Burrell, C. K., J. Eisert, and T. J. Osborne, 2009, Phys. Rev. A **80**, 052319.
35. Burrell, C. K., and T. J. Osborne, 2007, Phys. Rev. Lett. **99**, 167201.
36. Calabrese, P., and J. Cardy, 2004, J. Stat. Mech. , P06002.
37. Calabrese, P., and J. Cardy, 2005, J. Stat. Mech. , P04010.
38. Calabrese, P., and J. Cardy, 2006, Int. J. Quant. Inf. **4**, 429.
39. Calabrese, P., and J. Cardy, 2006, Phys. Rev. Lett. **96**, 136801.
40. Calabrese, P., and J. Cardy, 2007, J. Stat. Mech. , P10004.
41. Calabrese, P., and J. Cardy, 2007, J. Stat. Mech. , P06008.
42. Calabrese, P., and A. Lefevre, 2008, Phys. Rev. A **78**, 032329.
43. Callan, C., and F. Wilczek, 1994, Phys. Lett. B **333**, 55.
44. Cardy, J. L., O. A. Castro-Alvaredo, and B. Doyon, 2007, J. Stat. Phys. **130**, 129.
45. Casini, H., and M. Huerta, 2004, Phys. Lett. B **600**, 142.
46. Casini, H., and M. Huerta, 2007, Nucl. Phys. B **764**, 183.
47. Castelnovo, C., and C. Chamon, 2007, Phys. Rev. B **76**, 184442.
48. Cavalcanti, D., A. Ferraro, A. Garcia-Saez, and A. Acin, 2008, Phys. Rev. A **78**, 012335.
49. Christandl, M., and A. Winter, 2004, J. Math. Phys. **45**, 829.
50. Chung, M.-C., and I. Peschel, 2000, Phys. Rev. B **62**, 4191.
51. Corboz, P., G. Evenbly, F. Verstraete, and G. Vidal, 2010, Phys. Rev. A (R) **81**, 010303.
52. Corboz, P., and G. Vidal, 2009, Phys. Rev. B **80**, 165129.
53. Cramer, M., 2006, PhD thesis, University of Potsdam.
54. Cramer, M., C. M. Dawson, J. Eisert, and T. J. Osborne, 2008, Phys. Rev. Lett. **100**, 030602.
55. Cramer, M., and J. Eisert, 2006, New J. Phys. **8**, 71.
56. Cramer, M., J. Eisert, and M. B. Plenio, 2007, Phys. Rev. Lett. **98**, 220603.
57. Cramer, M., J. Eisert, M. B. Plenio, and J. Dreissig, 2006, Phys. Rev. A **73**, 012309.
58. Cramer, M., A. Flesch, I. P. McCulloch, U. Schollwöck, and J. Eisert, 2008, Phys. Rev. Lett. **101**, 063001.
59. Daley, A. J., C. Kollath, U. Schollwöck, and G. Vidal, 2004, J. Stat. Mech. , P04005.
60. Das, S., and S. Shankaranarayanan, 2006, Phys. Rev. D **73**, 121701.
61. Das, S., S. Shankaranarayanan, and S. Sur, 2008, Phys. Rev. D **77**, 064013.
62. Datta, A., and G. Vidal, 2007, Phys. Rev. A **75**, 042310.
63. Dawson, C. M., J. Eisert, and T. J. Osborne, 2008, Phys. Rev. Lett. **100**, 130501.
64. De Chiara, G., S. Montangero, P. Calabrese, and R. Fazio, 2006, J. Stat. Mech. , P03001.
65. De Pasquale, A., and P. Facchi, 2009, Phys. Rev. A **80**, 032102.
66. Ding, L., N. Bray-Ali, R. Yu, and S. Haas, 2008, Phys. Rev. Lett. **100**, 215701.
67. Dong, S., E. Fradkin, R. G. Leigh, and S. Nowling, 2008, JHEP **5**, 016.
68. Dukelsky, J., M. A. Martin-Delgado, T. Nishino, and G. Sierra, 1998, Europhys. Lett. **43**, 457.
69. Dür, W., L. Hartmann, M. Hein, M. Lewenstein, and H. J. Briegel, 2005, Phys. Rev. Lett. **94**, 097203.
70. Eisert, J., 2001, PhD thesis, University of Potsdam.
71. Eisert, J., 2006, Phys. Rev. Lett. **97**, 260501.
72. Eisert, J., and M. Cramer, 2005, Phys. Rev. A **72**, 042112.
73. Eisert, J., and T. J. Osborne, 2006, Phys. Rev. Lett **97**, 150404.
74. Eisert, J., and M. B. Plenio, 1999, J. Mod. Opt. **46**, 145.
75. Eisert, J., and M. B. Plenio, 2003, Int. J. Quant. Inf **1**, 479.
76. Eisert, J., M. B. Plenio, J. Hartley, and S. Bose, 2004, Phys. Rev. Lett. **93**, 190402.
77. Eisler, V., D. Karevski, T. Platini, and I. Peschel, 2008, J. Stat. Mech. , P01023.
78. Eisler, V., and I. Peschel, 2007, J. Stat. Mech. , P06005.
79. Evenbly, G., and G. Vidal, 2007, arXiv:0710.0692.
80. Evenbly, G., and G. Vidal, 2009, Phys. Rev. Lett. **102**, 180406.
81. Facchi, P., G. Florio, C. Invernizzi, and S. Pascazio, 2008, arXiv:0808.0600.
82. Fagotti, M., and P. Calabrese, 2008, Phys. Rev. A (R) **78**, 010306.
83. Fannes, M., B. Haegeman, and M. Mosonyi, 2003, J. Math. Phys. **44**, 6005.
84. Fannes, M., B. Nachtergaele, and R. F. Werner, 1992, Lett. Math. Phys. **25**, 249.
85. Farkas, S., and Z. Zimboras, 2005, J. Math. Phys. **46**, 123301.
86. Farkas, S., and Z. Zimboras, 2007, J. Math. Phys. **48**, 102110.
87. Ferraro, A., D. Cavalcanti, A. Garcia-Saez, and A. Acin, 2008, Phys. Rev. Lett. **100**, 080502.
88. Flammia, S. T., A. Hamma, T. L. Hughes, and X.-G. Wen, 2009, Phys. Rev. Lett. **103**, 261601.
89. Foong, S. K., and S. Kanno, 1994, Phys. Rev. Lett. **72**, 1148.
90. Fradkin, E., and J. E. Moore, 2006, Phys. Rev. Lett. **97**, 050404.
91. Franca, V. V., and K. Capelle, 2008, Phys. Rev. A **77**, 062324.
92. Franchini, F., A. R. Its, B.-Q. Jin, and V. E. Korepin, 2007, J. Phys. A **40**, 8467.
93. Franchini, F., A. R. Its, and V. E. Korepin, 2008, J. Phys. A **41**, 025302.
94. Freedman, M. H., A. Kitaev, M. J. Larsen, and Z. Wang, 2003, Bull. Amer. Math. Soc. **40**, 31.
95. Furukawa, S., and G. Misguich, 2007, Phys. Rev. B **75**, 214407.
96. Gioev, D., and I. Klich, 2006, Phys. Rev. Lett. **96**, 100503.
97. Giovannetti, V., S. Montangero, M. Rizzi, and R. Fazio, 2009, Phys. Rev. A **79**, 052314.
98. Gottesman, D., 1997, PhD Thesis, CalTech .
99. Gottesman, D., and M. B. Hastings, 2009, arXiv:0901.1108.
100. Grimmett, G., T. J. Osborne, and P. Scudo, 2007, J. Stat. Phys. **131**, 305.
101. Gross, D., and J. Eisert, 2007, Phys. Rev. Lett. **98**, 220503.
102. Gross, D., J. Eisert, N. Schuch, and D. Perez-Garcia, 2007, Phys. Rev. A **76**, 052315.
103. Gu, Z.-C., M. Levin, and X.-G. Wen, 2008, Phys. Rev. A **78**, 205116.
104. Gu, Z.-C., M. L. B. Swingle, and X.-G. Wen, 2009, Phys. Rev. B **79**, 085118.
105. H. Casini, M. H., 2005, J. Stat. Mech. , P05012.
106. Hamma, A., R. Ionicioiu, and P. Zanardi, 2005, Phys. Rev. A **71**, 022315.
107. Hamma, A., W. Zhang, S. Haas, and D. A. Lidar, 2008, Phys. Rev. B **77**, 155111.
108. Haque, M., O. S. Zozulya, and K. Schoutens, 2007, Phys. Rev. Lett. **98**, 060401.
109. Hartmann, M., J. Prior, S. Clark, and M. B. Plenio, 2008.
110. Hastings, M. B., 2004, Phys. Rev. B **69**, 104431.
111. Hastings, M. B., 2007, J. Stat. Mech. , P08024.
112. Hastings, M. B., 2007, Phys. Rev. B **76**, 035114.
113. Hastings, M. B., 2007, Phys. Rev. B **76**, 035114.
114. Hastings, M. B., 2008, Phys. Rev. B **77**, 144302.
115. Hastings, M. B., and T. Koma, 2006, Commun. Math. Phys. **265**, 781.
116. Hastings, M. B., and S. Michalakis, 2009, arXiv:0911.4706.
117. Hawking, S. W., 1974, Nature **248**, 30.
118. Hawking, S. W., J. Maldacena, and A. Strominger, 2001, J. High Energy Phys. **05**, 001.
119. Hein, M., W. Dür, J. Eisert, R. Raussendorf, M. V. den Nest, and H.-J. Briegel, 2006, quant-ph/0602096.
120. Hein, M., J. Eisert, and H. J. Briegel, 2004, Phys. Rev. A **69**,

062311.
121. Hieida, Y., K. Okunishi, and Y. Akutsu, 1999, New J. Phys. **1**, 7.
122. Holzhey, C., F. Larsen, and F. Wilczek, 1994, Nucl. Phys. B **424**, 443.
123. 't Hooft, G., 1985, Nucl. Phys. B **256**, 727.
124. Horn, R. A., and C. R. Johnson, 1985, *Matrix analysis* (Cambridge University Press).
125. Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki, 2009, Rev. Mod. Phys. **81**, 865.
126. Hoyos, J. A., A. P. Viera, N. Laflorencie, and E. Miranda, 2007, Phys. Rev. B **76**, 174425.
127. Hübener, R., L. H. C. Kruszynska, W. Dür, F. Verstraete, J. Eisert, and M. B. Plenio, 2009, Phys. Rev. A **79**, 022317.
128. Iblisdir, S., J. I. Latorre, and R. Orus, 2007, Phys. Rev. Lett. **98**, 060402.
129. Iblisdir, S., D. Perez-Garcia, M. Aguado, and J. Pachos, 2009, Phys. Rev. B **79**, 134303.
130. Irani, S., 2009, arXiv:0901.1107.
131. Isacsson, A., and O. F. Syljuasen, 2006, Phys. Rev. E **74**, 035601.
132. Its, A. R., B.-Q. Jin, and V. E. Korepin, 2005, J. Phys. A **38**, 2975.
133. Its, A. R., B.-Q. Jin, and V. E. Korepin, 2007, Fields Institute Communications **50**, 151.
134. Jin, B.-Q., and V. E. Korepin, 2004, J. Stat. Phys. **116**, 79.
135. Jonathan, D., and M. B. Plenio, 1999, Phys. Rev. Lett. **83**, 1455.
136. Jordan, J., R. Orus, G. Vidal, F. Verstraete, and J. I. Cirac, 2008, Phys. Rev. Lett. **101**, 250602.
137. Kargarian, M., 2008, Phys. Rev. A **78**, 062312.
138. Keating, J. P., and F. Mezzadri, 2005, Phys. Rev. Lett. **94**, 2005.
139. Keyl, M., T. Matsui, D. Schlingemann, and R. F. Werner, 2006, Rev. Math. Phys. **18**, 935.
140. Kitaev, A., and J. Preskill, 2006, Phys. Rev. Lett. **96**, 110404.
141. Kitaev, A. Y., 2003, Ann. Phys. (N.Y.) **303**, 2.
142. Kollath, C., A. Läuchli, and E. Altman, 2007, Phys. Rev. Lett. **98**, 180601.
143. Korepin, V. E., 2003, cond-mat/0311056.
144. Kraus, C. V., N. Schuch, F. Verstraete, and J. I. Cirac, 2009.
145. Laflorencie, N., 2005, Phys. Rev. B (R) **72**, 140408.
146. Latorre, J. I., R. Orus, E. Rico, and J. Vidal, 2005, Phys. Rev. A **71**, 064101.
147. Latorre, J. I., E. Rico, and G. Vidal, 2004, Quant. Inf. Comp. **4**, 48.
148. Levin, M., and X.-G. Wen, 2006, Phys. Rev. Lett. **96**, 110405.
149. Li, H., and F. D. M. Haldane, 2008, Phys. Rev. Lett. **101**, 010504.
150. Li, S.-H., Q.-Q. Shi, and H.-Q. Zhou, 2010, e-print arXiv:1001.3343.
151. Li, W., L. Ding, R. Yu, T. Roscilde, and S. Haas, 2006, Phys. Rev. B **73**, 064406.
152. Libby, R., 1990, PhD thesis, University of California, Santa Cruz.
153. Lieb, E. H., and D. W. Robinson, 1972, Commun. Math. Phys. **28**, 251.
154. Lieb, E. H., T. Schultz, and D. Mattis, 1961, Ann. Phys. **16**, 407.
155. Lin, Y.-C., F. Igloi, and H. Rieger, 2007, Phys. Rev. Lett. **99**, 147202.
156. Marcovitch, S., A. Retzker, M. B. Plenio, and B. Reznik, 2009, Phys. Rev. A **80**, 012325.
157. Martin-Delgado, M. A., M. Roncaglia, and G. Sierra, 2001, Phys. Rev. B **64**, 075117.
158. Masanes, L., 2009, Phys. Rev. A **80**, 052104.
159. Montvay, I., and G. Münster, 1994, *Quantum Fields on a Lattice*, Cambridge Monographs on Mathematical Physics (Cambridge University Press).
160. Murg, V., F. Verstraete, and J. I. Cirac, 2007, Phys. Rev. A **75**, 033605.
161. Nachtergaele, B., and R. Sims, 2006, Commun. Math. Phys. **265**, 119.
162. Nachtergaele, B., and R. Sims, 2007, Commun. Math. Phys. **276**, 437.
163. Nielsen, M. A., 1999, Phys. Rev. Lett. **83**, 436.
164. Niggeman, H., A. Klümper, and J. Zittartz, 1997, Z. Phys. B **104**, 103.
165. Nussinov, Z., and G. Ortiz, 2009, Ann. Phys. **324**, 977.
166. Orus, R., 2005, Phys. Rev. A **71**, 052327.
167. Orus, R., 2008, Phys. Rev. Lett. **100**, 130502.
168. Orus, R., J. I. Latorre, J. Eisert, and M. Cramer, 2006, Phys. Rev. A (R) **73**, 060303.
169. Osborne, T. J., 2006, Phys. Rev. Lett. **97**, 157202.
170. Osborne, T. J., 2007, Phys. Rev. A **75**, 032321.
171. Osborne, T. J., and M. A. Nielsen, 2002, Phys. Rev. A **66**, 032110.
172. Osterloh, A., L. Amico, G. Falci, and R. Fazio, 2002, Nature (London) **416**, 608.
173. Östlund, S., and S. Rommer, 1995, Phys. Rev. Lett. **75**, 3537.
174. Page, D. N., 1993, Phys. Rev. Lett. **71**, 1291.
175. Papanikolaou, S., K. S. Raman, and E. Fradkin, 2007, Phys. Rev. B **76**, 224421.
176. Peschel, I., 2004, J. Stat. Mech. , P12005.
177. Peschel, I., M. Kaulke, and O. Legeza, 1999, Ann. Phys. (Leipzig) **8**, 153.
178. Peschel, I., X. Wang, M. Kaulke, and K. H. (Eds.), 1999, *Density-matrix renormalization, a new numerical method in physics*, Lecture Notes in Physics (Springer).
179. Peschel, I., and J. Zhao, 2005, J. Stat. Mech. , P11002.
180. Pfeifer, R. N. C., G. Evenbly, and G. Vidal, 2009, Phys. Rev. A (R) **79**, 040301.
181. Pineda, C., T. Barthel, and J. Eisert, 2009, e-print arXiv:0905.0669.
182. Plenio, M. B., 2005, Phys. Rev. Lett. **95**, 090503.
183. Plenio, M. B., 2007, J. Mod. Opt. **54**, 349.
184. Plenio, M. B., J. Eisert, J. Dreissig, and M. Cramer, 2005, Phys. Rev. Lett. **94**, 060503.
185. Plenio, M. B., J. Hartley, and J. Eisert, 2004, New J. Phys. **6**, 36.
186. Plenio, M. B., and S. Virmani, 2007, Quant. Inf. Comp. **7**, 1.
187. Prosen, T., and I. Pizorn, 2007, Phys. Rev. A **76**, 032316.
188. Raussendorf, R., and H. J. Briegel, 2001, Phys. Rev. Lett. **86**, 5188.
189. Refael, G., and J. E. Moore, 2004, Phys. Rev. Lett. **93**, 260602.
190. Refael, G., and J. E. Moore, 2007, cond-mat/0703038.
191. Riera, A., and J. I. Latorre, 2006, Phys. Rev. A **74**, 052326.
192. Rizzi, M., S. Montangero, and G. Vidal, 2008, Phys. Rev. A **77**, 052328.
193. Rodriguez, I. D., and G. Sierra, 2009, Phys. Rev. B **80**, 153303.
194. Rodriguez, K., S. R. Manmana, M. Rigol, R. M. Noack, and A. Muramatsu, 2006, New J. Phys. **8**, 169.
195. Ryu, S., and T. Takayanagi, 2006, J. High Energy Phys. **08**, 045.
196. Ryu, S., and T. Takayanagi, 2006, Phys. Rev. Lett. **96**, 181602.
197. Santachiara, R., 2006, J. Stat. Mech. , L06002.
198. Schlingemann, D., and R. F. Werner, 2002, Phys. Rev. A **65**, 012308.
199. Schollwöck, U., 2005, Rev. Mod. Phys. **77**, 259.
200. Schuch, N., J. I. Cirac, and F. Verstraete, 2008, Phys. Rev. Lett.

