

Comments on defining entanglement entropy

Jennifer Lin^a, Đorđe Radičević^{b,*,1}

^a School of Natural Sciences, Institute for Advanced Study, Princeton, NJ, USA

^b Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5, Canada

Received 10 December 2019; received in revised form 15 July 2020; accepted 17 July 2020

Available online 30 July 2020

Editor: Stephan Stieberger

Abstract

We revisit the issue of defining the entropy of a spatial region in a broad class of quantum theories. In theories with explicit regularizations, working within an elementary but general algebraic framework applicable to matter and gauge theories alike, we give precise path integral expressions for three known types of entanglement entropy that we call full, distillable, and gauge-invariant. For a class of gauge theories that do not necessarily have a regularization in our framework, including Chern-Simons theory, we describe a related approach to defining entropies based on locally extending the Hilbert space at the entangling edge, and we discuss its connections to other calculational prescriptions. Based on results from both approaches, we conjecture that it is always the full entanglement entropy that is calculated by standard holographic techniques in strongly coupled conformal theories.

© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP³.

Contents

1. Introduction and summary	2
2. Formal preliminaries	4
3. Algebras and boundary conditions	7
3.1. Spin chains and scalar matter	8
3.2. Fermions	9

* Corresponding author.

E-mail addresses: jenlin@ias.edu (J. Lin), djordje@brandeis.edu (Đ. Radičević).

¹ Current Institution: Brandeis University, Waltham, MA, USA.

3.3.	Abelian gauge theories	11
3.4.	Nonabelian gauge theories	12
4.	Subalgebras and boundary conditions	14
5.	Entropies of reduced density matrices	16
6.	From subalgebras to path integrals	18
7.	Miscellaneous comments	22
7.1.	Gauge vs. matter theories	22
7.2.	The edge mode action	22
7.3.	Lattice, continuum, and conformal boundary conditions	23
7.4.	Universal terms	24
7.5.	Holographic entanglement entropy	25
8.	Local extensions of the Hilbert space in gauge theories	26
8.1.	Definition	27
8.1.1.	On the lattice	27
8.1.2.	In the continuum	29
8.2.	Path integral boundary conditions	31
8.2.1.	The replica trick in solvable gauge theories	31
8.2.2.	Conformal anomaly for a $U(1)$ gauge theory	32
8.2.3.	Kabat contact term in the geometric entropy	33
8.3.	Short remarks	34
8.3.1.	Emergence	34
8.3.2.	Duality	34
9.	Conclusion	34
	CRedit authorship contribution statement	35
	Declaration of competing interest	35
	Acknowledgements	35
	References	35

1. Introduction and summary

Studying the structure of entanglement in quantum field theory and many-body systems has long been the focus of a tremendous amount of activity.² One of the most fundamental measures of entanglement in such systems is the entropy associated to a spatial subregion. This paper is fully devoted to exploring certain subtle issues that arise in defining this quantity.

It is common — but imprecise — to define the entanglement entropy $S_{\mathbb{V}}$ of a spatial region \mathbb{V} in a quantum state ρ as the von Neumann entropy, $-\text{Tr} \rho_{\mathbb{V}} \log \rho_{\mathbb{V}}$, of the reduced density matrix $\rho_{\mathbb{V}}$ obtained by tracing out the degrees of freedom outside of \mathbb{V} , i.e. as $\rho_{\mathbb{V}} = \text{Tr}_{\bar{\mathbb{V}}} \rho$. The issue is that this definition assumes that the Hilbert space factorizes into a direct product of degrees of freedom on \mathbb{V} and its complement $\bar{\mathbb{V}}$. This assumption fails in many cases, including in strictly continuum QFTs (see the remarkably clear note [11] and references within) and in any lattice gauge theory [12–18].

A more precise approach is to define $S_{\mathbb{V}}$ as the entropy of an algebra of operators $\mathcal{A}_{\mathbb{V}}$ supported on \mathbb{V} (see [19,20] and references therein). If the Hilbert space does factorize and if $\mathcal{A}_{\mathbb{V}}$ is the *maximal* algebra supported on \mathbb{V} , its entropy agrees with the one defined via tracing out.

² See, for instance, [1–7] for a very incomplete sampling of important insights; see [8] for a venerable review, [9] for a modern QFT-oriented review, and [10] for a succinct and prescient summary of many relevant ideas.

If the Hilbert space does not factorize but the theory is fully regularized (e.g. it is defined on a finite lattice), the algebraic definition of entropy is meaningful even though tracing out is not defined. If the theory is defined directly in the continuum, the entropy associated to a subalgebra is not definable — but related quantities, such as relative entropy, will still be well-defined algebraically.

A catch in the above definition is that there is no unique algebra supported on a given region. For instance, even if the Hilbert space factorizes, if $\mathcal{A}_{\mathbb{V}}$ is not the maximal algebra on \mathbb{V} , the associated entropy will not generically agree with the tracing-out prescription. Moreover, for every non-maximal algebra there will be several natural entropic measures that can be associated to it. This is well known and has been discussed in, for example, refs. [14,21].

In this paper we will explore these algebraic choices in more generality. We will focus on theories with explicit regularizations: spin systems, fermion and scalar lattice field theories, and lattice gauge theories. (It is perhaps worth stressing that all of these theories require similar choices in their definitions of entropies; the issues we explore are *not* fundamentally due to the non-factorizability of gauge theory Hilbert spaces.) Our goal is to clarify *which* entanglement entropy one talks about when computing it using various prescriptions that exist in the literature. Many of the ideas here are not new; here we publicize them further and show how they are connected to each other.

Our main focus will be on replica trick path integrals, which are perhaps the most useful way of computing entanglement entropy in field theories. One of our main results is an explicit dictionary that, for a fixed entangling region \mathbb{V} , translates between certain choices of algebras and boundary conditions in these path integrals. (Basic aspects of this connection were presented in [22].) We will identify algebras that correspond to open, Dirichlet, and Neumann boundary conditions on the entanglement edge $\partial\mathbb{V}$ in path integrals with matter fields. In gauge theories, we will similarly identify algebras corresponding to natural gauge-invariant boundary conditions that we will call electric and magnetic. All of these boundary conditions are imposed on the entanglement edge only on a *single* time slice, and at other times the fields are unconstrained.

We will also demonstrate several new points. One is that when a particular choice of algebra $\mathcal{A}_{\mathbb{V}}$ corresponds to a particular type of a boundary condition, then it generically corresponds to a combination of *all* possible conditions of that type. For instance, an algebra corresponding to Dirichlet conditions for a scalar field $\varphi(x)$ will have a natural entropy calculated by a (weighted) sum over all possible boundary conditions $\{k(x_{\parallel})\}$ of the type $\lim_{x \rightarrow x_{\parallel}} \varphi(x) = k(x_{\parallel})$ for $x_{\parallel} \in \partial\mathbb{V}$. Another important lesson will be that the choices of algebras have relatively little to do with the entangling edge itself. While some choices of $\mathcal{A}_{\mathbb{V}}$ will correspond to boundary conditions of various sorts on $\partial\mathbb{V}$, other choices will correspond to “boundary” conditions (or constraints) in the interior of \mathbb{V} . The choices in the definition of entanglement entropy that we discuss here are thus much more general than the ambiguities in introducing a regulator for the conical defect in the replica path integral, and our analysis holds even in these more general cases.

The analysis described so far, being rather microscopic and micromanaging in nature, will not apply to theories that do not have known lattice discretizations. Examples include many field theories of interest, including chiral theories in even spacetime dimensions and Chern-Simons theory. However, a different prescription for calculating entanglement entropy allows us to say a bit about some of these theories, too [12,13,17,18,23,24]. The procedure in question involves local extensions of the Hilbert space that make it factorize on the entanglement edge: instead of regularizing these theories fully, we add the minimal amount of degrees of freedom needed to regularize the entangling edge alone. This allows us to apply the original tracing-out prescription to e.g. Chern-Simons theory, and to draw some instructive parallels to the algebraic approach.

Having thus whetted the reader's appetite, we overview the structure of the paper. In section 2, we set the notation and carefully review some dry (but important) facts about finite systems. We set up the analysis so that it works for *arbitrary* lattice theories, so our results will have obvious analogs in continuum scalar, matter, and gauge theories with known actions and field content.

In section 3, we connect different operator algebras on spatial lattices \mathbb{M} to different types of boundary conditions on boundaries $\partial\mathbb{M}$ of these lattices. This is a warm-up: we focus on pure states, and there are no entangling regions, reduced density matrices, replica tricks, etc. This analysis is also important unto itself, e.g. for the purposes of analyzing exact dualities of quantum theories on manifolds with boundaries.

In section 4 we perform a similar analysis for algebras on subregions $\mathbb{V} \subset \mathbb{M}$. The only difference between scalar and gauge theories arises at this step: for scalar theories, maximal subalgebras will never have a center, while for gauge theories, maximal gauge-invariant subalgebras always will. We will establish how analyzing the center of a chosen algebra $\mathcal{A}_{\mathbb{V}}$ tells us all we need to know about the type of boundary condition that will be obeyed by the appropriate reduced density matrix.

In section 5 we introduce several measures of entanglement that can be associated to the reduced density matrices constructed so far. In section 6 we show how these entropies are computed using replica trick path integrals, and we demonstrate how choices of algebras are reflected on the path integral side. In these sections we will focus on three natural types of entropies: the *full* entropy, the *distillable* entropy, and the *gauge-invariant* entropy. (These will be three out of many possible entropies one can define.) We will describe how they are related to each other and give precise path integral expressions for each of them.

In section 7 we present several important remarks that connect our analysis to other approaches of calculating the entanglement entropy. For instance, we comment on how choices of algebras in conformal field theories are reflected by conformal boundary conditions in the 2D Ising model, based on the analysis in [25]. In the context of CFTs more generally, we also argue that the full and gauge-invariant entanglement entropies are natural candidates for the types of entropies that may contain universal information on trace anomalies, meaning that either *could* agree with the holographic entanglement entropy of Ryu and Takayanagi [5].

In section 8 we return to the discussion of entropies in theories where a lattice realization is unknown, where we employ local extensions of the Hilbert space. We review a cross-section of the existing literature on entanglement entropy in gauge theories, present some archetypical calculations using the extended Hilbert space, and suggest how they fit into our general paradigm. Building on this, we extend our earlier arguments and conjecture that it is precisely the *full* entanglement entropy that contains the universal CFT information obtained using holography and other techniques. We conclude and stress a few open problems in section 9.