100, 250501.
201. Schuch, N., J. I. Cirac, and M. M. Wolf, 2006, Commun. Math. Phys. **267**, 65.
202. Schuch, N., M. M. Wolf, F. Verstraete, and J. I. Cirac, 2007, Phys. Rev. Lett. **98**, 140506.
203. Schuch, N., M. M. Wolf, F. Verstraete, and J. I. Cirac, 2008, Phys. Rev. Lett. **100**, 030504.
204. Schuch, N., M. M. Wolf, F. Verstraete, and J. I. Cirac, 2008, Phys. Rev. Lett. **100**, 040501.
205. Schuch, N., M. M. Wolf, K. G. H. Vollbrecht, and J. I. Cirac, 2008, New J. Phys. **10**, 033032.
206. Sen, S., 1996, Phys. Rev. Lett. **77**, 1.
207. Srednicki, M., 1993, Phys. Rev. Lett. **71**, 66.
208. Stelmachovic, P., and V. Buzek, 2004, Phys. Rev. A **70**, 032313.
209. Susskind, L., 1995, J. Math. Phys. **36**, 6377.
210. Susskind, L., and J. Uglum, 1994, Phys. Rev. D **50**, 2700.
211. Unanyan, R. G., M. Fleischhauer, and D. Bruss, 2007, Phys. Rev. A (R) **75**, 040302.
212. Vedral, V., and M. B. Plenio, 1998, Phys. Rev. A **57**, 1619.
213. Verstraete, F., and J. I. Cirac, 2004, cond-mat/0407066.
214. Verstraete, F., and J. I. Cirac, 2004, Phys. Rev. A **70**, 060302.
215. Verstraete, F., and J. I. Cirac, 2006, Phys. Rev. B **73**, 094423.
216. Verstraete, F., J. I. Cirac, and V. Murg, 2008, Adv. Phys. **57**, 143.
217. Verstraete, F., J. J. Garcia-Ripoll, and J. I. Cirac, 2004, Phys. Rev. Lett. **93**, 207204.
218. Verstraete, F., D. Porras, and J. I. Cirac, 2004, Phys. Rev. Lett. **93**, 227205.
219. Verstraete, F., M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, 2006, Phys. Rev. Lett. **96**, 220601.
220. Vidal, G., 1999, Phys. Rev. Lett. **83**, 1046.
221. Vidal, G., 2004, Phys. Rev. Lett. **93**, 040502.
222. Vidal, G., 2007, Phys. Rev. Lett. **99**, 220405.
223. Vidal, G., J. I. Latorre, E. Rico, and A. Kitaev, 2003, Phys. Rev. Lett. **90**, 227902.
224. Vidal, G., and R. F. Werner, 2002, Phys. Rev. A **65**, 032314.
225. Vidal, J., S. Dusuel, and T. Barthel, 2007, J. Stat. Mech. , P01015.
226. Viola, T. M., J. Preskill, A. Strominger, and S. P. Trivedi, 1994, Phys. Rev. D **50**, 3987.
227. Vojta, M., 2003, Rep. Prog. Phys. **66**, 2069.
228. Wen, X.-G., 1989, Phys. Rev. B **40**, 7387.
229. Wen, X.-G., 1990, Int. J. Mod. Phys. B **4**, 239.
230. Wen, X.-G., 1995, Adv. Phys. **44**, 405.
231. Weston, R., 2006, J. Stat. Mech. , L03002.
232. White, S. R., 1992, Phys. Rev. Lett. **69**, 2863.
233. White, S. R., and A. Feiguin, 2004, Phys. Rev. Lett. **93**, 076401.
234. White, S. R., and R. M. Noack, 1992, Phys. Rev. Lett. **68**, 3487.
235. Wilson, K. G., 1975, Rev. Mod. Phys. **47**, 773.
236. Witten, E., 1989, Commun. Math. Phys. **121**, 351.
237. Witten, E., 1998, Adv. Theor. Math. Phys. **2**, 253.
238. Wolf, M. M., 2006, Phys. Rev. Lett. **96**, 010404.
239. Wolf, M. M., F. Verstraete, M. B. Hastings, and J. I. Cirac, 2008, Phys. Rev. Lett. **100**, 070502.
240. Yu, R., H. Saleur, and S. Haas, 2008, Phys. Rev. B (R) **77**, 140402.
241. Zhou, H.-Q., T. Barthel, J. O. Fjaerestad, and U. Schollwöck, 2006, Phys. Rev. A (R) **74**, 050305.
242. Znidaric, M., T. Prosen, and I. Pizorn, 2008, Phys. Rev. A **78**, 022103.
243. Zozulya, O., M. Haque, K. Schoutens, and E. H. Rezayi, 2007, Phys. Rev. B **76**, 125310.
244. Zozulya, O. S., M. Haque, and K. Schoutens, 2008, Phys. Rev. A **78**, 042326.
245. Zurek, W. H., U. Dorner, and P. Zoller, 2005, Phys. Rev. Lett. **95**, 105701.
246. Zyczkowski, K., P. Horodecki, A. Sanpera, and M. Lewenstein, 1998, Phys. Rev. A **58**, 883.

Thermalization and its mechanism for generic isolated quantum systems*

Marcos Rigol,[1,2] Vanja Dunjko,[1,2] and Maxim Olshanii[2]

[1]*Department of Physics & Astronomy, University of Southern California, Los Angeles, CA 90089, USA*
[2]*Department of Physics, University of Massachusetts Boston, Boston, MA 02125, USA*

Time dynamics of isolated many-body quantum systems has long been an elusive subject. Very recently, however, meaningful experimental studies of the problem have finally become possible [1, 2], stimulating theoretical interest as well [3, 4, 5, 6, 7]. Progress in this field is perhaps most urgently needed in the foundations of quantum statistical mechanics. This is so because in generic isolated systems, one expects [8, 9] nonequilibrium dynamics on its own to result in thermalization: a relaxation to states where the values of macroscopic quantities are stationary, universal with respect to widely differing initial conditions, and predictable through the time-tested recipe of statistical mechanics. However, it is not obvious what feature of many-body quantum mechanics makes quantum thermalization possible, in a sense analogous to that in which dynamical chaos makes classical thermalization possible [10]. For example, dynamical chaos itself cannot occur in an isolated quantum system, where time evolution is linear and the spectrum is discrete [11]. Underscoring that new rules could apply in this case, some recent studies even suggested that statistical mechanics may give wrong predictions for the outcomes of relaxation in such systems [4, 5]. Here we demonstrate that an isolated generic quantum many-body system does in fact relax to a state well-described by the standard statistical mechanical prescription. Moreover, we show that time evolution itself plays a merely auxiliary role in relaxation and that thermalization happens instead at the level of individual eigenstates, as first proposed by J. M. Deutsch [12] and M. Srednicki [13]. A striking consequence of this *eigenstate thermalization* scenario, confirmed below for our system, is that the knowledge of a *single* many-body eigenstate suffices to compute thermal averages—any eigenstate in the microcanonical energy window will do, as they all give the same result.

If we pierce an inflated balloon inside a vacuum chamber, very soon we find the released air uniformly filling the enclosure and the air molecules attaining the Maxwell velocity distribution whose width depends only on their total number and energy. Different balloon shapes, placements, or piercing points all lead to the same spatial and velocity distributions. Classical physics explains this *thermodynamical universality* as follows [10]: almost all particle trajectories quickly begin looking alike, even if their initial points are very different, because nonlinear equations drive them to explore ergodically the constant-energy manifold, covering it uniformly with respect to precisely the microcanonical measure. However, if the system possesses further conserved quantities *functionally independent* from the Hamiltonian and each other, then time evolution is confined to a highly restricted hypersurface of the energy manifold. Hence, microcanonical predictions fail and the system does not thermalize.

On the other hand, in isolated quantum systems not only is dynamical chaos absent due to the linearity of time evolution and the discreteness of spectra [11], but it is also not clear under what conditions conserved quantities provide independent constraints on relaxation dynamics. To begin with, any operator commuting with a generic and thus nondegenerate Hamiltonian is functionally dependent on it [14], implying that the conservation of energy is the only independent constrain. On the other hand, even when operators are functionally dependent, their expectation values—considered as functionals of states—generally are not: for example, two states may have the same mean energies but different mean square-energies. For nondegenerate Hamiltonians a maximal set of constants of motion with functionally independent expectation values is as large as the dimension of the Hilbert space; examples include the projectors $\widehat{P}_\alpha = |\Psi_\alpha\rangle\langle\Psi_\alpha|$ to the energy eigenstates [14] and the integer powers of the Hamiltonian [5].

The current numerical and analytic evidence from integrable systems suggests that there exists a minimal set of independent constraints whose size is much smaller than the dimension of the Hilbert space but may still be much greater than one. In our previous work [3] (with V. Yurovsky) we showed that an integrable isolated one-dimensional system of lattice hard-core bosons relaxes to an equilibrium characterized not by the usual but by a *generalized* Gibbs ensemble. Instead of just the energy, the Gibbs exponent contained a linear combination of conserved quantities—the occupations of the eigenstates of the corresponding Jordan-Wigner fermions—whose number was still only a tiny fraction of the dimension of the Hilbert space. Yet this ensemble works, while the usual one does not, for a wide variety of initial conditions [15] as well as for a fermionic system [16]; it also explains a recent experimental result, the absence of thermalization in the Tonks-Girardeau gas [1]. Thus, while at least some constraints beyond the conservation of energy must be kept, it turns out one needs only a relatively limited number of additional conserved quantities with functionally independent expectation values; adding still further ones is redundant.

Since it is not clear which sets of conserved quantities—and some are always present—constrain relaxation and which do not, it becomes even more urgent to determine whether isolated generic quantum systems relax to the usual thermal state. The theoretical attention to this question has in fact been increasing recently, because of the high levels of iso-

*Published in Nature **452**, 854-858 (17 April 2008); 10.1038/nature06838. http://www.nature.com/nature/journal/v452/n7189/abs/nature06838.html

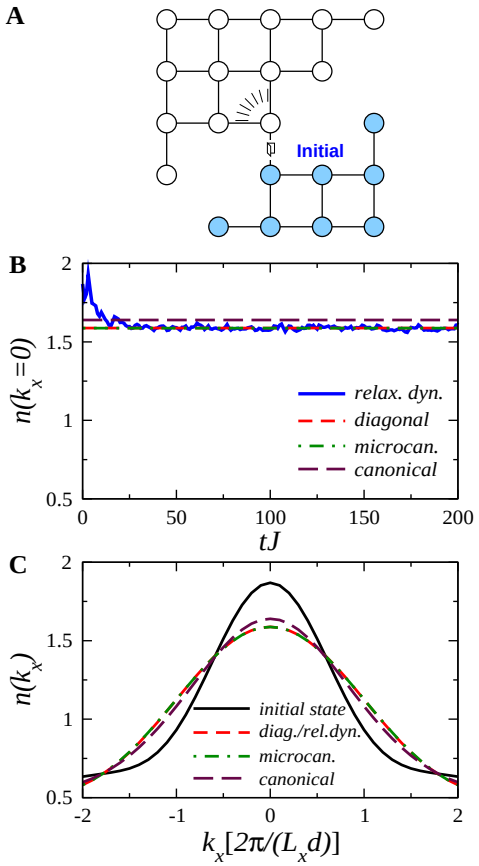


FIG. 1: **Relaxation dynamics**. **a**, Two-dimensional lattice on which five hard-core bosons propagate in time. The bosons are initially prepared in the ground state of the sub-lattice in the lower-right corner and released through the indicated link. **b**, The corresponding relaxation dynamics of the marginal momentum distribution center $[n(k_x = 0)]$ compared with the predictions of the three ensembles. In the microcanonical case, we averaged over all eigenstates whose energies lie within a narrow window (see Supplementary Discussion) $[E_0 - \Delta E, E_0 + \Delta E]$, where $E_0 \equiv \langle\psi(0)|\widehat{H}|\psi(0)\rangle = -5.06J$, $\Delta E = 0.1J$, and J is the hopping parameter. The canonical ensemble temperature is $k_{\rm B}T = 1.87J$, where $k_{\rm B}$ is the Boltzmann constant, so that the ensemble prediction for the energy is E_0. **c**, Full momentum distribution function in the initial state, after relaxation, and in the different ensembles. Here d is the lattice constant and $L_x = 5$ the lattice width.

lation [1, 2, 17] and control [18, 19] possible in experiments with ultracold quantum gases. However, despite numerous studies of specific models there is not yet consensus on how or even if relaxation to the usual thermal values occurs for nonintegrable systems [7]. Common wisdom says that it does [8, 9], but some recent numerical results suggest otherwise, either under certain conditions [4] or in general [5].

In order to study relaxation of an isolated quantum systems, we considered the time evolution of five hard-core bosons with additional weak nearest-neighbor repulsions, on a 21-site two-dimensional lattice, initially confined to a portion of the lattice and prepared in their ground state there. Figure 1a shows the exact geometry (see also Supplementary Discussion); the relaxation dynamics begins when the confinement is lifted. Expanding the initial state wavefunction in the eigenstate basis of the final Hamiltonian $\widehat{H}$ as $|\psi(0)\rangle = \sum_\alpha C_\alpha|\Psi_\alpha\rangle$, the many-body wavefunction evolves as $|\psi(t)\rangle = e^{-i\widehat{H}t}|\psi(0)\rangle = \sum_\alpha C_\alpha e^{-iE_\alpha t}|\Psi_\alpha\rangle$, where the E_α's are the eigenstate energies. Thus obtaining numerically-exact results for all times required the full diagonalization of the 20,349-dimensional Hamiltonian. The quantum-mechanical mean of any observable $\widehat{A}$ evolves as

$$\langle\widehat{A}(t)\rangle \equiv \langle\psi(t)|\widehat{A}|\psi(t)\rangle = \sum_{\alpha,\,\beta} C^\star_\alpha C_\beta e^{i(E_\alpha - E_\beta)t} A_{\alpha\beta}\,, \quad (1)$$

with $A_{\alpha\beta} = \langle\Psi_{\alpha'}|\widehat{A}|\Psi_\alpha\rangle$. The long-time average of $\langle\widehat{A}(t)\rangle$ is then

$$\overline{\langle\widehat{A}\rangle} = \sum_\alpha |C_\alpha|^2 A_{\alpha\alpha}\,. \quad (2)$$

Note that if the system relaxes at all, it must be to this value. We find it convenient to think of Eq. (2) as stating the prediction of a "diagonal ensemble," $|C_\alpha|^2$ corresponding to the weight $|\Psi_\alpha\rangle$ has in the ensemble. In fact, this ensemble is precisely the generalized Gibbs ensemble introduced in Ref. [3] if as integrals of motion one takes all the projection operators $\widehat{P}_\alpha = |\Psi_\alpha\rangle\langle\Psi_\alpha|$. Using these as constraints on relaxation dynamics, the theory gives $\hat{\rho}_c = \exp\left(-\sum_{\alpha=1}^{D}\lambda_\alpha\hat{P}_\alpha\right)$, with $\lambda_\alpha = -\ln(|C_\alpha|^2)$, and D the dimension of the Hilbert space. (Notice, however, that for the integrable system treated in Ref. [3], the generalized Gibbs ensemble was defined using a different, *minimal* set of independent integrals of motion, whose number was equal to the number of lattice sites $N \ll D$.)

Now if the quantum-mechanical mean of an observable behaves thermally it should settle to the prediction of an appropriate statistical-mechanical ensemble. For our numerical experiments we chose to monitor the marginal momentum distribution along the horizontal axis $n(k_x)$ and its central component $n(k_x = 0)$ (see Supplementary Discussion). Figures 1b and 1c demonstrate that both relax to their microcanonical predictions. The diagonal ensemble predictions are indistinct from these, but the canonical ones, although quite close, are not. This is an indication of the relevance of finite size effects, which may be the origin of some of the apparent deviations from thermodynamics seen in the recent numerical studies of Refs. [4] and [5].

The statement that the diagonal and microcanonical ensembles give the same predictions for the relaxed value of $\widehat{A}$ reads

$$\sum_\alpha |C_\alpha|^2 A_{\alpha\alpha} = \langle A\rangle_{\rm microcan.}(E_0) \stackrel{\rm def.}{=} \frac{1}{\mathcal{N}_{E_0,\,\Delta E}} \sum_{\substack{\alpha \\ |E_0 - E_\alpha| < \Delta E}} A_{\alpha\alpha}\,, \quad (3)$$

where E_0 is the mean energy of the initial state, ΔE is the half-width of an appropriately chosen (see Supplementary

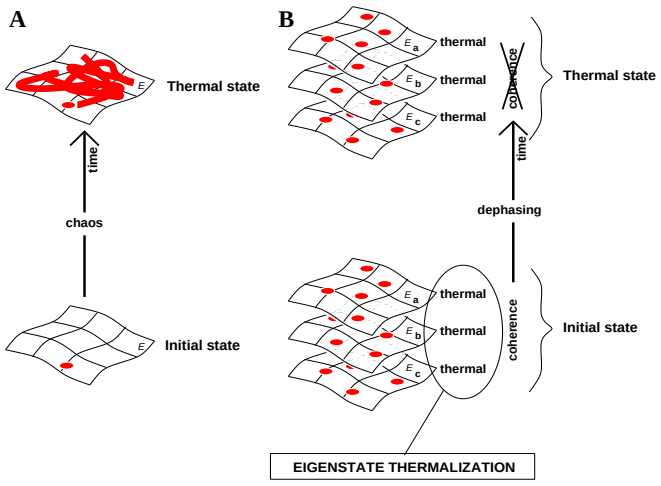


FIG. 2: **Thermalization in classical vs quantum mechanics**. **a**, In classical mechanics, time evolution constructs the thermal state from an initial state that generally bears no resemblance to the former. **b**, In quantum mechanics, according to the eigenstate thermalization hypothesis, every eigenstate of the Hamiltonian always implicitly contains a thermal state. The coherence between the eigenstates initially hides it, but time dynamics reveals it through dephasing.

Discussion) energy window centered at E_0, and the normalization $\mathcal{N}_{E_0,\,\Delta E}$ is the number of energy eigenstates with energies in the window $[E_0 - \Delta E,\, E_0 + \Delta E]$. Thermodynamical universality is evident in this equality: while the left hand side depends on the details of the initial conditions through the set of coefficients C_α, the right hand side depends only on the total energy, which is the same for many different initial conditions. Three mechanisms suggest themselves as possible explanations of this universality (assuming the initial state is sufficiently narrow in energy, as is normally the case—see Supplementary Discussion):

(i) Even for eigenstates close in energy, there are large eigenstate-to-eigenstate fluctuations of both the eigenstate expectation values $A_{\alpha\alpha}$ and of the eigenstate occupation numbers $|C_\alpha|^2$. However, for physically interesting initial conditions, the fluctuations in the two quantities are uncorrelated. A given initial state then performs an unbiased sampling of the distribution of the eigenstate expectation values $A_{\alpha\alpha}$, resulting in Eq. (3).

(ii) For physically interesting initial conditions, the eigenstate occupation numbers $|C_\alpha|^2$ practically do not fluctuate at all between eigenstates that are close in energy. Again, Eq. (3) immediately follows.

(iii) The eigenstate expectation values $A_{\alpha\alpha}$ practically do not fluctuate at all between eigenstates that are close in energy. In that case Eq. (3) holds for literally *all* initial states narrow in energy.

J. M. Deutsch and M. Srednicki have independently proposed the scenario (iii), dubbed the

> Eigenstate thermalization hypothesis (ETH)
> [Deutsch[12] (1991), Srednicki[13] (1994)].
> The expectation value $\langle\Psi_\alpha|\widehat{A}|\Psi_\alpha\rangle$ of a few-body observable $\widehat{A}$ in an eigenstate of the Hamiltonian $|\Psi_\alpha\rangle$, with energy E_α, of a large interacting many-body system equals the thermal (microcanonical in our case) average $\langle A\rangle_{\text{microcan.}}(E_\alpha)$ of $\widehat{A}$ at the mean energy E_α:
>
> $$\langle\Psi_\alpha|\widehat{A}|\Psi_\alpha\rangle = \langle A\rangle_{\text{microcan.}}(E_\alpha). \qquad (4)$$

The ETH suggests that classical and quantum thermal states have very different natures, as depicted in Fig. 2. While at present there are no general theoretical arguments supporting the ETH, some results do exist for restricted classes of systems. To begin with, Deutsch [12] showed that the ETH holds in the case of an integrable Hamiltonian weakly perturbed by a single matrix taken from a random Gaussian ensemble. Next, nuclear shell model calculations have shown that individual wavefunctions reproduce thermodynamic predictions [20]. Then there are rigorous proofs that some particular quantum systems, whose classical counterparts are chaotic, satisfy the ETH in the semiclassical limit [21, 22, 23, 24]. More generally, for low density billiards in the semiclassical regime, the ETH follows from Berry's conjecture [13, 25], which in turn is believed to hold in semiclassical classically-chaotic systems [26]. Finally, at the other end of the chaos-integrability continuum, in systems solvable by Bethe ansatz, observables are smooth functions of the integrals of motion. This allows the construction of single energy eigenstates that reproduce thermal predictions [27].

In Figs. 3a-c we demonstrate that the ETH *is* in fact the mechanism responsible for thermal behavior in our nonintegrable system. Fig. 3c additionally shows that scenario (ii) mentioned above plays no role, because the fluctuations in the eigenstate occupation numbers $|C_\alpha|^2$ are large. Thermal behavior also requires that both the diagonal and the chosen thermal ensemble have sufficiently narrow energy distributions $\rho(E)$ [= probability distribution $\times$ the density of states], so that in the energy region where the energy distributions $\rho(E)$ are appreciable, the derivative of the curve eigenstate expectation value $A_{\alpha\alpha}$ vs the energy (here $n(k_x = 0)$ vs the energy) does not change much; see Supplementary Discussion. As shown in Fig. 3b, this holds for the microcanonical and diagonal ensembles but not for the canonical ensemble, explaining the failure of the latter to describe the relaxation in Fig. 1. Note that the fluctuations of the eigenstate occupation numbers $|C_\alpha|^2$ in Fig. 3b are lowered by the averaging involved in the computation of the density of states (compare with Fig. 3c).

To strengthen the case for the ETH, we tested another observable. We chose it with the following consideration in mind: in our system interactions are local in space, and momentum distribution is a global, approximately spatially additive property. Thus one might wonder if the ETH for momentum distribution arises through some spatial averaging mechanism (we thank the anonymous referee 2 for bringing this point to our attention). It does not: for our final test of the ETH we chose an observable that is manifestly local in space, the expectation value of the occupation of the central site of the

lattice. We again found that the ETH holds true (3% relative standard deviation of eigenstate-to-eigenstate fluctuations).

On the other hand, Figs. 3d-f show how the ETH *fails* for an isolated one-dimensional *integrable* system. The latter consists of five hard-core bosons initially prepared in their ground state in an 8-site chain, one of the ends of which we then link to one of the ends of an adjoining (empty) 13-site chain to trigger relaxation dynamics. As Fig. 3e shows, $n(k_x)$ as a function of energy is a broad cloud of points, meaning that the ETH is not valid; Fig. 3f shows that scenario (ii) does not hold either.