2. Formal preliminaries

Consider a Hilbert space \mathcal{H} composed of N copies of a K -dimensional *target space* \mathcal{H}_0 ,

$$\mathcal{H} = \bigotimes_{i=1}^N \mathcal{H}_i \simeq \mathcal{H}_0^{\otimes N}, \quad D \equiv \dim \mathcal{H} = K^N. \quad (1)$$

The use of a regular direct product means that this system is bosonic. For fermionic theories we should use graded products. Instead, we exploit the fact that fermionic systems can be dualized to bosonic systems in which the target has dimension $K = 2$; this is possible even in higher

dimensions, with only mild assumptions on the regularity of spatial lattices [26,27]. This way we can always work with a Hilbert space \mathcal{H} defined as a conventional direct product.

The maximal algebra of operators that acts on this space is the algebra of complex matrices $\mathbb{C}^{D \times D}$. Recall that an algebra is a vector space (in this case, over \mathbb{C}) equipped with and closed under a vector product (in this case, matrix multiplication). This maximal algebra has D^2 complex dimensions. It is very convenient to express the basis as the set of all possible products of a small number of generators. Due to the direct product structure of \mathcal{H} , it is natural to pick the basis of $\mathbb{C}^{D \times D}$ to be generated by operators of the form

$$\mathbb{1}_1 \otimes \mathbb{1}_2 \otimes \dots \otimes \mathbb{1}_{i-1} \otimes \mathcal{O}_i \otimes \mathbb{1}_{i+1} \otimes \dots \otimes \mathbb{1}_N, \quad (2)$$

where \mathcal{O}_i is an operator that generates a basis of $\mathbb{C}^{K \times K}$, the maximal algebra on the i 'th target space. For brevity, we will drop the \otimes signs and factors of $\mathbb{1}$, so \mathcal{O}_i will denote the entire direct product (2).

Only two generators are needed to generate the full basis of \mathcal{H}_i , and we will call them the generalized position and momentum generators Φ_i and Π_i at location i . One simple choice is

$$\Phi_i = \begin{bmatrix} 1 & & & \\ & e^{2\pi i/K} & & \\ & & \ddots & \\ & & & e^{2\pi i(K-1)/K} \end{bmatrix}, \quad \Pi_i = \begin{bmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & \ddots & \end{bmatrix}. \quad (3)$$

For a clock model (a discretized version of the compact scalar theory) these are actual position and momentum operators. A particular Hamiltonian may describe theories very different from this one, e.g. a σ -model whose target space is a (discretization of) a curved manifold, but any such theory will have an operator algebra isomorphic to the one generated by the above Φ_i and Π_i .

From now on, “generators of an algebra” will refer to a set of operators \mathcal{G} whose all possible products form a basis that spans the entire algebra. For a general case of N sites with a K -dimensional target space on each site, the canonical choice for generators of $\mathbb{C}^{D \times D}$ is $\mathcal{G} = \{\Phi_i, \Pi_i\}_{i=1, \dots, N}$. For a spin system, the target space has $K = 2$ and the canonical generators from eq. (3) are the Pauli matrices, $\Phi_i = Z_i$ and $\Pi_i = X_i$. For spinless complex fermions, one typically chooses $\mathcal{G}_{\text{ferm}} = \{\psi_i, \psi_i^\dagger\}_{i=1, \dots, N}$, with $\{\psi_i, \psi_i^\dagger\} \propto \mathbb{1}$ and $\{\psi_i, \psi_j\} = \{\psi_i, \psi_j^\dagger\} = 0$. The algebra generated by $\mathcal{G}_{\text{ferm}}$ is isomorphic to that of a $K = 2$ bosonic system.³

A pure state $|\psi\rangle \in \mathcal{H}$ induces expectation values $\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle$ of all operators in $\mathbb{C}^{D \times D}$. More generally, a density matrix is a Hermitian operator $\rho \in \mathbb{C}^{D \times D}$ that induces expectation values via

$$\langle \mathcal{O} \rangle = \text{Tr}(\rho \mathcal{O}). \quad (4)$$

Conversely, any set of expectation values induces a density matrix. To see this, it is very useful to choose an orthonormal basis $\{\mathcal{O}_a\}_{a=1, \dots, D^2}$ for $\mathbb{C}^{D \times D}$ where all \mathcal{O}_a are invertible and, except the identity, traceless. The orthonormality is defined with respect to the natural trace inner product, such that $\text{Tr}(\mathcal{O}_a^{-1} \mathcal{O}_b) = D \delta_{ab}$ for all a, b . One example is the canonical basis generated by Φ_i and Π_i in eq. (3). The explicit expansion of the density matrix in any such basis is

³ This isomorphism can be made very explicit in one spatial dimension, where it is known as the Jordan-Wigner transformation. It maps a chain of fermions ψ_v for $v = 1, \dots, N$ to a spin chain, via $\psi_v = (X_v + iY_v) \prod_{u=1}^{v-1} Z_u$.

$$\rho = \frac{1}{D} \sum_{a=1}^{D^2} \left\langle \mathcal{O}_a^{-1} \right\rangle \mathcal{O}_a. \quad (5)$$

Let us now consider non-maximal algebras $\mathcal{A} \subset \mathbb{C}^{D \times D}$. For our purposes, the relevant objects are unital $*$ -algebras \mathcal{A} . These are algebras that contain the identity and that are closed under Hermitian conjugation.⁴ In practice, we will only look at non-maximal algebras generated by different products of generalized positions and momenta Φ_i and Π_i , e.g. $\{\Phi_i \Phi_{i+1}, \Pi_i\}_{i=1, \dots, N}$ or $\{\Phi_i, \Pi_i\}_{i=1, \dots, M}$ for $M < N$. Not only will algebras thus obtained all be unital and $*$, the bases generated by these generators will also automatically be orthonormal, with all basis operators traceless (except the identity) and invertible.

Given a non-maximal, unital $*$ -algebra \mathcal{A} spanned by a basis $\{\mathcal{O}_{a'}\}$ of the kind just described, there exists a unique density matrix $\rho \in \mathcal{A}$ that reproduces any list of specified expectations of elements of \mathcal{A} via $\rho = \frac{1}{D} \sum_{a'} \left\langle \mathcal{O}_{a'}^{-1} \right\rangle \mathcal{O}_{a'}$. At this level of generality, ρ is still a $D \times D$ matrix, i.e. an operator acting on the full Hilbert space \mathcal{H} .

An important fact is that any operator that is *not* in \mathcal{A} and has zero inner product with all operators in \mathcal{A} will have a vanishing expectation in state ρ . This is an abstract form of Elitzur's theorem [28]. A reduction of the algebra from $\mathbb{C}^{D \times D}$ to \mathcal{A} is intimately related to gauging degrees of freedom, i.e. to imposing constraints on the space of allowed density matrices.

As a concrete and nontrivial example of these ideas, let us explore the connection between non-maximal algebras and gauging in the context of pure \mathbb{Z}_2 gauge theories in $d = 2$ spatial dimensions.⁵ This will also set the stage for our discussion of gauge theories in general. Here we start with the full (ungauged) Hilbert space \mathcal{H} , given as in eq. (1) by a tensor product of a $K = 2$ bosonic Hilbert space over each link ℓ of a lattice \mathbb{M} . This is the setup of Kitaev's toric code [29]: no gauge constraint has been introduced yet.⁶ The maximal algebra of observables is canonically generated by Pauli matrices on links, $\{Z_\ell, X_\ell\}$. Now, consider the algebra \mathcal{A} generated by $\{X_\ell, W_f\}$, where $W_f = \prod_{\ell \in f} Z_\ell$ are products of Z operators along faces (plaquettes) f .⁷ This algebra has a huge center generated by Gauss operators $G_v = \prod_{\ell \supset v} X_\ell$, one per vertex v of the lattice. Elements of \mathcal{A} are precisely the familiar *gauge-invariant* operators.

⁴ An algebra encodes the allowed operations that can be performed on a given quantum system. We require that \mathcal{A} be unital because we should always be allowed to leave the system as it is, by acting on it with the identity. We require that \mathcal{A} be a $*$ -algebra because ρ is necessarily Hermitian, so operators \mathcal{O} that appear in the expansion (5) of ρ must either be Hermitian themselves, or must come in pairs of Hermitian conjugates. If an operator is in \mathcal{A} but its conjugate is not, then it cannot appear in (5). Any density matrix in \mathcal{A} thus also belongs to the maximal subalgebra of \mathcal{A} that is a $*$ -algebra. Therefore the entire discussion may be phrased from the outset for $*$ -algebras alone.

⁵ Henceforth we always use the lowercase d for the number of spatial dimensions.

⁶ From a traditional gauge theory perspective, one could refer to \mathcal{H} as a (*globally*) *extended Hilbert space*, the understanding being that only the gauge-invariant states are physical and that \mathcal{H} simply provides a convenient embedding for the space of gauge-invariant states. The distinction between gauge-invariant and full/globally extended Hilbert spaces does not influence any correlation functions of the gauge theory but does appear in calculations of entanglement entropy, leading us to distinguish between gauge-invariant and full entropies in section 5. In sections 2 through 7 we will use the globally extended Hilbert space whenever discussing lattice gauge theories, without further comment. In section 8 we will say more about this choice and define measures of entanglement referring to only the gauge-invariant Hilbert space. These entropies will be equivalent to what we call full and gauge-invariant entropies when an explicit regularization is available.

⁷ On lattices of nontrivial topology, one may add to \mathcal{A} gauge-invariant products of operators along noncontractible cycles. If we do not, \mathcal{A} will contain a set of central generators associated to one-form symmetries. These operators are crucial for understanding dualities, for instance, but we will ignore them here.

The fact that the gauge-invariant algebra has a center means that any density matrix $\rho \in \mathcal{A}$ must be block-diagonal in a Hilbert space basis that diagonalizes the center. In other words, the original Hilbert space splits into superselection sectors, one for each of the two possible eigenvalues ± 1 of each G_v . For a connected lattice with N vertices and no special boundary conditions, this means that there are 2^{N-1} sectors (the product of G_v over all sites is identically the identity, hence there is one central element fewer than would be naively expected). For every state whose density matrix is gauge-invariant, the expectation of any operator orthogonal to \mathcal{A} — say, Z_ℓ — is zero. This is the original Elitzur's theorem.

In a \mathbb{Z}_2 gauge theory, the Gauss law $G_v = \mathbb{1}$ is usually imposed at each vertex. (Sometimes background sources are inserted, meaning that $G_v = -\mathbb{1}$ is imposed on a finite number of sites.) This operator equation is tantamount to restricting our attention to just one superselection sector of the full Hilbert space \mathcal{H} . (Note that Elitzur's theorem holds even without this restriction to a single sector.) In the toric code, and more generally in condensed matter gauge theories, these gauge constraints are imposed dynamically, by adding $-\sum_v G_v$ to the Hamiltonian and then focusing on the low-energy sector; in particle physics gauge theories, this constraint is imposed by fiat.