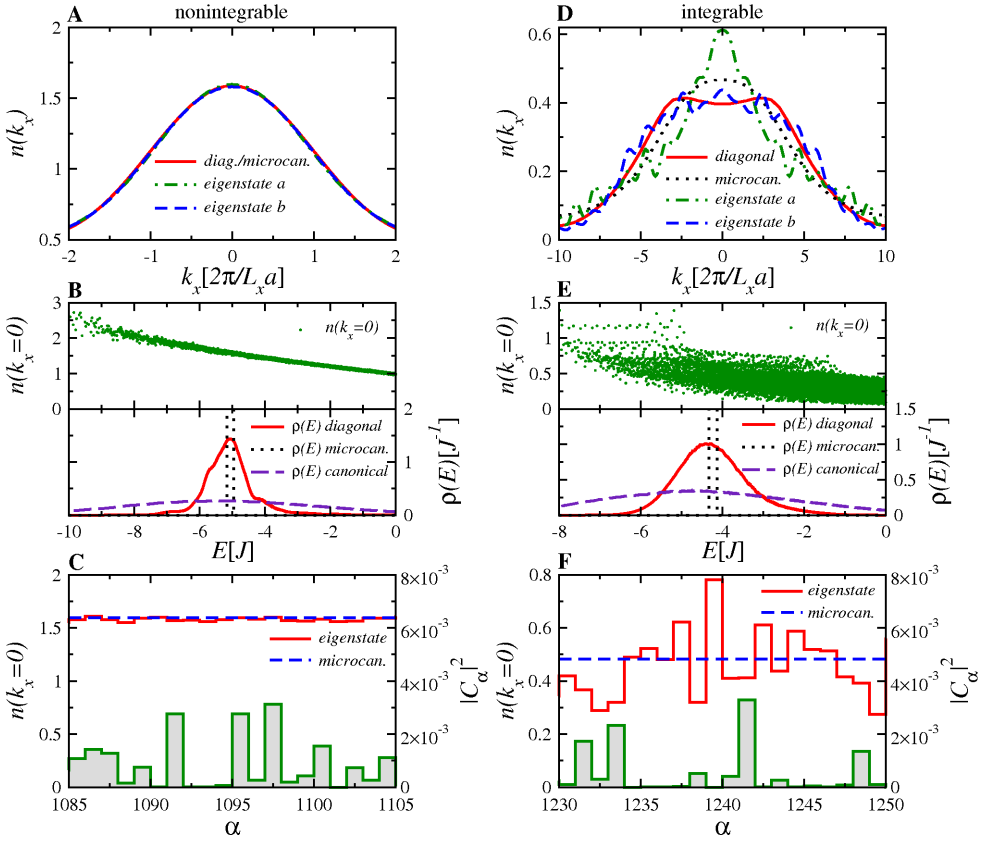


FIG. 3: **Eigenstate thermalization hypothesis**. **a**, In our nonintegrable system, the momentum distribution $n(k_x)$ for two typical eigenstates with energies close to E_0 is identical to the microcanonical result, in accordance with the ETH. **b**, Upper panel: $n(k_x = 0)$ eigenstate expectation values as a function of the eigenstate energy resemble a smooth curve. Lower panel: the energy distribution $\rho(E)$ of the three ensembles considered in this work. **c**, Detailed view of $n(k_x = 0)$ (left labels) and $|C_\alpha|^2$ (right labels) for 20 eigenstates around E_0. **d**, In the integrable system, $n(k_x)$ for two eigenstates with energies close to E_0 and for the microcanonical and diagonal ensembles are very different from each other, i.e., the ETH fails. **e**, Upper panel: $n(k_x = 0)$ eigenstate expectation value considered as a function of the eigenstate energy gives a thick cloud of points rather than resembling a smooth curve. Lower panel: energy distributions in the integrable system are similar to the nonintegrable ones depicted in **b**. **f**, Correlation between $n(k_x = 0)$ and $|C_\alpha|^2$ for 20 eigenstates around E_0. It explains why in **d** the microcanonical prediction for $n(k_x = 0)$ is larger than the diagonal one.

Nevertheless, one may still wonder if in this case scenario (i) might hold—if the averages over the diagonal and the microcanonical energy distributions shown in Fig. 3e might agree. Figure 3d shows that this does not happen. This is so because, as shown in Fig. 3f, the values of $n(k_x = 0)$ for the most-occupied states in the diagonal ensemble (the largest values of eigenstate occupation numbers $|C_\alpha|^2$) are always smaller than the microcanonical prediction, and those of the least-occupied states, always larger. Hence, the usual thermal predictions fail because the correlations between the values of $n(k_x = 0)$ and $|C_\alpha|^2$ preclude unbiased sampling of the latter by the former. These correlations have their origin in the nontrivial integrals of motion that make the system integrable and that enter the *generalized* Gibbs ensemble, which was introduced in Ref. [3] as appropriate for formulating statistical mechanics of isolated integrable systems. In the nonintegrable case shown in Fig. 3c, $n(k_x = 0)$ is so narrowly distributed that it does not matter whether or not it is correlated with $|C_\alpha|^2$ (we have in fact seen no correlations in the nonintegrable case).

The thermalization mechanism outlined thus far explains why long-time averages converge to their thermal predictions. A striking aspect of Fig. 1b, however, is that the time fluctuations are so small that after relaxation the thermal prediction works well at every instant of time. Looking at Eq. (1), one might think this is so because the contribution of the off-

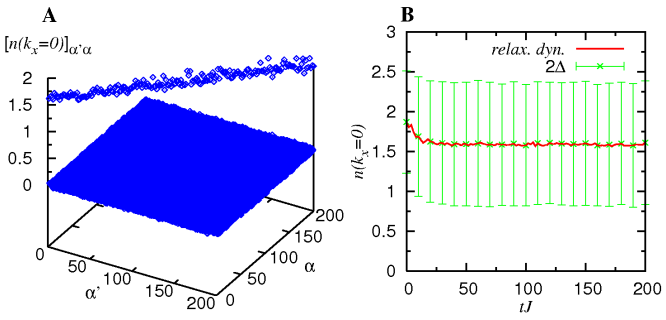


FIG. 4: **Temporal vs quantum fluctuations**. **a**, Matrix elements of the observable of interest, $n(k_x = 0)$, as a function of state indices; the eigenstates of the Hamiltonian are indexed in the order of diminishing overlap with the initial state. The dominance of the diagonal matrix elements is apparent. **b**, The same time evolution as in Fig. 1b with the error bars showing the quantum fluctuations $n(k_x = 0) \pm \Delta$ with $\Delta = [\langle \widehat{n}^2(k_x = 0)\rangle - \langle \widehat{n}(k_x = 0)\rangle^2]^{1/2}$, which are clearly much larger than the temporal fluctuations of $n(k_x = 0)$.

diagonal terms gets attenuated by temporal dephasing, which results from the generic incommensurability of the frequencies of the oscillating exponentials. However, this attenuation only scales as the root of the number of dephasing terms, and is exactly compensated by their larger number: if the number of eigenstates that have a significant overlap with the initial state is N_{states}, then typical $C_\alpha \sim 1/\sqrt{N_{\text{states}}}$, and the sum over off-diagonal terms in Eq. (1) finally does not scale down with N_{states}:

$$\sum_{\substack{\alpha,\,\beta \\ \alpha\neq\beta}} \frac{e^{i(E_\alpha - E_\beta)t}}{N_{\text{states}}} A_{\alpha\beta} \sim \frac{\sqrt{N^2_{\text{states}}}}{N_{\text{states}}} A^{\text{typical}}_{\alpha\beta} \sim A^{\text{typical}}_{\alpha\beta} \qquad (5)$$

Hence, were the magnitude of the diagonal and off-diagonal terms comparable, their contributions would be comparable as well, and time fluctuations of the average would be of the order of the average. However, this is not the case and thus

$$A^{\text{typical}}_{\substack{\alpha\beta \\ \alpha\neq\beta}} \ll A^{\text{typical}}_{\alpha\alpha}. \qquad (6)$$

Figure 4a confirms this inequality for the matrix elements of the momentum distribution in our system. We should mention that there is an *a priori* argument—admittedly in part dependent on certain hypotheses about chaos in quantum billiards—in support of this inequality for the case when the mean value of $\widehat{A}$ in an energy eigenstate is comparable to the quantum fluctuation of $\widehat{A}$ in that state [28].

On the other hand, the thermalization we see appears to be working a bit *too well*: in a system as small as ours, one would expect measurement-to-measurement fluctuations to be much larger than what Fig. 1b suggests. Indeed, as we show in Figure 4b, the fluctuations that one would actually measure would be dominated by the quantum fluctuations of the time-dependent state. The rather large size of the quantum fluctuations relative to the thermal mean value is of course particular to small systems; however, the dominance of the quantum fluctuations over the temporal fluctuations of quantum expectation values is not and is actually expected for generic systems in the thermodynamic limit [29].

We have demonstrated that, in contrast to the integrable case, the nonequilibrium dynamics of a generic isolated quantum system does lead to standard thermalization. We verified that this happens through the eigenstate thermalization mechanism, a scenario J. M. Deutsch [12] demonstrated for the case of an integrable quantum Hamiltonian weakly perturbed by a single matrix taken from a random Gaussian ensemble and M. Srednicki [13] compellingly defended for the case of rarefied semiclassical quantum billiards, and which both authors conjectured to be valid in general. Our results, when combined with the others we mentioned [12, 13, 20, 21, 22, 23, 24, 25, 26, 27], constitute strong evidence that eigenstate thermalization indeed generally underlies thermal relaxation in isolated quantum systems. Therefore, to understand the existence of universal thermal time-asymptotic states, one should study operator expectation values in individual eigenstates. This is a problem that is linear, time-independent, and conceptually far simpler than any arising in the research—currently dominating the field—on the nonlinear dynamics of semiclassical systems. Among the fundamental open problems of statistical mechanics that could benefit from the linear time-independent perspective are the nature of irreversibility, the existence of a KAM-like threshold [30] in quantum systems, and the role of conserved quantities in the approach to equilibrium. Finally, having a clear conceptual picture for the origins of thermalization may make it possible to engineer new, "unthermalizable" states of matter [12], with further applications in quantum information and precision measurement.

We thank A. C. Cassidy, K. Jacobs, A. P. Young, and E. J. Heller for helpful comments. We acknowledge financial support from the National Science Foundation and the Office of Naval Research. We are grateful to the USC HPCC center where our numerical computations were performed.

SUPPLEMENTARY DISCUSSION

1. The Hamiltonian and the numerical calculations.

In a system of units where $\hbar = 1$ the Hamiltonian reads

$$\widehat{H} = -J \sum_{\langle i,j\rangle} \left(\hat{b}^\dagger_i \hat{b}_j + \text{h.c.}\right) + U \sum_{\langle i,j\rangle} \hat{n}_i \hat{n}_j \qquad (7)$$

where $\langle i, j\rangle$ indicates that the sums run over all nearest-neighbor pairs of sites, J is the hopping parameter, and U the nearest-neighbor repulsion parameter that we always set to $0.1J$. The hard-core boson creation ($\hat{b}^\dagger_i$) and annihilation ($\hat{b}_j$) operators commute on different sites, $[\hat{b}_i, \hat{b}^\dagger_j] = [\hat{b}_i, \hat{b}_j] = [\hat{b}^\dagger_i, \hat{b}^\dagger_j] = 0$ for all i and $j \neq i$, while the hard-core condition imposes the canonical anticommutation relations on the

same site, $\{\hat{b}_i,\, \hat{b}_i^\dagger\} = 1$, and $(\hat{b}_i)^2 = (\hat{b}_i^\dagger)^2 = 0$ for all i. Here $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ is the density operator.

An exact study of the nonequilibrium dynamics for *all* time scales requires a full diagonalization of the many-body Hamiltonian (7). We are able to fully diagonalize—essentially to machine precision—matrices of dimension $D \sim 20,000$, and so we consider $N = 5$ hard-core bosons on a 5×5 lattice with four sites missing ($D = 20,349$); see Fig. 5. All the eigenstates of the Hamiltonian are used for the time evolution

$$|\psi(t)\rangle = \exp\left[-i\widehat{H}t\right]|\psi(0)\rangle = \sum_\alpha C_\alpha \exp\left[-iE_\alpha t\right]|\Psi_\alpha\rangle\,,$$

where $|\psi(t)\rangle$ is the time-evolving state, $|\psi(0)\rangle$ the initial state, $|\Psi_\alpha\rangle$ the eigenstates of the Hamiltonian with the energies E_α, and $C_\alpha = \langle\Psi_\alpha|\psi(0)\rangle$. Our initial state is the ground state of the five bosons when they are confined to the lower part of the lattice (the colored part in Fig. 5. The time evolution begins with the opening of the link shown in Fig. 5, which allows the bosons to expand over the whole lattice. The position of the missing sites was chosen so that we only open a single link to start the relaxation dynamics. The motivation for this will become apparent in the last paragraph below.

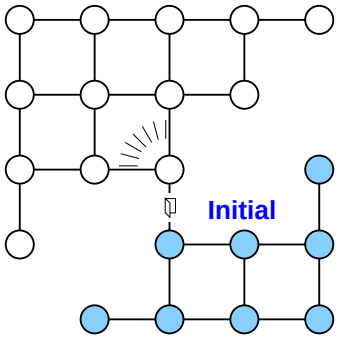


FIG. 5: **The lattice for the dynamics**. Two-dimensional lattice on which the particles propagate in time. The initial state is the ground state of 5 hard-core bosons confined to the sub-lattice in the lower right-hand corner, and the time evolution starts after the opening of the link indicated by the door symbol.

As the principal observables of interest we chose the marginal momentum distribution along the horizontal axis $n(k_x) = \sum_{k_y} n(k_x, k_y)$ and in particular its central component $n(k_x = 0)$, quantities readily measurable in actual experiments with ultracold quantum gases [19]. Here the full two-dimensional momentum distribution is $n(k_x, k_y) = 1/L^2 \sum_{i,j} e^{-i2\pi\mathbf{k}(\mathbf{r}_i - \mathbf{r}_j)/L} \langle \hat{b}_i^\dagger \hat{b}_j \rangle$, where $L = L_x = L_y = 5$ are the linear sizes of the lattice. The position $\mathbf{r}_i = (i_x\, d\,, i_y\, d)$ involves the lattice constant d.

2. The microcanonical ensemble in a small system.

To compute the microcanonical ensemble predictions, we have averaged over all eigenstates whose energies lie within a narrow window $[E_0 - \Delta E, E_0 + \Delta E]$, with $E_0 \equiv \langle\psi(0)|\widehat{H}|\psi(0)\rangle = \langle\psi(t)|\widehat{H}|\psi(t)\rangle = -5.06J$. Since our systems are small there is generally no meaning to the limit $\Delta E \to 0$, because small enough windows may fail to contain even a single eigenstate. Instead, one should show that the microcanonical predictions are robust with respect to the choice of the width of the energy window. In Fig. 6 we demonstrate this robustness in a neighborhood of $\Delta E = 0.1J$, a value that seems to be an appropriate choice given the data presented in the inset of the same figure. There we show the dependence on ΔE of the predictions for $n(k_x = 0)$ given by the "left-averaged" and the "right-averaged" microcanonical ensembles, by which we mean that the microcanonical windows are chosen as $[E_0 - \Delta E, E_0]$ and $[E_0, E_0 + \Delta E]$, respectively. We see that for $\Delta E \lesssim 0.1J$, the two microcanonical predictions are almost independent of the value of ΔE. The main panel in Fig. 6 shows that the microcanonical values of $n(k_x)$ for $\Delta E = 0.05J$ and for $\Delta E = 0.1J$ are indistinguishable.

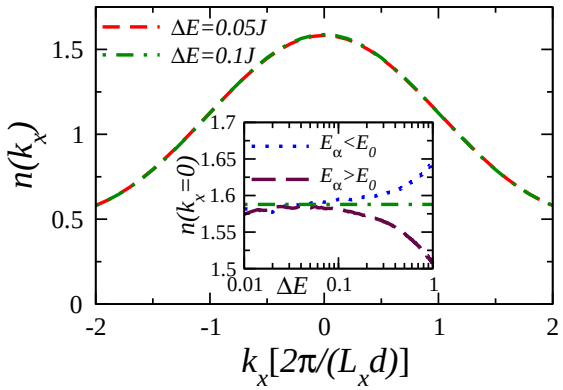


FIG. 6: **Microcanonical ensemble**. Microcanonical momentum distribution function for two different values of ΔE. Inset: Microcanonical predictions for $n(k_x = 0)$ calculated using the left ($[E_0 - \Delta E, E_0]$) and the right ($[E_0, E_0 + \Delta E]$) averages as functions of ΔE.

3. Eigenstate thermalization and the width of the energy distribution.

The eigenstate thermalization alone is not sufficient to ensure an agreement between the predictions of the diagonal and thermal ensembles. As discussed in Ref. [13], it is also necessary that both distributions be sufficiently narrow. More specifically, one must require for both ensembles

$$(\Delta E)^2\, |A''(E)/A(E)| \ll 1, \tag{8}$$

where ΔE is the width of the energy distribution in the ensemble, and $A(E)$ is the dependence of the expectation value of the observable $A_{\alpha\alpha} = \langle\Psi_\alpha|\widehat{A}|\Psi_\alpha\rangle$ on the energy E_α of the Hamiltonian-operator eigenstate $|\Psi_\alpha\rangle$. Note that because of

eigenstate thermalization, $A(E)$ is a smooth function of energy. For the thermodynamical ensembles the condition (8) is always satisfied in the thermodynamic limit. We now show that it is also satisfied for the diagonal ensemble in the thermodynamic limit.

If one considers an observable a that is the intensive counterpart of A, all conclusions obtained for a can be extended to the original observable A via trivial rescaling. For example, for our principal observable of interest, $n(k_x)$, the corresponding intensive variable is the momentum density $\xi(p_x)$ normalized as $\int dp_x\, \xi(p_x) = 1$. Notice that in this case $\xi(p_x) = n(k_x) L_x d/(2\pi N)$.

For a, the condition in (8) reduces to

$$(\Delta\epsilon)^2 |a''(\epsilon)/a(\epsilon)| \ll 1, \tag{9}$$

where $\epsilon \equiv E/N$. For sufficiently large systems the dependence of a on ϵ is independent of the system size. Hence, in order to justify the validity of (9) it is sufficient to prove that the width of the distribution of the energy per particle in the diagonal ensemble converges to zero for large linear sizes L of the system:

$$\Delta\epsilon \overset{L\to\infty}{\longrightarrow} 0\,. \tag{10}$$

Suppose that initially our system is prepared in an eigenstate $|\Psi_0\rangle$ of a Hamiltonian $\widehat{H}_0$ and that at time $t = 0$ the Hamiltonian is suddenly changed to $\widehat{H}$:

$$\widehat{H}_0 \to \widehat{H} = \widehat{H}_0 + \widehat{W},$$

where $\widehat{W}$ is the difference between the new and the old Hamiltonians. Within this scenario, the energy width

$$\Delta E = \sqrt{\sum_\alpha E_\alpha^2 |C_\alpha|^2 - \left(\sum_\alpha E_\alpha |C_\alpha|^2\right)^2}$$

of the diagonal ensemble becomes equal to the variance of the new energy in the state $|\Psi_0\rangle$:

$$\Delta E = \Delta H \equiv \sqrt{\langle\Psi_0|\widehat{H}^2|\Psi_0\rangle - \langle\Psi_0|\widehat{H}|\Psi_0\rangle^2}\,.$$

It is now straightforward to show that the variance of $\widehat{H}$ equals the variance of $\widehat{W}$:

$$\Delta H = \Delta W.$$

In order to deduce how ΔW scales in the thermodynamic limit, we assume that $\widehat{W}$ is a sum of local operators $\hat{w}(j)$ over some region of the lattice σ (a single point, a straight line, the whole lattice, etc.):

$$\widehat{W} = \sum_{j\in\sigma} \hat{w}(j).$$

Here $\hat{w}(j)$ is a polynomial of creation and annihilation operators localized at the points $j + \Delta j$, where $|\Delta j|$ is limited from the above by a finite number that does not scale with the system size. The mean square of $\widehat{W}$ can be written as

$$\begin{aligned}\langle\Psi_0|\widehat{W}^2|\Psi_0\rangle = {} & \langle\Psi_0|\widehat{W}|\Psi_0\rangle^2 \\ & + \sum_{j_1,j_2\in\sigma} \left[\langle\Psi_0|\hat{w}(j_1)\hat{w}(j_2)|\Psi_0\rangle\right. \\ & \left. - \langle\Psi_0|\hat{w}(j_1)|\Psi_0\rangle\langle\Psi_0|\hat{w}(j_2)|\Psi_0\rangle\right]. \end{aligned} \tag{11}$$

In the absence of long-range correlations the expression in brackets tends to zero for large distances between j_1 and j_2. Therefore, the whole second term on the right-hand-side of (11) scales as L^{d_σ}, where d_σ is the dimensionality of the sublattice σ and L is the linear size of the lattice. The variance of $\widehat{W}$ scales the same way:

$$(\Delta W)^2 \overset{L\to\infty}{\propto} L^{d_\sigma}.$$

Retracing our steps, we arrive at the conclusion that the energy width $\Delta\epsilon$ indeed tends to zero in the thermodynamic limit:

$$\Delta\epsilon \overset{L\to\infty}{\propto} \frac{1}{L^{d_L - d_\sigma/2}},$$

where $d_L \geq d_\sigma$ is the dimensionality of the whole lattice.

Note that for the two-dimensional lattice considered in this paper the role of $\widehat{W}$ is played by the hopping energy of the "door" link. An analysis similar to the one above shows that increasing the number of "door" links will lead to an increase in the width $\Delta\epsilon$, proportional to the square root of the number of "door" links. This is why in our 2D experiment, we have chosen the position of the missing sites to be the one in Fig. 5, so that only a single link is opened during the time evolution.

[1] Kinoshita, T., Wenger, T. & Weiss, D. S. A quantum Newton's cradle. *Nature* **440**, 900–903 (2006).

[2] Hofferberth, S., Lesanovsky, I., Fischer, B., Schumm, T. & Schmiedmayer, J. Non-equilibrium coherence dynamics in one-dimensional Bose gases. *Nature* **449**, 324–327 (2007).

[3] Rigol, M., Dunjko, V., Yurovsky, V. & Olshanii, M. Relaxation in a completely integrable many-body quantum system: An ab initio study of the dynamics of the highly excited states of 1d lattice hard-core bosons. *Phys. Rev. Lett.* **98**, 050405 (2007).

[4] Kollath, C., Läuchli, A. & Altman, E. Quench dynamics and nonequilibrium phase diagram of the Bose-Hubbard model. *Phys. Rev. Lett.* **98**, 180601 (2007).

[5] Manmana, S. R., Wessel, S., Noack, R. M. & Muramatsu, A. Strongly correlated fermions after a quantum quench. *Phys. Rev. Lett.* **98**, 210405 (2007).

[6] Burkov, A. A., Lukin, M. D. & Demler, E. Decoherence dynamics in low-dimensional cold atom interferometers. *Phys. Rev. Lett.* **98**, 200404 (2007).

[7] Calabrese, P. & Cardy, J. Quantum quenches in extended systems. *J. Stat. Mech.* P06008 (2007).

[8] Sengupta, K., Powell, S. & Sachdev, S. Quench dynamics across quantum critical points. *Phys. Rev. A* **69**, 053616 (2004).

[9] Berges, J., Borsányi, S. & Wetterich, C. Prethermalization. *Phys. Rev. Lett.* **93**, 142002 (2004).

[10] Gallavotti, G. *Statistical Mechanics: A Short Treatise* (Springer, Berlin, 1999).

[11] Krylov, N. S. *Works on the Foundation of Statistical Physics* (Princeton Univ. Press, Princeton, 1979).

[12] Deutsch, J. M. Quantum statistical mechanics in a closed system. *Phys. Rev. A* **43**, 2046–2049 (1991).

[13] Srednicki, M. Chaos and quantum thermalization. *Phys. Rev. E* **50**, 888–901 (1994).

[14] Sutherland, B. *Beautiful Models* (World Scientific, Singapore, 2004).

[15] Rigol, M., Muramatsu, A. & Olshanii, M. Hard-core bosons on optical superlattices: Dynamics and relaxation in the superfluid and insulating regimes. *Phys. Rev. A* **74**, 053616 (2006).

[16] Cazalilla, M. A. Effect of suddenly turning on interactions in the Luttinger model. *Phys. Rev. Lett.* **97**, 156403 (2006).

[17] Jin, D. S., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Collective excitations of a Bose-Einstein condensate in a dilute gas. *Phys. Rev. Lett.* **77**, 420–423 (1996).

[18] Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. W. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39–44 (2002).

[19] Mandel, O. *et al.* Controlled collisions for multi-particle entanglement of optically trapped atoms. *Nature* **425**, 937–940 (2003).

[20] Horoi, M., Zelevinsky, V. & Brown, B. A. Chaos vs thermalization in the nuclear shell model. *Phys. Rev. Lett.* **74**, 5194–5197 (1995).

[21] Shnirelman, A. I. Ergodic properties of eigenfunctions. *Usp. Mat. Nauk* **29**, 181–182 (1974).

[22] Voros, A. *Stochastic Behavior in Classical and Quantum Hamiltonian Systems* (Springer, Berlin, 1979).

[23] de Verdière, Y. C. Ergodicité et fonctions propres du Laplacien. *Commun. Math. Phys.* **102**, 497–502 (1985).

[24] Zelditch, S. Uniform distribution of eigenfunctions on compact hyperbolic surfaces. *Duke Math. J.* **55**, 919–941 (1987).

[25] Heller, E. J. & Landry, B. R. Statistical properties of many particle eigenfunctions. *J. Phys. A* **40**, 9259–9274 (2007).

[26] Berry, M. V. Regular and irregular semiclassical wavefunctions. *J. Phys. A* **10**, 2083–2091 (1977).

[27] Korepin, V. E., Bogoliubov, N. M. & Izergin, A. G. *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge Univ. Press, Cambridge, 1993).

[28] Srednicki, M. Does quantum chaos explain quantum statistical mechanics? Preprint at ⟨http://arxiv.org/abs/cond-mat/9410046⟩ (1994).

[29] Srednicki, M. Thermal fluctuations in quantized chaotic systems. *J. Phys. A* **29**, L75–L79 (1996).

[30] José, J. V. & Saletan, E. J. *Classical Dynamics: A Contemporary Approach* (Cambridge Univ. Press, Cambridge, 1998).

cond-mat/9809360

The approach to thermal equilibrium in quantized chaotic systems

Mark Srednicki*
Department of Physics, University of California, Santa Barbara, CA 93106

Abstract

We consider many-body quantum systems that exhibit quantum chaos, in the sense that the observables of interest act on energy eigenstates like banded random matrices. We study the time-dependent expectation values of these observables, assuming that the system is in a definite (but arbitrary) pure quantum state. We induce a probability distribution for the expectation values by treating the zero of time as a uniformly distributed random variable. We show explicitly that if an observable has a nonequilibrium expectation value at some particular moment, then it is overwhelmingly likely to move towards equilibrium, both forwards and backwards in time. For deviations from equilibrium that are not much larger than a typical quantum or thermal fluctuation, we find that the time dependence of the move towards equilibrium is given by the Kubo correlation function, in agreement with Onsager's postulate. These results are independent of the details of the system's quantum state.

Typeset using REVTEX

*E–mail: mark@tpau.physics.ucsb.edu

I. INTRODUCTION

Many-body systems typically exhibit certain dynamical properties that are studied under the subject headings of thermodynamics and statistical mechanics. These properties include the following:

1) Given an arbitrary initial state, the system almost always evolves towards an identifiable condition known as thermal equilibrium, and then remains in this condition at almost all subsequent times.

2) When the system is in thermal equilibrium, observables of interest take on values that depend only on the nature of the system and its total energy, but not on any other details of the specific state of the system.

3) When the system is in thermal equilibrium, the measured value of an observable of interest at any particular time fluctuates about its equilibrium value, with fluctuations whose amplitude is suppressed by a factor of $N^{-1/2}$, where N is the number of degrees of freedom.

4) During the approach to thermal equilibrium, the values of observables of interest are governed by equations that are not time reversal invariant. These equations typically depend on the values of other observables, possibly at different times. The information about the system that is included in these equations is not sufficient to reconstruct the complete physical state of the system.

5) Often (but not always), these equations are markovian; that is, they depend only on the values of the observables in question, and their first time derivatives, at any given moment.

There is a vast literature on the derivation of these properties from an underlying deterministic, time reversal invariant dynamics, classical or quantum. In this paper (closely related earlier work includes [1–16]), we explore to what extent these properties can be deduced as consequences of *quantum chaos*. This means that we will assume that the energy eigenvalues and (more importantly) the matrix elements of typical observables have certain properties. These properties are believed (and, in some cases, rigorously proven) to hold for a canonically quantized system whose classical phase space is fully chaotic at the energies of interest, and they are likely to hold at least approximately for a broader array of systems.

The outline of this paper is as follows. In Section 2 we state our basic assumptions, and briefly discuss their origins in quantum chaos theory. Some previous work that is directly relevant is summarized in Section 3. Sections 4 and 5 present new results concerning the approach to thermal equilibrium. Section 6 discusses the main conclusions.