The gauge theory example is an instance of a more general phenomenon: a non-maximal algebra always has an associated center, and eigenvalues of generators of the center label different superselection sectors. All density matrices are direct sums (statistical mixtures) of smaller density matrices that act only within specific sectors. Sometimes, as with gauge-invariant algebras and states obeying the Gauss law in the preceding paragraph, only one superselection sector will be populated. At other times, as with more general subalgebras and states to be discussed below, multiple sectors will be populated. In those cases, within each sector, its labels can be interpreted as constraints. When the center generators have support only on the edges of a spatial lattice, these sector labels denote different boundary conditions and are called *edge modes*.

3. Algebras and boundary conditions

Consider a system defined on a spatial lattice \mathbb{M} with boundary.⁸ There exist several families of algebras that are supported on all of \mathbb{M} , but that differ in the choice of generators present at the boundary $\partial\mathbb{M}$. The purpose of this section is to show which families naturally correspond to which familiar sets of (lattice) boundary conditions.

As a warm-up, we start with a spin chain with N sites. Here we will study families of algebras obtained by removing edge generators from the maximal algebra $\mathbb{C}^{D \times D}$ with $D = 2^N$. The generalization to higher dimensions and arbitrary compact scalar theories is immediate.⁹ Fermion systems in one spatial dimension will be examined next, and they also generalize to higher dimensions and parafermionic models. Finally, we will show how edge modes arise in gauge theories, where subalgebras of the non-maximal, gauge-invariant algebra \mathcal{A} are considered.

⁸ The boundary of a 2d lattice is the set of all links that do not belong to exactly two faces. On a general triangulated d -manifold, the boundary is the set of all codimension-one simplices which do not belong to two codimension-zero simplices, together with any nonzero-codimension simplices that do not belong to any codimension-zero simplices.

⁹ Noncompact scalars, on the other hand, do not have an algebraic formulation unless they are embedded into a compact scalar theory. We will only discuss compact target spaces in this paper.

3.1. Spin chains and scalar matter

A spin chain has a maximal algebra generated by Pauli operators X_v and Z_v on each site v . (An analogous discussion applies if we start from other generators, say X_v and Y_v .) Consider now the algebra $\mathcal{A}_{D,O}$ generated by

$$\mathcal{G}_{D,O} = \{X_1, \dots, X_N, Z_2, \dots, Z_N\}. \quad (6)$$

The difference from the maximal algebra is that one generator from the edge of the system, Z_1 , is removed. The center of $\mathcal{A}_{D,O}$ is generated by X_1 , the remaining generator on the edge. All pure states with density matrices in this reduced algebra thus must be eigenstates of X_1 . Moreover, a general density matrix will be a statistical mixture of $X_1 = \mathbb{1}$ and $X_1 = -\mathbb{1}$ states. Thus algebras missing one generator on the edge correspond to (mixtures of) states with definite values of spins at that edge. In the case of $\mathcal{A}_{D,O}$, there are thus Dirichlet boundary conditions on one edge, and open boundary conditions on the other edge.

We can also consider “coarse graining” generators on two adjacent sites near the edge, so the remaining ones are only sensitive to some (but not all) degrees of freedom near $\partial\mathbb{M}$. For instance, take an algebra that cannot measure the z -component (magnetization) of individual spins at sites 1 and 2, but that can measure the parity of the *total* magnetization on these sites. This algebra is generated by

$$\mathcal{G}_{N,O} = \{X_1, \dots, X_N, Z_1 Z_2, Z_3, Z_4, \dots, Z_N\}. \quad (7)$$

The resulting algebra, $\mathcal{A}_{N,O}$, has the same number of generators as the Dirichlet-open one, $\mathcal{A}_{D,O}$. Its center is generated by $X_1 X_2$, and so states whose density matrices belong to this algebra are (statistical mixtures of) eigenstates of $X_1 X_2$. If eigenvalues of X_v are labeled by $e^{i\pi\phi_v}$ with $\phi_v \in \mathbb{Z} \bmod 2$, then this algebra contains density matrices of states with definite values of $\phi_1 - \phi_2 \sim \partial\phi_1$. This means that Neumann boundary conditions are imposed on one end of the chain, while the other end has open boundary conditions. (See Fig. 1.)

It is straightforward to find algebras that correspond to other combinations of open, Neumann, and Dirichlet conditions on edges of the system. For instance, Dirichlet conditions on both edges correspond to the algebra generated by

$$\mathcal{G}_{D,D} = \{X_1, \dots, X_N, Z_2, \dots, Z_{N-1}\}. \quad (8)$$

It is also possible to choose algebras where different fields are subject to boundary conditions at the edges. Consider the generating set

$$\mathcal{G}_{D,\dot{D}} = \{X_1, \dots, X_{N-1}, Z_2, \dots, Z_N\}. \quad (9)$$

The corresponding algebra has a center generated by X_1 and Z_N , so the density matrices in it correspond to states with definite values of ϕ_1 and ϕ_N . This notation is inspired by the fact that Z_N is the momentum operator conjugate to the position X_N , so its eigenvalues, schematically, correspond to time derivatives of the field ϕ_N at the edge.

The same considerations apply in higher dimensions. Different pairing up of generators on the boundary can be performed to obtain algebras that correspond to fixed derivatives of fields parallel or perpendicular to the boundary. (Links are perpendicular to $\partial\mathbb{M}$ if one of their vertices is on $\partial\mathbb{M}$ and the other one is not.) It is also straightforward to generalize to bosonic systems with more complicated target spaces: all the operators Z_v and X_v can be replaced by Φ_v and Π_v from eq. (3) or by their inverses, as appropriate.



Fig. 1. Depictions of generating operators for various choices of algebras for spin chains discussed in the text. Algebras generated by these sets are all supported on the entire lattice, i.e. there are no sites on which all operators act trivially. Red indicates locations of central generators. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

3.2. Fermions

The examples given in this section will focus on fermionic systems in one spatial dimension. As with spin systems, our conclusions can be generalized to higher dimensions, though there will be some additional subtleties we will discuss later. Consider any theory of complex fermions in $d = 1$, defined on N sites labeled $v = 1, \dots, N$. The maximal algebra is generated by $\{\psi_v, \psi_v^\dagger\}_{v=1, \dots, N}$, but unlike the spin operators X_v and Z_v , the fermionic basis operators neither form a group nor are they all traceless — and, most importantly, the generators on different sites all anticommute. This system can, however, be dualized to a bosonic spin system via the *nonlocal* Jordan-Wigner map. By thus bosonizing the fermions, the machinery from the last section can be imported wholesale.

This procedure might seem cumbersome, but matters are greatly simplified by the requirement that the Hamiltonian always be bosonic, i.e. that all fermionic systems have a fermion parity symmetry, generated by the product of operators $2\psi_v^\dagger\psi_v - 1$ over all sites. If our attention is restricted to states of definite fermion parity and not to their superpositions, then only operators made out of an even number of fermionic generators ψ_v and ψ_v^\dagger will have nonzero expectations. In other words, the algebra of operators for fermionic systems with conserved fermion parity can be taken to be the non-maximal one generated by fermion bilinears $\psi_v^\dagger\psi_v$, $\psi_v\psi_{v+1}$, and so on. These operators map *locally* under the Jordan-Wigner transformation, even in higher dimensions [26,27], and the dual algebra (acting on the bosonic Hilbert space of Ising spins) is generated by

$$\mathcal{G}^f = \{X_1X_2, X_2X_3, \dots, X_{N-1}X_N, Z_1, \dots, Z_N\}. \quad (10)$$

We will refer to this algebra as the *maximal fermionic algebra* on N sites, and we will assume that we only work with states of definite fermion parity. The center of this algebra is generated by the fermion parity operator, which is $\prod_{v=1}^N Z_v$ in the bosonic language. (See Fig. 2.)

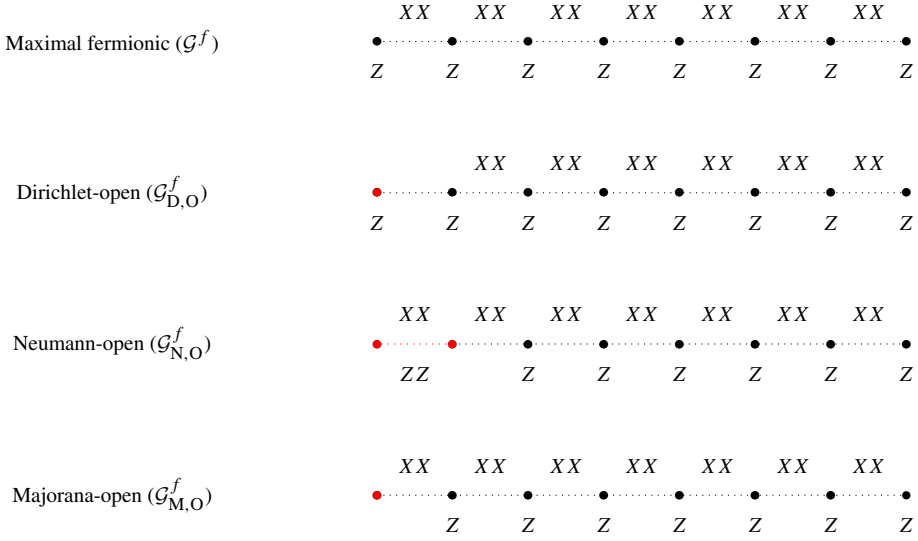


Fig. 2. Generators of fermionic algebras discussed in the text, presented using their bosonized equivalents. Each algebra commutes with the product of all Z 's, i.e. with the total fermion parity $(-1)^F$. Locations of all *additional* central generators are red. In the case of Majorana boundary conditions, the removal of Z_1 is taken to mean that $(-1)^F$ is still a central generator, but that an overall constraint is imposed on superselection sector weights such that $(-1)^F = 0$ hold as an operator equation.

What edge modes or boundary conditions can be imposed here? A natural first step is the removal of the operator $X_1 X_2$ from \mathcal{G}^f , resulting in the generating set

$$\mathcal{G}_{D,O}^f = \{X_2 X_3, \dots, X_{N-1} X_N, Z_1, \dots, Z_N\}. \quad (11)$$

This adds another central generator, Z_1 , to the existing one (the fermion parity). The extra superselection sectors that ensue are labeled by the number of fermions on the edge. This is the fermionic dual to Dirichlet boundary conditions.

The other natural alternative is

$$\mathcal{G}_{N,O}^f = \{X_1 X_2, \dots, X_{N-1} X_N, Z_1 Z_2, Z_3, \dots, Z_N\}. \quad (12)$$

The additional central generator here is $X_1 X_2$, just like in the Neumann boundary condition for spins. In the fermionic language this central generator is rather nontrivial: it corresponds to $(\psi_1^\dagger - \psi_1)(\psi_2^\dagger + \psi_2)$. However, in terms of Majorana fermions χ_v and χ'_v , defined via $\psi_v = \chi_v + i\chi'_v$, this central generator is simply the Majorana hopping operator $i\chi'_1 \chi_2$. (In this Majorana basis, the central generator corresponding to Dirichlet conditions is $i\chi'_1 \chi_1$.)