II. QUANTUM CHAOS

We assume that the quantum system of interest is bounded and isolated, with N degrees of freedom, where $N \gg 1$. Since the system is bounded, the energy eigenvalues are discrete, and since it is isolated, its time evolution is governed by the Schrodinger equation. Let E_α denote the energy eigenvalue corresponding to the energy eigenstate $|\alpha\rangle$; the state of the system at any time t is then given by

$$|\psi_t\rangle = \sum_\alpha c_\alpha \, e^{-iE_\alpha t} \, |\alpha\rangle \, . \tag{2.1}$$

We emphasize that $|\alpha\rangle$ is an eigenstate of the full, many-body hamiltonian. The c_α's specify the state at any one time (say, $t = 0$), and we assume the usual normalization

$$\langle\psi_t|\psi_t\rangle = \sum_\alpha |c_\alpha|^2 = 1\,. \tag{2.2}$$

Note that we have set $\hbar = 1$ to simplify the notation. We will, however, point out how various quantities scale with $\hbar$ when this information is useful.

We now make two key assumptions about the system.

Our first assumption is not strictly necessary, but it will simplify some of the subsequent analysis. We assume that if any two sums of equal numbers of energy eigenvalues are equal,

$$E_{\alpha_1} + \ldots + E_{\alpha_n} = E_{\beta_1} + \ldots + E_{\beta_n}\,, \tag{2.3}$$

then $\{\beta_1, \ldots, \beta_n\}$ is a permutation of $\{\alpha_1, \ldots, \alpha_n\}$; that is, the corresponding eigenstates are the same for both sides. In particular, $E_\alpha = E_\beta$ implies that $|\alpha\rangle = |\beta\rangle$, so that all the energy eigenvalues are nondegenerate. This assumption is expected to hold in general for a quantized chaotic system in which all unitary symmetries (such as invariance under any discrete or continuous reflection or rotation) have been removed by suitable changes (such as by putting the system in an irregularly shaped box). In this case, the energy eigenvalues have the same statistical properties as the eigenvalues of gaussian random matrices [17]. The system may or may not be invariant under the anti-unitary transformation of time reversal; we will assume that it is, because one of the most interesting aspects of thermodynamics is the appearance of an "arrow of time" even when the underlying dynamics is time reversal invariant.

Our second assumption is the crucial one. Let A be a hermitian operator corresponding to an observable of interest that is a smooth, $\hbar$-independent function of the classical coordinates and momenta. We assume that the matrix elements of A in the energy eigenstate basis take the form [18,6]

$$A_{\alpha\beta} = \mathcal{A}(E)\delta_{\alpha\beta} + e^{-S(E)/2} f(E,\omega) R_{\alpha\beta}\,. \tag{2.4}$$

All the factors in this formula need explanation.

First, for notational convenience we have defined

$$E \equiv \tfrac{1}{2}(E_\alpha + E_\beta) \qquad \text{and} \qquad \omega \equiv E_\alpha - E_\beta\,. \tag{2.5}$$

$S(E)$ is the thermodynamic entropy at energy E, given by

$$e^{S(E)} \equiv E \sum_\alpha \delta_\varepsilon(E - E_\alpha)\,, \tag{2.6}$$

where $\delta_\varepsilon(x)$ is a Dirac delta function that has been smeared just enough to render $S(E)$ monotonic. $\mathcal{A}(E)$ and $f(E,\omega)$ are smooth functions of their arguments whose physical implications will be the main focus of this paper. Finally, $R_{\alpha\beta}$ is a numerical factor that varies erratically with α and β. It is helpful to think of the real and imaginary parts of $R_{\alpha\beta}$ as random variables, each with zero mean and unit variance. Without loss of generality, we can take $f(E,\omega)$ to be real, positive, and an even function of ω; then hermiticity of A implies $R_{\beta\alpha} = R^*_{\alpha\beta}$. Also, in many cases of physical interest, $R_{\alpha\beta}$ is purely real.

Eq. (2.4) is semiclassical in nature; the factor of $e^{-S(E)/2}$ scales like $\hbar^{(N-1)/2}$. Thus the validity of eq. (2.4) requires $\hbar$ to be "small," which in practice means that the energy E must be "large." The appropriate definitions of "small" and "large" are a key problem of quantum chaos theory. For systems with few degrees of freedom, it is now well established [19] that a necessary condition for quantum chaos is $\delta \ll \hbar/\tau_{\rm Th}$, where $\delta \sim e^{-S}E$ is the mean spacing between energy eigenvalues near E, and $\tau_{\rm Th}$ is the Thouless time [20] (roughly speaking, the time scale by which all diffusive classical processes have saturated). For many-body systems in general, less is known. Studies of nonlinearly coupled oscillators [7,14] and of fermions with pseudo-random single-particle energies and two-body matrix elements [11,15] both suggest a threshold energy for quantum chaos that goes to zero in the thermodynamic limit like $N^{-\nu}$, with $0 < \nu \leq 1$. We may therefore hope that eq. (2.4) will enjoy a wide range of validity.

An important feature of eq. (2.4) is that the general structure that it describes is preserved under multiplication [6]. Thus, for example, the matrix elements of any power of A are given by

$$(A^n)_{\alpha\beta} = \mathcal{A}_n(E)\delta_{\alpha\beta} + e^{-S(E)/2} f_n(E,\omega) R^{(n)}_{\alpha\beta}\,, \tag{2.7}$$

where $\mathcal{A}_n(E)$, $f_n(E,\omega)$, and $R^{(n)}_{\alpha\beta}$ can be expressed in terms of $\mathcal{A}(E)$, $f(E,\omega)$, and $R_{\alpha\beta}$. The precise relationship will not be needed, however; the key point is the generic character of eq. (2.4).

Finally, the function $\mathcal{A}(E)$ can be related to a standard expression in statistical mechanics: the equilibrium value of A, as given by the canonical thermal average

$$\langle A\rangle_T \equiv \frac{{\rm Tr}\, e^{-H/T} A}{{\rm Tr}\, e^{-H/T}}\,. \tag{2.8}$$

Here T is the temperature, and we have set Boltzmann's constant to one. To see the relation between $\langle A\rangle_T$ and $\mathcal{A}(E)$, we use eqs. (2.4) and (2.6) in eq. (2.8) to get

$$\langle A\rangle_T = \frac{\int_0^\infty \frac{dE}{E} e^{S(E)-E/T}\mathcal{A}(E)}{\int_0^\infty \frac{dE}{E} e^{S(E)-E/T}} + O(e^{-S/2})\,. \tag{2.9}$$

When N is large, the entropy is extensive: $S(E,N) = Ns(E/N) + O(\log N)$. In this case the integrals in eq. (2.9) can be evaluated by steepest descent. We see that their ratio is $\mathcal{A}(E) + O(N^{-1})$, where E is now fixed in terms of T by the steepest-descent condition $\partial S/\partial E = 1/T$; this also implies $E = \langle H\rangle_T$. Turning around eq. (2.9) then gives us the expression for $\mathcal{A}(E)$ that we want [6],

$$\mathcal{A}(E) = \langle A\rangle_T + O(N^{-1}) + O(e^{-S/2})\,. \tag{2.10}$$

We will always assume that N is large enough to make the indicated corrections negligible.

Note also that eq. (2.10) is consistent with, but not identical to, Shnirelman's theorem [21]. This theorem essentially states that $\mathcal{A}(E)$ is given by the classical, microcanonical average of A, up to corrections which are expected to be $O(\hbar^{1/2})$. Eq. (2.10), on the other hand, already includes corrections up to $O(\hbar^{(N-2)/2})$, but has in addition corrections of $O(N^{-1})$.

III. THERMAL EQUILIBRIUM

The expectation value of an observable A in the state specified by eq. (2.1) is given by

$$\begin{aligned} A_t &\equiv \langle \psi_t | A | \psi_t \rangle \\ &= \sum_{\alpha\beta} c_\alpha^* c_\beta \, e^{i(E_\alpha - E_\beta)t} A_{\alpha\beta} \, . \end{aligned} \tag{3.1}$$

We will take A_t as the main object of study. It is not obvious that this is the right thing to do, since short-time measurements do not generally yield quantum expectation values. However, our main goal is to compare with the results of conventional nonequilibrium statistical mechanics, in which time-dependent expectation values are the basic ingredients (see, e.g., [22–24], and Section V, below). A detailed discussion of the quantum measurement problem would be needed to address this issue properly, but this is beyond the scope of the present paper.

If we now take the infinite time average of A_t, we find

$$\begin{aligned} \overline{A} &\equiv \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau dt \; A_t \\ &= \sum_\alpha |c_\alpha|^2 A_{\alpha\alpha} \, , \end{aligned} \tag{3.2}$$

where the last line requires nondegeneracy of the energy eigenvalues. It also requires that the averaging time τ be much larger than the Heisenberg time $\tau_{\rm H} \equiv 2\pi\hbar/\delta \sim e^S$. This time scale is much too large to be physically relevant, and thus the infinite time average must be regarded as a purely theoretical device. Nevertheless, if the system comes to thermal equilibrium, then A_t should be near its equilibrium value $\langle A \rangle_T$ most of the time, and thus we should have $\overline{A} = \langle A \rangle_T$.

To check whether or not this is the case, we first substitute eq. (2.4) into eq. (3.2) to get

$$\overline{A} = \sum_\alpha |c_\alpha|^2 \mathcal{A}(E_\alpha) + O(e^{-S/2}) \, . \tag{3.3}$$

We now make a mild assumption about the state $|\psi_t\rangle$. We assume that the expectation value of the total energy

$$E \equiv \sum_\alpha |c_\alpha|^2 E_\alpha \tag{3.4}$$

has a quantum uncertainty

$$\Delta \equiv \Big[\sum_\alpha |c_\alpha|^2 \, (E_\alpha - E)^2 \Big]^{1/2} \tag{3.5}$$

that is small, in the sense that $\Delta^2 |\mathcal{A}''(E)/\mathcal{A}(E)| \ll 1$. This is a natural assumption when N is large, since states of physical interest typically have $\Delta \sim N^{-1/2} E$. If we now expand $\mathcal{A}(E_\alpha)$ in eq. (3.3) in powers of $E_\alpha - E$, we get $\overline{A} = \mathcal{A}(E) + O(\Delta^2)$; combining this with eq. (2.10), we find

$$\overline{A} = \langle A \rangle_T + O(\Delta^2) + O(N^{-1}) + O(e^{-S/2}) \,. \tag{3.6}$$

Thus we have shown that the infinite time average of A_t is indeed equal to its equilibrium value at the appropriate temperature. Note that this key property follows entirely from the matrix element structure of eq. (2.4), and does not depend on the details of the initial quantum state.

We now turn to an examination of the fluctuations of A_t about its equilibrium value $\overline{A}$. The mean squared amplitude of the these fluctuations is

$$\begin{aligned} \overline{(A_t - \overline{A}\,)^2} &= \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau dt\, (A_t - \overline{A}\,)^2 \\ &= \sum_{\alpha,\beta\neq\alpha} |c_\alpha|^2\, |c_\beta|^2\, |A_{\alpha\beta}|^2 \\ &= O(e^{-S}) \,. \end{aligned} \tag{3.7}$$

We see that the fluctuations of A_t about $\overline{A}$ are very small. This tells us that, whatever the initial value A_0 happens to be, A_t must eventually approach its equilibrium value, and then remain near it most of the time.

On the other hand, eq. (3.7) is too small to represent the expected thermal fluctuations of A. To find them, we must look at what are usually called quantum fluctuations. The mean squared amplitude of the quantum fluctuations is

$$\begin{aligned} \overline{\langle\psi_t|(A - \overline{A}\,)^2|\psi_t\rangle} &= \overline{(A^2)_t} - \overline{A}^2 \\ &= \langle A^2\rangle_T - \langle A\rangle_T^2 \\ &\quad + O(\Delta^2) + O(N^{-1}) + O(e^{-S/2}) \,. \end{aligned} \tag{3.8}$$

In the last line, we have used the fact that the matrix elements of A^2 have the same general structure as the matrix elements of A, as shown in eq. (2.7), and that this structure implies that the infinite time average is the same as the thermal average, as shown in eq. (3.6). Eq. (3.8) tells us that the quantum fluctuations in A have the right magnitude to be identified as thermal fluctuations [6].

Note, however, that $\langle A^2\rangle_T - \langle A\rangle_T^2$ is itself expected to be $O(N^{-1})$ for typical observables of interest (see, e.g., [25]), and so the first two correction terms on the right-hand side of eq. (3.8) are not necessarily smaller than the leading term. This is not a point of concern, however; we used the canonical ensemble to define thermal averaging, and the result would in general change by a factor of order one if we were to use instead the grand canonical or microcanonical ensemble. Since the exact size of the thermal fluctuations in any particular observable depends on the choice of ensemble, our claimed identification of quantum fluctuations as thermal fluctuations can be meaningful only up to a numerical factor. This is what is established in eq. (3.8).

Another point of interest is the nature of the classical limit. Recalling that $e^{-S} \sim \hbar^{N-1}$, we see that eq. (3.7) predicts that $\overline{(A_t - \overline{A}\,)^2}$ vanishes in the classical limit. This is in fact a reasonable result if the classical system is chaotic. To see why, first note that the classical limit of a generic quantum state (to the extent that such a thing exists at all) is a probability density on phase space. Then the time dependent expectation value A_t is given (in the classical limit) by $A_t = \int d^{2N}X\, U_t\rho(X)A(X)$, where X is a point in phase

space, $\rho(X)$ is the probability density associated with the quantum state at $t = 0$, and U_t is the Frobenius-Perron evolution operator for phase-space densities. When suitably regulated and renormalized, this formally unitary operator acquires eigenvalues (known as Ruelle resonances) inside the unit circle that are associated with irreversible decay to the ergodic distribution [26]. If both $A(X)$ and $\rho(X)$ are smooth functions, A_t approaches a fixed limit as $t \to \infty$, and $\overline{(A_t - \overline{A}\,)^2}$ vanishes. The mean squared amplitude of the classical thermal fluctuations is then given by the infinite time average of $\int d^{2N}X\, U_t\rho(X)[A(X) - \overline{A}\,]^2$, which is the classical limit of the quantum expression $\langle\psi_t|(A - \overline{A}\,)^2|\psi_t\rangle$.

IV. APPROACHING EQUILIBRIUM

We now turn to our main topic, the approach to thermal equilibrium. Suppose we have an initial state $|\psi_0\rangle = \sum_\alpha c_\alpha|\alpha\rangle$ such that the initial expectation value $A_0 = \langle\psi_0|A|\psi_0\rangle$ of an observable A is far from its equilibrium value $\overline{A} = \langle A\rangle_T$. What, then, is the behavior of $A_t = \langle\psi_t|A|\psi_t\rangle$ at later times?