It is interesting to examine the generating set obtained by removing Z_1 from \mathcal{G}^f ,

$$\mathcal{G}_{M,O}^f = \{X_1 X_2, \dots, X_{N-1} X_N, Z_2, Z_3, \dots, Z_N\}. \quad (13)$$

This is a natural object in terms of Majorana generators χ_v and χ'_v introduced in the preceding paragraph, as the corresponding algebra is generated by all Majorana bilinears built *without* χ_1 .¹⁰

¹⁰ It is also possible to remove bilinears involving χ'_1 from the generating set, which corresponds to removing both $X_1 X_2$ and Z_1 from \mathcal{G}^f . However, the resulting algebra is not supported on all sites — there are no nontrivial operators acting on site 1 — so we do not consider this case here.

This algebra has no center. Moreover, it does not contain the fermion parity operator $(-1)^F$. Thus, it only describes states with $(-1)^F = 0$ and so it cannot contain the density matrix of any state with definite fermion parity; it can only contain equal statistical mixes of states with $(-1)^F = \mathbb{1}$ and $(-1)^F = -\mathbb{1}$. We can say that the algebra generated by $\mathcal{G}_{\text{M},\text{O}}^f$ corresponds to “Majorana boundary conditions.” We will return to such kinds of boundary conditions when discussing algebras on subregions.

The notion of Dirichlet and Neumann subalgebras of the maximal fermionic algebra can be extended to higher dimensions. The procedure is the same in spirit as the one outlined above: the algebra of fermion bilinears is dualized to a (non-maximal) subalgebra acting on a \mathbb{Z}_2 bosonic system, and appropriate subalgebras with additional central generators at the edges are then considered. The complication is that duals of fermionic systems in higher dimensions are gauge theories with non-standard Gauss laws [26,27]. Nevertheless, the fermion bilinears all dualize into local operators, and the procedure is conceptually straightforward and merely technically involved.

3.3. Abelian gauge theories

Let us now consider a \mathbb{Z}_2 gauge theory in $d = 2$. (Generalizations to \mathbb{Z}_K and $U(1)$ theories are straightforward.) To recap, the full Hilbert space is a product of $K = 2$ spaces on each link ℓ of the lattice \mathbb{M} . The algebra \mathcal{A} of gauge-invariant operators, described in section 2, is generated by electric fields X_ℓ on links, magnetic fields W_f on faces, and Wilson loops W_c along homologically distinct noncontractible 1-cycles. Pure density matrices in \mathcal{A} must describe eigenstates of Gauss operators $G_v = \prod_{\ell \supset v} X_\ell$ on vertices, as these operators generate the center of \mathcal{A} . Gauge-invariant states are those with $G_v = \mathbb{1}$ for all v .

What are natural ways to reduce this algebra by removing operators at the edges of the system? A subalgebra of \mathcal{A} without some of the Gauss operators will not contain density matrices of any pure gauge-invariant states; it will necessarily only contain mixtures of states with $G_v = \mathbb{1}$ and $G_v = -\mathbb{1}$ for those Gauss operators G_v absent from the subalgebra. This is analogous to what happened with fermionic subalgebras that did not contain the fermion parity operator. If our goal is to formulate gauge theory subalgebras that correspond to boundary conditions on *pure* states, the subalgebras must then contain all Gauss operators. Appropriate subalgebras will thus be obtained either by removing magnetic generators or by removing electric generators while keeping at least those products of theirs that form Gauss operators.

Consider first the generating set $\mathcal{G}_{\text{E}(\ell_1)}$ obtained by removing the magnetic field operator W_{f_1} from a plaquette f_1 that contains the link $\ell_1 \in \partial\mathbb{M}$, see Fig. 3. The removal of W_{f_1} causes the operator X_{ℓ_1} to become a central generator. Thus pure states whose density matrices are in $\mathcal{A}_{\text{E}(\ell_1)}$ have a definite electric field on the boundary link ℓ_1 . We can similarly remove magnetic field operators W_{f_2}, W_{f_3}, \dots from other plaquettes that contain boundary links ℓ_2, ℓ_3 , and so on. If all the magnetic operators from boundary faces are removed, we will say that the remaining generators generate the *electric algebra*, and will denote it by \mathcal{A}_{E} . This algebra contains density matrices of all states with definite electric fields at the boundary. In the more common continuum notation, these are states with definite values of the electric field E_{\parallel} parallel to the boundary. Note that because of the Gauss law at each boundary site, knowing E_{\parallel} as a function of the position along the boundary is the same as knowing E_{\perp} , the electric field flowing into the boundary at each boundary site.

The analysis of what happens when certain electric field operators are removed from \mathcal{A} is a bit more involved. The simplest way to obtain a nontrivial subalgebra that contains all the boundary

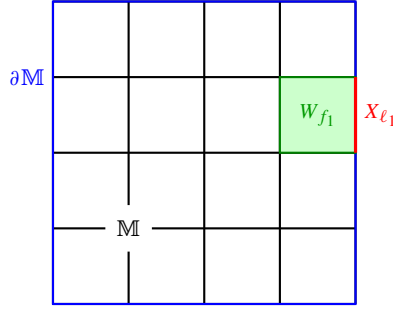


Fig. 3. A \mathbb{Z}_2 gauge theory on a lattice \mathbb{M} with boundary $\partial\mathbb{M}$ (drawn in blue). There is a central generator G_v at every vertex of this lattice, including on edge sites $v \in \partial\mathbb{M}$. Boundary plaquettes are those faces that contain boundary links; for instance f_1 contains a boundary link $\ell_1 \in \partial\mathbb{M}$. When W_{f_1} is removed from the algebra, X_{ℓ_1} becomes another central generator. It corresponds to the electric field operator parallel to the boundary.

Gauss operators is to consider the generating set \mathcal{G}_M that contains none of the individual X_ℓ operators on the boundary of the lattice, but that does contain all the Gauss operators G_v for edge sites v . Pure states whose density matrices are in this algebra \mathcal{A}_M are eigenstates of the Wilson loop along the boundary of the region; in continuum notation, this corresponds to a fixed value of $\oint A$ along the boundary. These states hence have definite values of total magnetic flux through the entire spatial surface. We will call the corresponding \mathcal{A}_M the *magnetic algebra*.

Other admissible subalgebras are obtained by removing more electric operators from plaquettes f_i on the boundary in such a manner that every boundary plaquette has only products of an even number of electric operators acting on its links. This way individual magnetic fields W_{f_i} on the boundary can become generators of the center. The pure density matrices in such algebras $\mathcal{A}_{M(f_i)}$ correspond to states with definite magnetic fields at the boundary.

We introduced Neumann and Dirichlet boundary conditions for 1d fermions as Jordan-Wigner duals of those boundary conditions for bosons. Given that 2d Abelian gauge theories can be dualized to bosonic theories via the Kramers-Wannier transform, it may appear natural to compare the electric and magnetic conditions of \mathbb{Z}_2 theories to Dirichlet and Neumann conditions in dual spin systems. However, this is not straightforward. A gauge theory on a lattice with boundary is dual to a spin system only if it contains no operators on the boundary — Gauss operators included [30,31]. A duality between boundary conditions can indeed be established, but the paradigm developed in this section would need to be altered to account for gauge theories without boundary operators. We will not explore this issue here.

Finally, we note that in relativistic gauge theories there exists another terminology for boundary terms. Any boundary condition that fixes the field strength $F_{\mu\nu}$ with μ and ν both parallel to the boundary is a Dirichlet condition, while fixing $F_{\mu\nu}$ with μ or ν perpendicular to the boundary is a Neumann condition. This means that our magnetic conditions can be either Neumann or Dirichlet in $d \geq 3$; in $d = 2$ magnetic conditions can only be Neumann, and in $d = 1$ they are not defined. Our electric boundary conditions can only be Dirichlet in $d \geq 2$ (if we are fixing X_ℓ for $\ell \in \partial\mathbb{M}$) but can always be Neumann (if we are fixing X_ℓ for $\ell \perp \partial\mathbb{M}$). (See Table 1.)

3.4. Nonabelian gauge theories

The direct approach indicated in section 2 is not very convenient for nonabelian gauge theories, so we adjust it as follows [17,32]. We start from a theory in which on each link ℓ we have

Table 1

A summary of correspondences between algebras and boundary conditions in scalar, fermion, and (Abelian) gauge theories. Most common boundary conditions are listed.

	Generators of the algebra	Generators of the center	Boundary conditions
Spin chains	$X_1, \dots, X_N, Z_1, \dots, Z_N$	None	Open
	$X_1, \dots, X_N, Z_2, \dots, Z_{N-1}$	X_1, X_N	Dirichlet
	$X_1, \dots, X_N,$ $Z_1 Z_2, Z_3, \dots, Z_{N-2}, Z_{N-1} Z_N$	$X_1 X_2, X_{N-1} X_N$	Neumann
Fermion chains	$X'_1 X_1, X'_2 X_2, \dots, X'_N X_N,$ $X'_1 X_2, X'_2 X_3, \dots, X'_{N-1} X_N$	$(-1)^F \equiv (i X'_1 X_1) \cdots (i X'_N X_N)$	Open
	$X'_1 X_1, \dots, X'_N X_N,$ $X'_2 X_3, \dots, X'_{N-2} X_{N-1}$	$(-1)^F, X'_1 X_1, X'_N X_N$	Dirichlet
	$X'_1 X_1 X'_2 X_2, X'_{N-1} X_{N-1} X'_N X_N,$ $X'_3 X_3, \dots, X'_{N-2} X_{N-2},$ $X'_1 X_2, X'_2 X_3, \dots, X'_{N-1} X_N$	$(-1)^F, X'_1 X_2, X'_{N-1} X_N$	Neumann
\mathbb{Z}_2 gauge theory (any dimension)	$X_\ell, W_f \equiv \prod_{\ell \subset f} Z_\ell$ for all ℓ, f	$G_v \equiv \prod_{\ell \supset v} X_\ell$ for all v	Open
	All X_ℓ, W_f except $W_f, f \in \partial \mathbb{M}$	All G_v , and X_ℓ for $\ell \in \partial \mathbb{M}$	Electric
	All X_ℓ, W_f except $X_\ell, \ell \in \partial \mathbb{M}$	All G_v , and $W_{\partial \mathbb{M}} \equiv \prod_{\ell \in \partial \mathbb{M}} Z_\ell$	Magnetic
Compact scalar (any dimension)	$\Phi_v \equiv e^{i\phi_v}, \Pi_v$ for all vertices v	None	Open
	All Φ_v, Π_v except Π_v for $v \in \partial \mathbb{M}$	All Φ_v for $v \in \partial \mathbb{M}$	Dirichlet
	All Φ_v , all Π_v in the bulk, only $\Pi_v \Pi_u$ for links $(v, u) \perp \partial \mathbb{M}$	All $\Phi_v^{-1} \Phi_u$ for $(v, u) \perp \partial \mathbb{M}$	Neumann
$U(1)$ gauge theory (any dimension)	$X_\ell \equiv e^{\frac{\partial}{\partial A_\ell}}, W_f \equiv e^{i \oint_f A}$ for all ℓ, f	$G_v \equiv e^{\oint_v * \frac{\partial}{\partial A}}$ for all v	Open
	All X_ℓ, W_f except $W_f, f \in \partial \mathbb{M}$	All G_v , and X_ℓ for $\ell \in \partial \mathbb{M}$	Electric
	All X_ℓ, W_f except $X_\ell, \ell \in \partial \mathbb{M}$	All G_v , and W_f for $f \in \partial \mathbb{M}$	Magnetic

a Hilbert space \mathcal{H}_ℓ spanned by $\{|U\rangle_\ell\}$ for all U in a given gauge group G . Now, on each link we define an infinite series of finite-dimensional Hilbert spaces, \mathcal{H}_ℓ^R , associated to irreducible representations R of G . A basis of \mathcal{H}_ℓ^R is

$$|R_{\alpha\beta}\rangle_\ell \propto \int_G dU R_{\alpha\beta}(U) |U\rangle_\ell, \quad \alpha, \beta = 1, \dots, \dim R. \quad (14)$$

Here we omit normalization factors and take dU to be the Haar measure on G . The dimension of \mathcal{H}_ℓ^R is $(\dim R)^2$. The vectors $\{|R_{\alpha\beta}\rangle_\ell\}$ for all R, α, β span \mathcal{H}_ℓ .¹¹

States $|R_{\alpha\beta}\rangle_\ell$ have definite electric flux in representation R , but their products over all ℓ are generally not gauge-invariant. Gauss operators that implement gauge transformations by $\Lambda \in G$ are

$$G_v^\Lambda = \prod_u L_{vu}^\Lambda, \quad (15)$$

where the product goes over all neighboring vertices of v , and $L_{vu}^\Lambda |U\rangle_\ell$ equals $|\Lambda U\rangle_\ell$ or $|U \Lambda^{-1}\rangle_\ell$, depending on whether the link ℓ is oriented from v to u or vice versa.¹² The oper-

¹¹ A finite approximation to a nonabelian gauge theory can be obtained by keeping only representations below a certain dimension, but then the algebra of operators may only approximately close under multiplication.

¹² In terms of states $|R_{\alpha\beta}\rangle$, these two types of operators act on α and β by the matrices $R(\Lambda^{-1})$ and $R^T(\Lambda)$, respectively. For example, $L_\ell^\Lambda |R_{\alpha\beta}\rangle_\ell \propto \int_G dU R_{\alpha\alpha'}(\Lambda^{-1}) R_{\alpha'\beta}(U) |U\rangle_\ell \equiv |R_{\alpha\alpha'}^{-1}(\Lambda) R_{\alpha'\beta}\rangle_\ell$.

ators L_{vu}^Λ are nonabelian generalizations of operators built out of electric flux generators X_ℓ from the previous subsection, but unlike them L_{vu}^Λ are not gauge-invariant. It is common to view G as a Lie group and to employ its generators T^a to write

$$L_{vu}^\Lambda = e^{i\theta^a J_{vu}^a} \quad \text{for} \quad \Lambda = e^{i\theta^a T^a}, \quad (16)$$

with the understanding that the electric flux generators J_{vu}^a act either from left or from right and have sign $+$ or $-$, depending on the orientation of the link [17].

Gauge-invariant states are products of states $|R_{\alpha\beta}\rangle_\ell$ in the same representation along a cycle, with indices α, β, \dots contracted between neighboring links in this cycle. Gauge-invariant operators are Wilson loops W_c^R in representations R on cycles c , and Casimirs built out of electric flux generators J_{uv} defined in eq. (16). The analysis of boundary conditions in non-maximal algebras now proceeds analogously to the Abelian case: removing e.g. all W_f^R from a face f near $\partial\mathbb{V}$ turns all the Casimirs on the boundary link from f into central generators. Conversely, removing all Casimirs and local Wilson loops from $\partial\mathbb{V}$ makes all the $W_{\partial\mathbb{V}}^R$'s become central.

4. Subalgebras and boundary conditions

The analysis so far has dealt with algebras that are supported on the whole lattice \mathbb{M} , and a particular emphasis was placed on understanding which pure density matrices belong to which algebras. In this section we will deal with algebras $\mathcal{A}_\mathbb{V}$ of operators supported on a *subregion* $\mathbb{V} \subset \mathbb{M}$. We assume that we are given a density matrix ρ in the full algebra \mathcal{A} — or, equivalently, a list of expectation values $\langle \mathcal{O} \rangle$ of all basis elements $\mathcal{O} \in \mathcal{A}$. Each pair $(\rho, \mathcal{A}_\mathbb{V})$ uniquely induces a reduced density matrix via an analog of eq. (5),

$$\rho(\mathcal{A}_\mathbb{V}) \equiv \frac{1}{D} \sum_{a'} \left(\mathcal{O}_{a'}^{-1} \right) \mathcal{O}_{a'}, \quad \text{where} \quad \mathcal{A}_\mathbb{V} = \text{span}\{\mathcal{O}_{a'}\}. \quad (17)$$

We seek to understand what kinds of different $\rho(\mathcal{A}_\mathbb{V})$'s can there be, given a state ρ and a region \mathbb{V} .

We will focus on algebras $\mathcal{A}_\mathbb{V}$ that differ by operators at the edges of \mathbb{V} , and the resulting set of different possible $\rho_\mathbb{V}$ will correspond to different boundary conditions that can be put at the entangling cut of the system. This is analogous to section 3. Our analysis applies more broadly, and the same techniques can be applied to algebras $\mathcal{A}_\mathbb{V}$ lacking generators in the interior of \mathbb{V} .

Any reduced density matrix in any subalgebra $\mathcal{A}_\mathbb{V}$ has the form

$$\rho(\mathcal{A}_\mathbb{V}) = \rho_\mathbb{V} \otimes \left(\frac{1}{D_{\bar{\mathbb{V}}}} \mathbb{1}_{\bar{\mathbb{V}}} \right), \quad (18)$$

where $\rho_\mathbb{V}$ is a properly normalized density matrix of a state in $\mathcal{H}_\mathbb{V}$ (the product of all target Hilbert spaces \mathcal{H}_0 associated to elements of \mathbb{V} that carry degrees of freedom), and $\mathbb{1}_{\bar{\mathbb{V}}}$ acts only on the complement $\mathcal{H}_{\bar{\mathbb{V}}}$ of dimension $D_{\bar{\mathbb{V}}}$. This means that the algebra $\mathcal{A}_\mathbb{V}$ only contains density matrices of states that are uniform mixtures of all possible field configurations outside \mathbb{V} . From now on we focus only on $\rho_\mathbb{V}$, for which there are three possibilities:

1. If $\mathcal{A}_\mathbb{V}$ is the maximal algebra $\mathbb{C}^{D_\mathbb{V} \times D_\mathbb{V}}$ on $\mathcal{H}_\mathbb{V}$, there will be open boundary conditions on the edges of the interval, and all density matrices $\rho_\mathbb{V}$ will be allowed.
2. If $\mathcal{A}_\mathbb{V}$ is not maximal but has a center generated by edge operators, $\rho_\mathbb{V}$ will split into supers-election sectors that contain states obeying a specific kind of boundary condition. Examples

include Dirichlet and Neumann conditions from the previous section. Relative weights of different sectors depend on the original state and may take arbitrary values.

3. If $\mathcal{A}_{\mathbb{V}}$ is not maximal and has no center generated by edge operators, we can add a central generator and require that its expectation value always be zero. The reduced density matrix will then split into blocks labeled by eigenvalues of this central generator, but there will be a constraint on the relative weights of superselection sectors that will force $\rho_{\mathbb{V}}$ to be impure. An example is the Majorana boundary condition from the previous section. In this case the field at the boundary is *random*, i.e. it is an equal statistical mixture between states with and without a fermion at the edge.

Boundary conditions may effectively depend on ρ in the following sense. If $\mathcal{A}_{\mathbb{V}}$ is the maximal algebra on \mathbb{V} , and if any operator that fails to commute with a given $\mathcal{O} \in \mathcal{A}_{\mathbb{V}}$ has zero expectation, then $\rho_{\mathbb{V}}$ will be block-diagonal with sectors labeled by eigenvalues of \mathcal{O} . For instance, if Π is a momentum operator conjugate to a position operator Φ , open boundary conditions are the same as Dirichlet conditions in a state where all expectations involving Π vanish (i.e. if $\Pi = 0$). As another example, if $\mathcal{A}_{\mathbb{V}}$ is the Dirichlet algebra of a spin system with Z_v a central generator, and if the state has $Z_v = 0$ for some $v \in \partial\mathbb{V}$, the Dirichlet boundary conditions will be the same as random (Majorana) ones: only a uniform mix of $Z_v = \mathbb{1}$ and $Z_v = -\mathbb{1}$ states will appear in $\rho_{\mathbb{V}}$. When specifying a set of boundary conditions, we will always specify them based on the algebra and not on the particular state in question.

Let us now apply these ideas to our usual test subject, the \mathbb{Z}_2 gauge theory. Gauge invariance has historically caused a lot of confusion when it came in contact with the study of entanglement. By the following analysis we wish to very forcefully point out that even when working with the full Hilbert space the results will be gauge-invariant, as long as only gauge-invariant algebras are studied. These points have been made in [30].