It is obvious that the answer depends on the details of the initial state. Without knowing them, we can only make a probabilistic analysis. There are two basic methods for doing so. One is to introduce a probability distribution for the initial state itself, and to average relevant quantities over it. The issue then becomes the justification of the procedure (e.g., maximum entropy) for choosing this particular distribution. The second method, which we will adopt, is to fix the initial state, and then study the values of interesting observables as functions of time. We treat the observation time as a uniformly distributed random variable, thus inducing a probability distribution for each observable. We then attempt to determine to what extent these probability distributions depend on the initial state. Ideally, there would be no dependence at all, thus rendering the choice of the initial state irrelevant.

The moments of the probability distribution for A_t (which is induced by assuming a uniform probability distribution for t) are given by the infinite time averages

$$\overline{(A_t)^n} = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau dt\ (A_t)^n\,. \tag{4.1}$$

Again we note that we are using the infinite time average simply as a mathematical tool; in this context, the fact that astronomically long averaging times are necessary is not relevant. By using eq. (3.1), and the nondegeneracy assumption discussed after eq. (2.3), we can express these moments as products of traces of powers of the matrix

$$\widetilde{A}_{\alpha\beta} \equiv c^*_\alpha A_{\alpha\beta} c_\beta\,. \tag{4.2}$$

It will simplify the notation considerably if we first shift the operator A by a constant, so that the infinite time average of A_t is zero. This entails no loss of generality, and so from here on we will take

$$\overline{A} = \mathrm{Tr}\,\widetilde{A} = 0\,. \tag{4.3}$$

The first few moments $\overline{(A_t)^n}$ can then be expressed as

$$
\begin{aligned}
\overline{(A_t)^2} &= \operatorname{Tr} \widetilde{A}^2 \,, \\
\overline{(A_t)^3} &= 2 \operatorname{Tr} \widetilde{A}^3 \,, \\
\overline{(A_t)^4} &= 3(\operatorname{Tr} \widetilde{A}^2)^2 + 6 \operatorname{Tr} \widetilde{A}^4 \,, \\
\overline{(A_t)^5} &= 20 \operatorname{Tr} \widetilde{A}^2 \operatorname{Tr} \widetilde{A}^3 + 24 \operatorname{Tr} \widetilde{A}^5 \,, \\
\overline{(A_t)^6} &= 15(\operatorname{Tr} \widetilde{A}^2)^3 + 90 \operatorname{Tr} \widetilde{A}^2 \operatorname{Tr} \widetilde{A}^4 + 40(\operatorname{Tr} \widetilde{A}^3)^2 + 120 \operatorname{Tr} \widetilde{A}^6 \,.
\end{aligned} \tag{4.4}
$$

Determining these coefficients is a straightforward but tedious combinatoric problem; it involves counting the number of different ways that the time-dependent phases can cancel against each other, and so survive the infinite time average. Generalizing from eq. (4.4), the coefficient of a term of the form $\prod_i (\operatorname{Tr} \widetilde{A}^{m_i})^{p_i}$ in the expansion of $\overline{(A_t)^n}$ is given by $n!/(\prod_i p_i! m_i^{p_i})$, where $n = \sum_i m_i p_i$.

We now wish to estimate the magnitude of $\operatorname{Tr} \widetilde{A}^m$ for different values of m. Necessary inputs include the quantum uncertainty Δ in the total energy E, given by eq. (3.5); the energy bandwidth W of the matrix $A_{\alpha\beta}$ near energy E, given by

$$
W \equiv \frac{\int_{-\infty}^{+\infty} d\omega \, |f(E,\omega)|^2}{|f(E,0)|^2} \,, \tag{4.5}
$$

where $f(E,\omega)$ is defined in eq. (2.4); the typical size a of the quantum/thermal fluctuations in A, given by

$$
a^2 \equiv \overline{(A^2)_t} = \sum_\alpha |c_\alpha|^2 \, (A^2)_{\alpha\alpha} = \sum_{\alpha\beta} |c_\alpha|^2 \, |A_{\alpha\beta}|^2 \,; \tag{4.6}
$$

and the inverse participation ratio

$$
\mathcal{I} \equiv \frac{1}{\sum_\alpha |c_\alpha|^4} \,, \tag{4.7}
$$

which counts the effective number of different energy levels that are present in the quantum state of the system. Also, $\mathcal{I}^{-1}$ can be regarded as the average value of $|c_\alpha|^2$. Given $\mathcal{I}$, it is helpful to define an effective level spacing between the participating energy eigenstates,

$$
\delta_{\text{eff}} \equiv \frac{\Delta}{\mathcal{I}} \,. \tag{4.8}
$$

In general, δ_{eff} must be greater than or equal to the actual level spacing $\delta \sim e^{-S} E$.

Before proceeding to evaluate $\operatorname{Tr} \widetilde{A}^m$, we must order the various energy scales. We expect that "typical" states of physical interest will have $\delta_{\text{eff}} \sim \delta$ and $\Delta \sim N^{-1/2} E \sim N^{1/2} T$, since these are properties of a thermal density matrix. We also expect that W will be independent of N. To see why, we turn to the formula for A_t, eq. (3.1). From it, we can see that the time scale for the initial evolution of A_t is either $\hbar/W$ or $\hbar/\Delta$, whichever is larger. (Before this time, no relative phases have changed significantly.) For an observable A of interest, this time scale should be finite and nonzero in the thermodynamic limit, and hence independent of N. Since we expect $\Delta \sim N^{1/2} T$, it must be W that is independent of N. We therefore conclude that

$$
\delta_{\text{eff}} \ll W \ll \Delta \ll E \tag{4.9}
$$

is the regime of interest.

With these considerations in place, we note that we can regard $\widetilde{A}$ as a banded random matrix of overall size $(\Delta/\delta_{\rm eff}) \times (\Delta/\delta_{\rm eff})$ and bandwidth $W/\delta_{\rm eff}$. Within the band of nonzero matrix elements, the magnitude of a typical entry is

$$\widetilde{A}_{\rm typ} \sim \mathcal{I}^{-1} (W/\delta)^{-1/2} a \,. \tag{4.10}$$

This comes from eq. (4.6), whose last equality demonstrates that a typical value of $|A_{\alpha\beta}|^2$ is $(W/\delta)^{-1} a^2$; the extra factor of $\mathcal{I}^{-1}$ in eq. (4.10) comes from the c_α's in eq. (4.2). We can now estimate $\mathrm{Tr}\, \widetilde{A}^m$ for even $m \equiv 2k$ as

$$\begin{aligned} \mathrm{Tr}\, \widetilde{A}^{2k} &\sim \left(\frac{\Delta}{\delta_{\rm eff}}\right) \left(\frac{W}{\delta_{\rm eff}}\right)^k \left(\widetilde{A}_{\rm typ}\right)^{2k} \\ &\sim \left(\frac{\Delta}{\delta_{\rm eff}}\right) \left(\frac{\delta_{\rm eff}}{\Delta}\right)^{2k} \left(\frac{\delta}{\delta_{\rm eff}}\right)^k a^{2k} \,. \end{aligned} \tag{4.11}$$

For odd $m \equiv 2k+1$, the expected value of $\mathrm{Tr}\, \widetilde{A}^{2k+1}$ can be positive or negative; a root-mean-square estimate of its magnitude is

$$\begin{aligned} \mathrm{Tr}\, \widetilde{A}^{2k+1} &\sim \left(\frac{\Delta}{\delta_{\rm eff}}\right)^{1/2} \left(\frac{W}{\delta_{\rm eff}}\right)^k \left(\widetilde{A}_{\rm typ}\right)^{2k+1} \\ &\sim \left(\frac{\Delta}{W}\right)^{1/2} \left(\frac{\delta_{\rm eff}}{\Delta}\right)^{2k+1} \left(\frac{\delta}{\delta_{\rm eff}}\right)^{(2k+1)/2} a^{2k+1} \,. \end{aligned} \tag{4.12}$$

However, this estimate should be regarded as less trustworthy than eq. (4.11).

Eqs. (4.11) and (4.12) imply that the probability distribution for A_t has the form

$$P(A_t) \propto \exp\left[-\left(\frac{\Delta}{\delta_{\rm eff}}\right) F_+\left(\frac{A_t}{(\delta/\delta_{\rm eff})^{1/2} a}\right) - \left(\frac{\Delta}{W}\right)^{1/2} F_-\left(\frac{A_t}{(\delta/\delta_{\rm eff})^{1/2} a}\right) \right] , \tag{4.13}$$

where $F_+(x)$ and $F_-(x)$ are even and odd functions (respectively) whose Taylor expansions involve purely numerical coefficients. Eq. (4.13) can be verified by via a Feynman-diagram expansion for the moments; the quadratic term from F_+ provides the propagator, and the remaining terms (in both F_+ and F_-) give the vertex coefficients.

For $A_t \lesssim (\delta/\delta_{\rm eff})^{1/2} a$, we can neglect F_- and all but the quadratic term in F_+; we then have

$$P(A_t) \propto \exp\left[-\frac{\xi_2}{2} \left(\frac{\Delta}{\delta}\right) \left(\frac{A_t}{a}\right)^2 \right] , \tag{4.14}$$

where $F_+(x) = \frac{1}{2}\xi_2 x^2 + O(x^4)$, and ξ_2 is a number of order one. Thus, for sufficiently small values of A_t, its probability distribution (as induced by a uniform distribution for t) is gaussian, and independent of the details of the initial state; only the energy width Δ of this state is relevant. Furthermore, we see again that the fluctuations of A_t about its infinite time average (which is zero, by construction) are suppressed by a factor of $(\delta/\Delta)^{1/2} \sim e^{-S/2}$.

For $A_t \gtrsim (\delta/\delta_{\rm eff})^{1/2} a$, the nonuniversal behavior of the functions $F_\pm$ becomes relevant. Initial states that are "typical" according to most any reasonable criterion would all have $\delta_{\rm eff} \sim \delta$. In this case the nonuniversal corrections are important for $A_t \gtrsim a$.

So far, our analysis has not addressed our original question: given the initial value A_0, what is the behavior of A_t at later times? To answer this question, we must compute the conditional probability $P(A_t|A_0)$ to find the value A_t at time t, given the value A_0 at time zero. By the usual rules of probability theory, this conditional probability can be expressed as

$$P(A_t|A_0) = \frac{P(A_t, A_0)}{P(A_0)}\,, \tag{4.15}$$

where $P(A_t, A_0)$ is the joint probability for the observable A to have the expectation values A_t at time t and A_0 at time zero. This joint probability is induced by assuming a uniform distribution for the initial time; hence the moments of $P(A_t, A_0)$ are given by $\overline{(A_{t+t'})^n (A_{t'})^m}$, where the time averaging is done with respect to t', with t held fixed. To compute these moments, we need an expansion analogous to eq. (4.4). Let us focus on the universal regime. In this case, the dominant terms in the expansion, as in eq. (4.4), are those with the largest number of traces. Keeping only these terms results in a gaussian distribution. As always for a gaussian distribution, it is completely determined by its second moments,

$$\begin{aligned} \overline{A_{t+t'} A_{t'}} &= \sum_{\alpha\beta} |\widetilde{A}_{\alpha\beta}|^2 \, e^{i(E_\alpha - E_\beta)t}\,, \\ \overline{(A_{t+t'})^2} &= \overline{(A_{t'})^2} = \sum_{\alpha\beta} |\widetilde{A}_{\alpha\beta}|^2\,. \end{aligned} \tag{4.16}$$

To streamline the notation a little we define a correlation function

$$C(t) \equiv \overline{A_{t+t'} A_{t'}} \Big/ \overline{(A_{t'})^2}\,. \tag{4.17}$$

Note that, by construction, $C(0) = 1$.

These considerations lead to

$$P(A_t, A_0) \propto \exp\left[-\frac{\xi_2}{2} \left(\frac{\Delta}{\delta}\right) \frac{(A_t)^2 + (A_0)^2 - 2C(t) A_t A_0}{[1 - C(t)]^2 a^2} \right], \tag{4.18}$$

which has the correct second moments. Then, eqs. (4.14), (4.15), and (4.18) give us the conditional probability

$$P(A_t|A_0) \propto \exp\left[-\frac{\xi_2}{2} \left(\frac{\Delta}{\delta}\right) \frac{[A_t - C(t) A_0]^2}{[1 - C(t)]^2 a^2} \right] \tag{4.19}$$

in the universal regime. Recalling that $\Delta/\delta \sim e^S$, eq. (4.19) shows us that it is overwhelmingly likely that we will find $A_t = C(t) A_0$. In the next section, we will see that the correlation function $C(t)$ does not depend on the quantum state of the system. Thus, for all practical purposes, A_t is fully determined just by the initial value A_0; no other information about the state of the system is needed.

For $A_0 \gtrsim (\delta/\delta_{\text{eff}})^{1/2} a$, the nonuniversal corrections can become important. We then expect a formula for $P(A_t|A_0)$ that is similar to eq. (4.13). Because the large prefactor of $\Delta/\delta \sim e^S$ should still be present, we still expect to get an effectively deterministic equation for A_t as a function of t and A_0, although it will no longer take the simple form $A_t = C(t) A_0$.

Returning to the universal regime, we need to study the properties of $C(t)$. This we do in the next section.

V. LINEAR RESPONSE

Eqs. (4.16) and (4.17) imply that the correlation function $C(t)$ is real, even, and has a maximum magnitude of one. Also, it is quasiperiodic on the scale of the Heisenberg time $\tau_{\rm H} = 2\pi\hbar/\delta \sim e^S$. However, this time scale is much too long to be of physical interest, and so we can justifiably ignore the quasiperiodicity of $C(t)$. Then, also using eqs. (2.4) and (4.2), we get

$$\begin{aligned} C(t) &\propto \sum_{\alpha\beta} |c_\alpha|^2 \, |c_\beta|^2 \, |A_{\alpha\beta}|^2 \, e^{i(E_\alpha - E_\beta)t} \\ &\propto \int_{-\infty}^{+\infty} d\omega \, |f(E,\omega)|^2 \, e^{i\omega t} \, . \end{aligned} \tag{5.1}$$

The last line shows that $C(t)$ does not depend on the initial state, and that the bandwidth W of $f(E,\omega)$ sets the time scale for $C(t)$.