Consider an algebra $\mathcal{A}_{\mathbb{V}}$ of gauge-invariant operators supported on a set \mathbb{V} , which in this case is a collection of links. No operator in $\mathcal{A}_{\mathbb{V}}$ has support outside of \mathbb{V} — in particular, no Gauss operator G_v for $v \in \partial\mathbb{V}$ is in $\mathcal{A}_{\mathbb{V}}$.¹³ The center of $\mathcal{A}_{\mathbb{V}}$ is generated by Gauss operators in the interior of \mathbb{V} and by boundary electric fields

$$E_v \equiv \prod_{\substack{\ell \supset v \\ \ell \in \mathbb{V}}} X_{\ell}. \quad (19)$$

If the original state ρ is gauge-invariant, then all Gauss operators will have $G_v = \mathbb{1}$, and the appropriate reduced density matrix $\rho_{\mathbb{V}}$ will contain only states that satisfy the Gauss law on all the interior points, i.e. it will be of the form

$$\rho_{\mathbb{V}} = \tilde{\rho}_{\mathbb{V}} \prod_{v \in \text{Int}(\mathbb{V})} \frac{\mathbb{1} + G_v}{2}. \quad (20)$$

The remaining central generators of $\mathcal{A}_{\mathbb{V}}$, E_v , ensure that the generic density matrix $\rho_{\mathbb{V}}$ describes a mixture of states with definite electric fluxes through $\partial\mathbb{V}$. This *almost* corresponds to electric boundary conditions on $\partial\mathbb{M}$ as described in the previous section. Recall that our original definition of electric boundary conditions involved specifying all electric fields on links in $\partial\mathbb{M}$. In that situation, all magnetic field generators were removed from plaquettes containing

¹³ We assume that \mathbb{V} is a subregion of \mathbb{M} that is far away from the boundary $\partial\mathbb{M}$ of our spatial lattice, if any exists.

boundary links, and $W_{\partial\mathbb{M}}$, the Wilson loop along the boundary of \mathbb{M} , was not in the algebra of operators dubbed \mathcal{A}_E . Here, however, all magnetic operators on the boundary are retained, and in particular $W_{\partial\mathbb{V}} \in \mathcal{A}_{\mathbb{V}}$. Thus, with electric boundary conditions on physical boundaries of the system, density matrices split into $2^{|\partial\mathbb{V}|}$ blocks; with electric boundary conditions on the entanglement edge, they split into $2^{|\partial\mathbb{V}|-b_0}$ blocks, where b_0 is the number of disconnected components of \mathbb{V} . We will still refer to these boundary conditions as electric, but this caveat must be kept in mind, as the b_0 “correction” gives rise to universal quantities such as the topological entanglement entropy.¹⁴

In nonabelian theories, the center of the gauge-invariant algebra is also given by the Gauss operators in the interior and by a set of operators on the boundary of \mathbb{V} . These boundary operators are defined as Casimirs built out of products over all interior vertices u neighboring $v \in \partial\mathbb{V}$,

$$E_v^\Lambda \equiv \prod_{u \in \mathbb{V}} L_{vu}^\Lambda. \quad (21)$$

How do these operators act in the basis $|R_{\alpha\beta}\rangle_\ell$ from eq. (14)? Within each \mathcal{H}_ℓ^R such that $v \in \ell \in \mathbb{V}$, we have

$$L_{vu}^\Lambda : |R_{\alpha\beta}\rangle_\ell \mapsto |R_{\alpha\alpha'}^{-1}(\Lambda)R_{\alpha'\beta}\rangle_\ell \quad \text{or} \quad |R_{\alpha\beta'}R_{\beta'\beta}(\Lambda)\rangle_\ell, \quad (22)$$

depending on the orientation of ℓ (recall that we are summing over repeated indices, and when the argument of R is dropped it is understood to be the integration variable U appearing in (14)). The product of these operators, E_v^Λ , has an analogous action on the *total* flux in representation R that enters the vertex v [17]. Thus, every central operator E_v^Λ in a nonabelian theory has an identity component when represented on the space $\bigotimes_\ell \mathcal{H}_\ell^R$, and hence the reduced density matrix will contain this product of identities, too. (We will give them a physical interpretation in section 8.) Beyond this, the analysis of nonabelian theories is analogous to the Abelian case: the existence of boundary central generators causes $\rho_{\mathbb{V}}$ to split into superselection sectors labeled by the tuple R , which are direct generalizations of Abelian electric fluxes labeled by k .

5. Entropies of reduced density matrices

The algebraic point of view developed in the previous sections allows us to start from any theory with a known lattice regularization and define the Hilbert space in the natural representation of any subalgebra of operators $\mathcal{A}_{\mathbb{V}}$. Roughly speaking, this space is the direct product $\mathcal{H}_{\mathbb{V}} = \bigotimes_{i \in \mathbb{V}} \mathcal{H}_i$ over all degree-of-freedom-carrying simplices i (vertices, links, faces, etc) on which at least one operator in $\mathcal{A}_{\mathbb{V}}$ acts nontrivially. More precisely, a non-maximal algebra $\mathcal{A}_{\mathbb{V}}$ restricts the possible quantum superpositions between the states in $\mathcal{H}_{\mathbb{V}}$, so the resulting restricted space will be denoted $\mathcal{H}_{\mathbb{V}}^*$. This restricted space $\mathcal{H}_{\mathbb{V}}^*$ has the same dimension as $\mathcal{H}_{\mathbb{V}}$. The only difference between them is that $\mathcal{H}_{\mathbb{V}}^*$ does not allow superpositions of states in different superselection sectors. If k labels these sectors, we can write $\mathcal{H}_{\mathbb{V}}^* = \bigoplus_k \mathcal{H}_{\mathbb{V}}^{(k)}$, where each $\mathcal{H}_{\mathbb{V}}^{(k)}$ is a smaller but more “traditional” kind of Hilbert space where all superpositions are allowed.

¹⁴ Non-maximal gauge-invariant algebras, obtained by removing gauge-invariant operators from the vicinity of $\partial\mathbb{V}$ in a manner described in section 3.3, can also be viewed as having random boundary conditions for some of the central generators E_v . We will not use this point of view in this paper, but it may be a useful fact when comparing entropies of different algebras in a fixed quantum state.

The natural entropy to be associated with $\rho_{\mathbb{V}}$ is the von Neumann entropy,

$$S_{\mathbb{V}} = -\text{Tr}(\rho_{\mathbb{V}} \log \rho_{\mathbb{V}}) = -\frac{\partial}{\partial n} \bigg|_{n=1} \text{Tr} \rho_{\mathbb{V}}^n. \quad (23)$$

The trace is taken over $\mathcal{H}_{\mathbb{V}}$ or, equivalently, over $\mathcal{H}_{\mathbb{V}}^*$.¹⁵ The latter formula is the basis of the replica trick, to be explored in the next section. We will call this quantity the *full entanglement entropy*.

We again stress that in a lattice gauge theory, we take $\mathcal{H}_{\mathbb{V}}$ to be the set of all states on links in \mathbb{V} — gauge-invariant or otherwise. The projectors in (20) make sure that only states obeying $G_v = \mathbb{1}$ for $v \in \text{Int}(\mathbb{V})$ contribute to the trace. More interestingly, $\rho_{\mathbb{V}}$ does not contain projectors that enforce $G_v = \mathbb{1}$ for $v \in \partial\mathbb{V}$. This is an important point whose ramifications we study below eq. (43) and in section 8.

If $\rho_{\mathbb{V}}$ is block diagonal in some basis, we can first compute $\text{Tr} \left(\bigoplus_k \rho_{\mathbb{V}}^{(k)} \right)^n = \sum_k \text{Tr} \left(\rho_{\mathbb{V}}^{(k)} \right)^n$ and then write

$$S_{\mathbb{V}} = -\frac{\partial}{\partial n} \bigg|_{n=1} \left[\sum_k \text{Tr} \left(\rho_{\mathbb{V}}^{(k)} \right)^n \right]. \quad (24)$$

If there is only one relevant superselection sector, i.e. if $\mathcal{A}_{\mathbb{V}}$ is the maximal algebra on $\mathcal{H}_{\mathbb{V}}$ or if all sectors but one have zero density matrices, this formula reduces to (23).

In the presence of superselection sectors, the entropy $S_{\mathbb{V}}$ is not the only interesting quantity to define. Let us first assume that we are working with matter or Abelian gauge theories. It is useful to define unit-trace density matrices $\bar{\rho}$ in each sector:

$$\rho_{\mathbb{V}}^{(k)} = p_k \bar{\rho}_{\mathbb{V}}^{(k)}, \quad p_k \equiv \text{Tr} \rho_{\mathbb{V}}^{(k)}. \quad (25)$$

The entropy associated to $\mathcal{A}_{\mathbb{V}}$ is now

$$S_{\mathbb{V}} = -\sum_k p_k \log p_k + \sum_k p_k S_{\mathbb{V}}^{(k)}, \quad (26)$$

with

$$S_{\mathbb{V}}^{(k)} = -\frac{\partial}{\partial n} \bigg|_{n=1} \text{Tr} \left(\bar{\rho}_{\mathbb{V}}^{(k)} \right)^n. \quad (27)$$

The quantity

$$S_{\mathbb{V}}^{\text{dist}} \equiv \sum_k p_k S_{\mathbb{V}}^{(k)} \quad (28)$$

is called the *distillable entanglement entropy*.¹⁶ For a generic state, it is different from the full entropy $S_{\mathbb{V}}$ whenever $\mathcal{A}_{\mathbb{V}}$ has central generators.

¹⁵ It is possible to take other definitions of the trace. For instance, it may be over the smallest faithful representation of $\mathcal{A}_{\mathbb{V}}$, or it may be over all of \mathcal{H} . Different choices will lead to entropies that differ by state-independent constants. It is important to only compare entropies defined using the same trace.

¹⁶ The edge terms are not distillable using gauge-invariant operations in the interiors of the entangling region [17], i.e. they are not convertible to entangled Bell pairs in an external reservoir of qubits via the local operations. The obstruction is that gauge-invariant operators in the interiors of the regions cannot change the superselection sector. See [18] for an explicit distillation protocol in a constrained spin system that models a lattice gauge theory.

In case we are working with nonabelian gauge theories, recall from the discussion around eq. (21) that $\rho_{\mathbb{V}}$ has more structure beyond splitting into sectors labeled by the representation R of the gauge group: within each sector, it is a tensor of a $\dim R \times \dim R$ identity matrix and a nontrivial matrix that we may call the proper reduced density matrix. Each representation R thus contributes an entropy of $\log \dim R$ to $S_{\mathbb{V}}^{(k)}$. (Here we are using k to denote all labels for superselection sectors; in general, R will be a subset of k , meaning that there may be multiple sectors k with the same label R .) In fact, it is natural to define $S_{\mathbb{V}}^{(k)}$ as the entropy of the proper reduced density matrix in sector k , as the $\log \dim R$ contribution is also not distillable. Hence, in a nonabelian gauge theory, we will decompose the full entanglement entropy as

$$S_{\mathbb{V}} = - \sum_k p_k \log p_k + \sum_k p_k S_{\mathbb{V}}^{(k)} + \sum_R p_R \log \dim R, \quad (29)$$

where p_R is the total weight of all sectors k that contain R .

With this convention,

$$S_{\mathbb{V}}^{\text{g-inv}} \equiv - \sum_k p_k \log p_k + \sum_k p_k S_{\mathbb{V}}^{(k)} \quad (30)$$

will be called the *gauge-invariant entropy* of \mathbb{V} . This is the entropy that we would get if the trace in (23) was taken in the representation associated to gauge-invariant degrees of freedom only, instead of using the representation on the full space $\mathcal{H}_{\mathbb{V}} = \bigotimes_{\ell} \mathcal{H}_{\ell}$. We caution the reader that this name is slightly misleading, as the full entropy (29) is gauge-invariant, too.