Eq. (5.1) can be compared with the results obtained through more standard methods. For example [22], we can give A a nonzero expectation value at $t = 0$ by supposing that, for $t < 0$, the system's hamiltonian was perturbed to $H + \lambda A$, where λ is a constant. For $t \le 0$, we assume that the quantum state of the system is described by a thermal density matrix $\rho_0 \sim e^{-(H+\lambda A)/T}$. For $t \ge 0$, this state is evolved forward in time with the original hamiltonian H, so that $\rho_t = e^{-iHt} \rho_0 e^{iHt}$; the time-dependent expectation value of A is then $A_t \equiv \mathrm{Tr}\, \rho_t A$. To leading order in λ, this procedure results in $A_t = C_{\rm Kubo}(t) A_0$, where

$$C_{\rm Kubo}(t) \propto \int_0^{1/T} d\mu \, \langle A_H(-i\mu) A_H(t) \rangle_T \, . \tag{5.2}$$

Here $A_H(t) \equiv e^{iHt} A e^{-iHt}$ is the operator A in the Heisenberg picture, and the angle brackets denote canonical thermal averaging as in eq. (2.8). Eq. (5.2) can be written in terms of the matrix elements $A_{\alpha\beta}$ as

$$\begin{aligned} C_{\rm Kubo}(t) &\propto \sum_{\alpha\beta} \frac{e^{-E_\alpha/T} - e^{-E_\beta/T}}{E_\beta - E_\alpha} \, |A_{\alpha\beta}|^2 \, e^{i(E_\alpha - E_\beta)t} \\ &\propto \int_{-\infty}^{+\infty} d\omega \, \frac{\sinh(\omega/2T)}{\omega} \, |f(E,\omega)|^2 \, e^{i\omega t} \, . \end{aligned} \tag{5.3}$$

This is not the same as eq. (5.1). Suppose, however, that the bandwidth W of $f(E,\omega)$ is smaller than the temperature T, and that the falloff of $f(E,\omega)$ for $\omega \gg W$ is fast enough to make the integral in eq. (5.3) converge. Then the factor of $\sinh(\omega/2T)/\omega$ will be approximately constant over the relevant range of ω, and so we will get

$$C_{\rm Kubo}(t) = C(t) + O(W^2/T^2) \, . \tag{5.4}$$

In this case, eq. (4.19) may be viewed as a verification of Onsager's postulate [27] that a random fluctuation (in the value of some quantity) will dissipate in the same way as an initial value that is produced by an applied force.

Let us now consider the circumstances under which the time evolution will be markovian. In the linear regime, and for $t > 0$, a markovian equation has the form $(d/dt)A_t = -\Gamma A_t$,

where the parameter Γ must be real and positive. Given $A_t = C(t)A_0$, this implies $C(t) = \exp(-\Gamma|t|)$. From eq. (5.1), we see that this corresponds to $|f(E,\omega)|^2 \propto 1/(\omega^2 + \Gamma^2)$, and hence $W \sim \Gamma$. However, $1/(\omega^2 + \Gamma^2)$ does not fall off fast enough at large ω for convergence of the integral in eq. (5.3), and so we expect some additional suppression of $|f(E,\omega)|^2$ when ω is greater than the temperature T. If it happens that Γ is much less than T, then the time evolution will still be approximately markovian, but with "memory" effects on time scales less than $\hbar/T$. This is consistent with other analyses [28,24], which find that the time evolution of expectation values is always non-markovian on time scales less than $\hbar/T$, essentially because of the energy-time uncertainty principle. The overall conclusion is that markovian time evolution is associated with an isolated pole in $|f(E,\omega)|^2$ at $\omega = \pm i\Gamma$, with $\Gamma \ll T$. However, this structure for $|f(E,\omega)|^2$ is not required by any general principles, and so non-markovian behavior on all time scales is an open possibility.

VI. DISCUSSION

Our major result is eq. (4.19), which gives the probability to find that an observable A has a quantum expectation value of A_t at time t, given that its expectation value is A_0 at time zero (and assuming that A has been shifted, if necessary, so that the infinite time average of A_t is zero). This probability is computed for a particular quantum state $|\psi_t\rangle$, but with the zero of time treated as a uniformly distributed random variable.

To understand the essential features of eq. (4.19), it is helpful to rewrite it more schematically as

$$P(A_t|A_0) \propto \exp\left\{-O(e^S)\left[A_t - C(t)A_0\right]^2/a^2\right\} . \tag{6.1}$$

Here $A_t \equiv \langle\psi_t|A|\psi_t\rangle$ is the time-dependent quantum expectation value of A; a^2 is the mean squared amplitude of the quantum fluctuations (which are also to be identified as thermal fluctuations) in A; S is the thermodynamic entropy at energy $E = \langle\psi_t|H|\psi_t\rangle$; and $C(t)$ is the correlation function (normalized to one at $t = 0$) given in eq. (5.1), which, under favorable circumstances, is the same as the Kubo correlation function $C_{\rm Kubo}(t)$ given in eq. (5.3). The range of validity of eq. (6.1) is $A_0 \lesssim (\delta/\delta_{\rm eff})^{1/2}a$, where $\delta \sim e^{-S}E$ is the mean energy-level spacing near E, and $\delta_{\rm eff} \geq \delta$ is the mean level spacing for those eigenstates that participate significantly in $|\psi_t\rangle$. Maximum participation (such as would be predicted by a thermal density matrix) corresponds to $\delta_{\rm eff} \sim \delta$.

Eq. (6.1) implies an effectively deterministic evolution equation for A_t, given only A_0 as input: $A_t = C(t)A_0$. The probability that this equation is not satisfied is $O(e^{-S})$. Since $C(t)$ is an even function of time, the effective evolution equation is invariant under time reversal. However, $C(t) \leq C(0)$ for all t, and $C(t)$ decays to zero if the bandwidth W, defined in eq. (4.5), is finite. (There will be quasiperiodic resurgences of $C(t)$ on the scale of the Heisenberg time $\tau_{\rm H} = 2\pi\hbar/\delta \sim e^S$, but this is much too long to be relevant.) Thus a nonzero initial value A_0 is overwhelmingly likely to evolve, both forwards and backwards in time, towards the equilibrium value of zero. This is of course entirely consistent with the heuristic picture of statistical mechanics originally proposed by Boltzmann.

If the envelope function $|f(E,\omega)|^2$, defined in eq. (2.4), has an isolated pole at $\omega = \pm i\Gamma$, then the time evolution will be approximately markovian, with $C(t) \simeq \exp(-\Gamma|t|)$. This is

consistent with $C(t) \simeq C_{\rm Kubo}(t)$ only if $\Gamma \ll T$, where T is the temperature corresponding to a thermodynamic energy of E. There will still be non-markovian "memory" effects on time scales less than $\hbar/T$, which is consistent with the results of other analyses [28,24].

To get eq. (6.1) [or, more precisely, (4.19)], we had to make a fairly strong assumption about the quantum matrix elements of A, eq. (2.4). However, this equation is well grounded in quantum chaos theory, and is likely to be at least approximately valid under a fairly wide range of circumstances.

Naturally it would be useful to extend our results to the nonlinear regime, and naturally this is very much harder to do. Still we hope to return to this question in future work. A combination of our methods (which easily demonstrate the ubiquity of thermal equilibrium) with projection-operator methods (which easily generate equations for expectation values, assuming appropriately forced thermal density matrices as the starting point [23,24]) might be a fruitful approach.

ACKNOWLEDGMENTS

I am grateful to Roland Ketzmerick, Jim Langer, and Doug Scalapino for helpful comments. This work was supported in part by NSF Grant PHY–22022.

REFERENCES

[1] Deutsch J M 1989 *Phys. Rev. A* **43** 2046

[2] Gaspard P 1992 *Quantum Chaos–Quantum Measurement* ed P Cvitanovic, I Percival and A Wirzba (Dordrecht: Kluwer)

[3] Montambaux G, Poilblanc D, Bellissard J and Sire C 1993 *Phys. Rev. Lett.* **70** 497
Poilblanc D, Ziman T, Bellissard J, Mila F and Montambaux G 1993 *Europhys. Lett.* **22** 537 (cond-mat/9301005)

[4] Gaspard P and van Ede van der Pals P 1994 *Phys. Rev. E* **49** 79
—— 1995 *Chaos Solitons Fractals* **5** 1183

[5] Srednicki M 1994 *Phys. Rev. E* **50** 888 (cond-mat/9403051)

[6] Srednicki M 1996 *J. Phys. A: Math. Gen.* **29** L75 (chao-dyn/9511001)

[7] Leitner D M and Wolynes P G 1996 *Phys. Rev. Lett.* **76** 216
—— 1996 *J. Chem. Phys.* **105** 11226

[8] Jain S R and Alonso D 1996 *Phys. Lett. B* **387** 812 (cond-mat/9608144)
—— 1997 *J. Phys. A: Math. Gen.* **30** 4993 (cond-mat/9609013)

[9] Gaspard P and Jain S R 1997 *Pramana J. Phys.* **48** 503 (cond-mat/9609187)

[10] Flambaum V V, Izrailev F M and Casati G 1996 *Phys. Rev. E* **54** 2136 (cond-mat/9607035)
Flambaum V V and Izrailev F M 1997 *Phys. Rev. E* **55** R13 (cond-mat/9610178)
—— 1997 *Phys. Rev. E* **56** 5144 (cond-mat/9707016)
Borgonovi F, Guarneri I, Izrailev F M and Casati G 1997 *Preprint* chao-dyn/9711005

[11] Jacquod Ph and Shepelyansky D L 1997 *Phys. Rev. Lett.* **79** 1837 (cond-mat/9706040)
Georgeot B and Shepelyansky D L 1997 *Phys. Rev. Lett.* **79** 4635 (cond-mat/9707231)

[12] Tasaki H 1998 *Phys. Rev. Lett.* **80** 1373 (cond-mat/9707253)

[13] Prosen T 1998 *Phys. Rev. Lett.* **80** 1808 (cond-mat/9707180)
—— 1998 *J. Phys. A: Math. Gen.* **31** L645 (cond-mat/9803358)
—— 1998 *Preprint* cond-mat/9808150

[14] Tsaur G-Y and Wang J 1998 *Phys. Rev. Lett.* **80** 3682

[15] Silvestrov P G 1998 *Preprint* cond-mat/9712036

[16] Casati G 1998 *Preprint* cond-mat/9801063

[17] Casati G, Valz-Gris F and Guarnieri I 1980 *Lett. N. Cim.* **28** 279
Bohigas O, Giannoni M-J and Schmit C 1984 *Phys. Rev. Lett.* **52** 1
Bohigas O 1991 *Chaos and Quantum Physics (Les Houches LII)* ed M-J Giannoni, A Voros and J Zinn-Justin (Amsterdam: North-Holland)
Guhr T, Mueller-Groeling A and Weidenmueller H A 1998 *Phys. Rep.* **299** 189 (cond-mat/9707301)

[18] Berry M V 1977 *J. Phys. A: Math. Gen.* **10** 2083
Pechukas P 1983 *Phys. Rev. Lett.* **51** 943
Chirikov B V 1985 *Phys. Lett. A* **108** 68
Feingold M and Peres A 1986 *Phys. Rev. A* **34** 591
Wilkinson M 1987 *J. Phys. A: Math. Gen* **20** 2415
Prosen T 1994 *Ann. Phys.* **235** 115
Eckhardt B and Main J 1995 *Phys. Rev. Lett.* **75** 2300 (chao-dyn/9508006)
Eckhardt B, Fishman S, Keating J, Agam O, Main J and Müller K 1995 *Phys. Rev. E* **52** 5893 (chao-dyn/9509017)

Hortikar S and Srednicki M 1998 *Phys. Rev. E* **57** 7313 (chao-dyn/9711020)
[19] Casati G and Chirikov B 1995 *Quantum Chaos* (New York: Cambridge Univ. Press).
Muzykantskii B A and Khmel'nitskii D E 1995 *JETP Lett.* **62** 76
Andreev A V, Agam O, Simons B D and Altshuler B L 1996 *Phys. Rev Lett.* **76** 3947 (cond-mat/9601001)
—— 1996 *Nucl. Phys.* **B482** 536 (cond-mat/9605204)
Borgonovi F, Casati G and Li B 1996 *Phys. Rev. Lett.* **77** 4744 (cond-mat/9610158)
[20] Thouless D J 1973 *J. Phys. C: Solid State* **6** L49
[21] Shnirelman A I 1974 *Ups. Mat. Nauk* **29** 181
Zelditch S 1987 *Duke Math J.* **55** 919
Colin de Verdière Y 1985 *Comm. Math. Phys.* **102** 497
—— 1991 *Chaos and Quantum Physics (Les Houches LII)* ed M-J Giannoni, A Voros and J Zinn-Justin (Amsterdam: North-Holland)
Helffer B, Martinez A and Robert D 1987 *Comm. Math. Phys.* **109** 313
Zelditch S and Zworski M 1996 *Comm. Math. Phys.* **175**, 673
Bäcker A, Schubert R and Stifter P, *Phys. Rev. E* **57** 5425 (chao-dyn/9709030)
[22] Kubo R 1957 *J. Phys. Soc. Japan* **12** 570
Kubo R, Toda M and Hashitsume N 1978 *Statistical Physics II* (Springer-Verlag, New York)
[23] Haake F 1973 *Statistical Treatment of Open Systems by Generalized Master Equations* (New York: Springer-Verlag)
Grabert H 1982 *Projection Operator Techniques in Nonequilibrium Statistical Mechanics* (New York: Springer-Verlag)
Fick E and Sauermann G 1990 *The Quantum Statistics of Dynamic Processes* (New York: Springer-Verlag)
[24] Rau J and Mueller B 1996 *Phys. Rep.* **272** 1 (nucl-th/9505009)
[25] Khinchin A Ia 1949 *Mathematical foundations of statistical mechanics* (New York: Dover)
[26] Ruelle D 1978 *Thermodynamic formalism* (Reading: Addison-Wesley)
Gaspard P, Nicolis G, Provata A and Tasaki S 1995 *Phys. Rev. E* **51** 74
[27] Onsager L 1931 *Phys. Rev.* **37** 405
[28] Danielewicz P 1984 *Ann. Phys. (N.Y.)* **152** 239