6. From subalgebras to path integrals

We will now show how the various entropies defined in the previous section can be computed using path integrals.¹⁷ Let $\{\phi^{(k)}\}$ be a basis of $\mathcal{H}_{\mathbb{V}}^*$ that diagonalizes all central elements of $\mathcal{A}_{\mathbb{V}}$, and let $\{|\vartheta\rangle\}$ be any basis of $\mathcal{H}_{\bar{\mathbb{V}}}$. The eigenvalues of central generators are collected in the tuple k . For instance, in a scalar theory with Dirichlet boundary conditions on the entangling edge, $\phi_i^{(k)}$ could be used to label the position eigenstate with field configuration ϕ_i for $i \in \text{Int}(\mathbb{V})$ and with boundary conditions $\phi_v = k_v$ for $v \in \partial \mathbb{V}$. In a gauge theory, instead of including in k the Gauss operator eigenvalues in $\text{Int}(\mathbb{V})$ (which are all equal to unity, cf. (20)), we will simply require that any $|\phi^{(k)}\rangle$ be invariant under all gauge transformations in $\text{Int}(\mathbb{V})$. This way in gauge theories the tuple k will contain only eigenvalues of central operators on the edge.

The wave function of the ground state $|\Psi\rangle$ of a theory with Euclidean Lagrangian L is given by

$$\begin{aligned} \Psi_k(\phi, \vartheta) &\equiv \langle \phi^{(k)}, \vartheta | \Psi \rangle = \int [d\varphi] e^{-\int_{-\infty}^0 L d\tau}, \\ \varphi_i(\tau = 0^-) &= \phi_i^{(k)} \text{ for } i \in \mathbb{V}, \quad \varphi_j(\tau = 0^-) = \vartheta_j \text{ for } j \in \bar{\mathbb{V}}. \end{aligned} \quad (31)$$

¹⁷ When working with theories with finite target spaces, like spin systems or fermions, we may either consider the time τ to be discrete, or we may take it to be continuous with the understanding that all integrals are constructed using the coherent state approach [33]. Either way, τ is always Euclidean. For gauge theories, it is most helpful to keep τ discrete, with fields A_0 living on temporal links. Integrating out the A_0 's in any Maxwell/Yang-Mills theory enforces local conservation of charge, making the path integral include only gauge-invariant configurations [34].

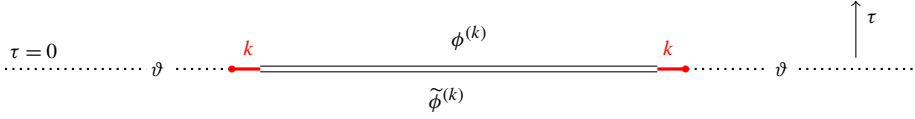


Fig. 4. A side view of the path integral boundary conditions on the $\tau = 0$ slice, illustrating eq. (33). At all other values of τ , the integral is unconstrained; even at $\tau = 0$, there is a sum over all values of fields ϑ in \mathbb{V} , i.e. outside the two red circles. The setup shown here calculates $\langle \tilde{\phi}^{(k)} | \rho_{\mathbb{V}}^{(k)} | \phi^{(k)} \rangle$ when central generators are at or near $\partial \mathbb{V}$; their locations are shown in red. The values of central generators are denoted k and are the same at $\tau = 0^+$ and $\tau = 0^-$. The replica trick calculates $\text{Tr}[\rho_{\mathbb{V}}^{(k)}]^n$ by taking n copies of this setup and constraining $\phi^{(k)}$ on the replica l to equal $\tilde{\phi}^{(k)}$ on replica $l + 1$, with $n + 1 \equiv 1$. The fields in the red region are then eigenstates of central generators with eigenvalue k on each replica.

The notation $\int^{(k)}$ means that the path integral variables within the region \mathbb{V} on the $\tau = 0$ slice are constrained to equal the eigenvalues k of central generators of $\mathcal{A}_{\mathbb{V}}$.¹⁸ The path integral is normalized so that

$$\sum_{k, \phi, \vartheta} \Psi_k^*(\phi, \vartheta) \Psi_k(\phi, \vartheta) = \sum_k \int^{(k)} [d\varphi] e^{-\int_{-\infty}^{\infty} L d\tau} = 1. \quad (32)$$

The reduced density matrix induced by the algebra $\mathcal{A}_{\mathbb{V}}$ is block-diagonal in the basis $\{|\phi^{(k)}\rangle\}$, and each block has matrix elements

$$\begin{aligned} \langle \tilde{\phi}^{(k)} | \rho_{\mathbb{V}}^{(k)} | \phi^{(k)} \rangle &= \sum_{\vartheta} \Psi_k^*(\phi, \vartheta) \Psi_k(\tilde{\phi}, \vartheta) = \int^{(k)} [d\varphi] e^{-\int_{-\infty}^{\infty} L d\tau}, \\ \varphi_i(\tau = 0^+) &= \phi_i^{(k)}, \quad \varphi_i(\tau = 0^-) = \tilde{\phi}_i^{(k)} \text{ for } i \in \mathbb{V}. \end{aligned} \quad (33)$$

On the other hand, the normalized density matrix in sector k (cf. eq. (25)) has matrix elements

$$\langle \tilde{\phi}^{(k)} | \bar{\rho}_{\mathbb{V}}^{(k)} | \phi^{(k)} \rangle = \int [d\varphi]_k e^{-\int_{-\infty}^{\infty} L d\tau}, \quad (34)$$

with the measure chosen so that $\text{Tr} \bar{\rho}_{\mathbb{V}}^{(k)} = 1$ for all k . See Fig. 4. For instance, for a spin chain with Dirichlet boundary conditions at the edge sites $\{v_1, v_2\}$ of the entangling interval \mathbb{V} , the normalization condition

$$\int^{(k)} [d\varphi]_k e^{-\int_{-\infty}^{\infty} L d\tau} = 1 \quad (35)$$

has $k = (k_{v_1}, k_{v_2}) \in \{\uparrow, \downarrow\} \times \{\uparrow, \downarrow\}$ labeling the four superselection sectors.

The sector weights p_k appearing in (26) can be calculated from the original path integral (33), as made clear by their definition (25):

$$p_k = \text{Tr} \rho_{\mathbb{V}}^{(k)} = \int^{(k)} [d\varphi] e^{-\int_{-\infty}^{\infty} L d\tau}. \quad (36)$$

¹⁸ More precisely, we should also allow for constraints on links emanating from $\partial \mathbb{V}$. Such constraints would play a role when computing, say, the entropy of the algebra $\mathcal{G}_{\mathbb{D}, \mathbb{D}}$ in eq. (9). They will also play a role when studying various entropies of gauge theories. We will henceforth include this possibility when referring to boundary conditions at $\tau = 0$.

Note that here the measure of the path integral is the same as in (31), meaning that when summed over all k the integrals in the above equation yield unity. The two measures are related by

$$p_k [d\varphi]_k = [d\varphi]. \quad (37)$$

These probabilities can also be obtained from the expectation values of central operators. Let $\{\mathcal{O}_i\}_{i=1,\dots,M}$ be the generators of the center of $\mathcal{A}_{\mathbb{V}}$, chosen so that their eigenvalues k_i are roots of unity. A general operator in the basis of the central algebra has the form $\mathcal{O}^{(m)} \equiv \prod_{i=1}^M \mathcal{O}_i^{m_i}$. Its expectation is

$$\langle \mathcal{O}^{(m)} \rangle = \sum_k p_k k_1^{m_1} \dots k_M^{m_M} \equiv \sum_k \mathbf{K}_{mk} p_k. \quad (38)$$

Note that the tuple m is conjugate to k , and the matrix \mathbf{K} is simply the kernel of an M -dimensional Fourier transform.¹⁹ This ensures that \mathbf{K} is always invertible, and we can write

$$p_k = \sum_m \mathbf{K}_{km}^{-1} \langle \mathcal{O}^{(m)} \rangle. \quad (39)$$

This expression can also be written as $p_k = \langle \prod_i P_{k_i} \rangle$, where $P_{k_i} \propto \sum_{m_i} k_i^{m_i} \mathcal{O}_i^{m_i}$ is the projector onto the k_i -eigenstate of the central generator \mathcal{O}_i .

Finally, let

$$p_k \equiv e^{-I(k)}. \quad (40)$$

The sum over individual contributions of sectors to the entropy $S_{\mathbb{V}}$ can be written as

$$S_{\mathbb{V}}^{\text{dist}} \equiv \sum_k p_k S_{\mathbb{V}}^{(k)} = - \frac{\partial}{\partial n} \bigg|_{n=1} \left[\sum_k e^{-I(k)} \int [d\varphi]_k^n e^{-\int_{(\mathbb{M} \times \mathbb{R})^n} L d\tau} \right], \quad (41)$$

where $(\mathbb{M} \times \mathbb{R})^n$ denotes the n -replicated Euclidean spacetime, and $[d\varphi]_k^n = [d\varphi_1]_k \dots [d\varphi_n]_k$ is the replicated measure for the path integral in sector k , normalized according to eq. (35). The same boundary condition k holds on each copy of the region \mathbb{V} on the $\tau = 0$ slice. The sum over boundary conditions (edge modes) does not depend on n and can be commuted past the derivative $\frac{\partial}{\partial n}$.

The full entropy associated to the algebra $\mathcal{A}_{\mathbb{V}}$, given by eq. (26), can be calculated by modifying the edge mode contribution in (44) to get

$$S_{\mathbb{V}} = - \frac{\partial}{\partial n} \bigg|_{n=1} \left[\sum_k e^{-nI(k)} \int [d\varphi]_k^n e^{-\int_{(\mathbb{M} \times \mathbb{R})^n} L d\tau} \right]. \quad (42)$$

According to eq. (37), $e^{-nI(k)} [d\varphi]_k^n = [d\varphi]^n$, so the full entropy can also be written as the path integral with the original measure on each replica, in agreement with eq. (24):

¹⁹ For example, in a \mathbb{Z}_2 system, the eigenvalues k_i of each central generator are ± 1 . They can be written as $e^{i\pi \kappa_i}$ for $\kappa_i \in \{0, 1\}$, and the expectations become $\langle \mathcal{O}^{(m)} \rangle = \sum_{\kappa} e^{i\pi \sum_i \kappa_i m_i} p_{\kappa}$. Note that $m_i = 0$ corresponds to the identity operator ($\mathcal{O}^{(0)} = \mathbb{1}$), which is always in the center. In subsection 7.2 we will show how this works for the case of a free compact scalar.

$$S_{\mathbb{V}} = -\frac{\partial}{\partial n} \bigg|_{n=1} \left[\sum_k \int [\mathrm{d}\varphi]^n e^{-\int_{(\mathbb{M} \times \mathbb{R})^n} L \, \mathrm{d}\tau} \right]. \quad (43)$$

We also record the expression (41) for $S_{\mathbb{V}}^{\mathrm{dist}}$ once eq. (37) is employed:

$$S_{\mathbb{V}}^{\mathrm{dist}} = -\frac{\partial}{\partial n} \bigg|_{n=1} \left[\sum_k \int [\mathrm{d}\varphi]^n e^{(n-1)I(k) - \int_{(\mathbb{M} \times \mathbb{R})^n} L \, \mathrm{d}\tau} \right]. \quad (44)$$

Let us clarify what these path integrals look like in an Abelian gauge theory. Conventionally, path integrals are defined over variables A_ℓ on both spatial and temporal links. On spatial links, e^{iA_ℓ} are eigenvalues of Z_ℓ , whereas — as mentioned in footnote 17 — A_ℓ on temporal links are Lagrange multipliers enforcing the gauge constraint. After imposing these constraints, we may gauge-fix the variables on temporal links to $A_\ell = 0$. We are left with time-independent gauge transformations; let us set them aside for a moment. In the axial gauge fixed so far, electric field eigenvalues are time derivatives $\partial_\tau A_\ell$ of variables on spatial links. Eigenvalues of the electric central generators E_v (defined in eq. (19)) are $\prod_{\ell \supset v, \ell \in \mathbb{V}} \exp \{i \partial_\tau A_\ell\}$ for each $v \in \partial \mathbb{V}$. In the path integral (43) associated to the maximal gauge-invariant algebra, within each $\int^{(k)}$ the fields are thus constrained to have

$$e^{i \sum_{\ell \supset v, \ell \in \mathbb{V}} \partial_\tau A_\ell(\tau=0)} = k_v, \quad v \in \partial \mathbb{V}. \quad (45)$$

(The integral over temporal A_ℓ 's emanating from $v \in \partial \mathbb{V}$ makes sure that field configurations on $\bar{\mathbb{V}}$ end up having a matching flux k at $\partial \bar{\mathbb{V}}$.) The conditions (45) are defined in axial gauge but are invariant under time-independent gauge transformations. Thus we can fix the remaining gauge freedom in any way we want, but we caution any prospective gauge-fixers that the replica structure makes it extremely natural to fix separately inside \mathbb{V} and outside of it.

In nonabelian theories, the same story as above plays out, with one significant difference. The path integral (43) for the maximal $\mathcal{A}_{\mathbb{V}}$ forces the gauge fields to have a definite representation of the electric flux through $\partial \mathbb{V}$ for each sector k . Unlike in the Abelian case, an eigenstate with a definite flux representation is not invariant under time-independedent gauge transformations on $v \in \partial \mathbb{V}$. (Recall from section 4 that any state in $\mathcal{H}_{\mathbb{V}}^R$ has one “dangling” index α at each edge site, resulting in $\rho_{\mathbb{V}}$ having factors of $\dim R \times \dim R$ identity matrices.) We can *choose* whether to fix the gauge on the edge; if we do not, these dangling degrees of freedom will get replicated and will contribute to $S_{\mathbb{V}}$. Gauge-fixing on $\partial \mathbb{V}$ thus reduces $S_{\mathbb{V}}$ down to the “gauge-invariant” part $S_{\mathbb{V}}^{\mathrm{g-inv}}$ given by (30). We will discuss the physical nature of this gauge-fixing in section 8.

Equation (43) should be taken as the regularized *definition* of the replica trick path integral. It gives a precise treatment of boundary conditions and distinguishes between entropies associated to algebras with different centers. In general, the entropy associated to a general subalgebra $\mathcal{A}_{\mathbb{V}}$ can be calculated algorithmically by following these steps:

1. Determine the center of $\mathcal{A}_{\mathbb{V}}$.
2. Determine the set of degrees of freedom on which $\mathcal{A}_{\mathbb{V}}$ acts nontrivially, and pick the basis of the associated space $\mathcal{H}_{\mathbb{V}}^*$ so that the center generators are all diagonal. This sets the variables to be used in the path integral.
3. Calculate the full entropy $S_{\mathbb{V}}$ via eq. (43). In a nonabelian theory, the same equation can be used to calculate the gauge-invariant entropy $S_{\mathbb{V}}^{\mathrm{g-inv}}$. It is obtained by fixing the time-dependent gauge transformations on $\partial \mathbb{V}$.

This procedure can be applied to any quantum theory, assuming we know how to discretize it such that we start off from a Hilbert space that factorizes.

7. Miscellaneous comments

7.1. Gauge vs. matter theories

As repeatedly stressed throughout this note, there is no procedural difference between calculating entropies in matter and in gauge theories. The principal difference between these theories is that the latter always have central generators — the Gauss operators. Moreover, in a gauge theory, the maximal gauge-invariant subalgebra supported on \mathbb{V} always has additional central generators in the form of edge electric flux operators E_v , defined above eq. (20). In contrast, the maximal subalgebra on a region \mathbb{V} has no center for matter theories.

In entropy calculations, Gauss operators in the interior of a region \mathbb{V} do not give rise to any sums over superselection sectors. This is because all gauge-invariant states are in the same, singlet sector in which $G_v = \mathbb{1}$ holds as an operator equation. No other sectors are populated. This means that any reduced density matrix in \mathbb{V} can be represented as a density matrix of gauge-invariant degrees of freedom times a projection to the singlet sector, eq. (20).

On the other hand, the existence of flux operators E_v on $v \in \partial\mathbb{V}$ means that gauge theories will naturally have nontrivial superselection sectors, and since these are supported on edges it is reasonable to refer to them as edge modes. (Here we assume that the “natural” algebra associated to a region is the maximal allowed algebra of operators on that region.) Using the “balanced center” procedure of [14] it is possible to find a non-maximal gauge-invariant algebra that has no edge modes. However, in a topological phase of a gauge theory this kind of algebra does not give rise to the usual topological entanglement entropy. In general, it is not known how universal data about the gauge theory are encoded in entropies of algebras with non-maximal centers.

7.2. The edge mode action

The general prescription in the previous section can be straightforwardly applied to $d = 1$ systems. However, $d > 1$ might seem hopeless, as the size of the center grows exponentially with $|\partial\mathbb{V}|$. Fortunately, there are several situations in higher dimensions when the superselection sectors can be tamed. These are either free or infinitely gapped theories of various kinds, and here we will enumerate some examples. To simplify the discussion, we will focus on Dirichlet algebras for matter theories and on electric algebras for gauge theories. In both cases a sum over sectors generically appears.

1. *Product states:* The sum over superselection sectors is trivial if only one sector has nonzero weight. In such a case there is no “classical” entropy from the $-\sum_k p_k \log p_k$ term in (26). In a matter theory, this happens in massive phases with product states, such as in the Ising model at large external field. The ground state has all spins pointing in the same direction, say $Z_v = \mathbb{1}$. If the center of $\mathcal{A}_{\mathbb{V}}$ is generated by Z ’s at the edge, only one superselection sector will be populated. Similarly, a lattice gauge theory at strong coupling will have a ground state with $X_\ell = \mathbb{1}$ on all links, and only one configuration of edge modes (electric fields piercing $\partial\mathbb{V}$) will appear in the reduced density matrix.

2. *Topological phases*: All gauge theories with a finite gauge group have a topological phase at weak coupling.²⁰ In this limit, central generators are the electric flux operators E_v , and topological states have the same p_k for each set k of eigenvalues of E_v , modulo global constraints (as discussed in section 3). A quick way to see this is to note that all electric flux operators have zero expectation values, so $p_k = \frac{1}{K^{M-1}}$ follows from eq. (39) (with M denoting the number of sites at $\partial\mathbb{V}$, and \mathbb{Z}_K being the gauge group).
3. *Symmetry protection*: The previous two examples dealt with extreme weak and strong coupling limits of simple theories. It is conceivable that with a sufficient amount of symmetry, only a few superselection sectors will be populated in more complicated models, as well. At the very least, we expect this to be the case in any symmetry-protected trivial (SPT) or symmetry-enhanced topological (SET) phases. We will see in section 8 that Chern-Simons theory naturally fits into this category, too. It would be interesting to discover other nontrivial theories in which the superselection sectors are simple due to a symmetry.
4. *Free theories*: In a free field theory the expectation values that enter eq. (39) are nontrivial, but they do factorize due to Wick's theorem. Consider a compact scalar, with central generators $Z_v \equiv e^{i\phi_v}$ for $v \in \partial\mathbb{V}$. The expectations of central elements can be written as $\langle e^{i\sum_v m_v \phi_v} \rangle \sim e^{-\sum_{v,u} G_{vu} m_v m_u}$. Plugging this back into (39), using $k_v \equiv e^{ik_v}$ and $\mathbf{K}_{km}^{-1} \sim e^{-i\sum_v \kappa_v m_v}$, and summing over m_v , we find sector probabilities $p_k \sim e^{-\sum_{v,u} G_{vu}^{-1} \kappa_v \kappa_u}$. This is an elementary way to derive the Gaussian distribution of edge modes that has played an important role in understanding the entanglement entropy in $d = 3$ Maxwell theory [21,35,36]. Note that the edge mode action $I(k) = -\log p_k$ is nonlocal. The action that appears in the replica trick, eq. (44), is thus not equivalent to the naive restriction of the original action to the subregion \mathbb{V} , as would be the case for an algebra $\mathcal{A}_{\mathbb{V}}$ without a center.

7.3. Lattice, continuum, and conformal boundary conditions

The discussion so far was centered on path integrals carried out over lattice fields. When such integrals are dominated by configurations that vary over length scales much greater than the lattice spacing, lattice boundary conditions can be replaced by continuum ones. If the field configurations of interest vary slowly and the theory remains weakly coupled, the type of boundary condition (open, Dirichlet, Neumann, electric, magnetic, etc) will be the same in the continuum and on the lattice, although the edge mode action $I(k)$ may be renormalized. For instance, in a free theory we can approximate the sum over edge modes k with an integral over boundary configurations $\varphi(x_{\parallel})$ for $x_{\parallel} \in \partial\mathbb{V}$, and these continuum edge modes will still have Gaussian weights, with an effective action $I[\varphi]$ [13]. This naive approach will not be possible for theories without an explicit Lagrangian description, for instance for strongly coupled theories defined as the infrared limits of known theories, where boundary conditions are more abstractly determined by bulk-to-boundary operator product expansions (see [37] for a modern take on this). It is not known how to generalize our discussion to these cases, though of course if we were powerful enough to compute path integrals directly in the UV, our analysis would apply.

In arbitrary $d = 1$ CFTs, information about boundary conditions can be captured in a different, more efficient way. Conformal boundary conditions are classified by their transformations under

²⁰ If the gauge group is continuous, the topological phase may be inaccessible. Consider the case of $U(1)$. We can imagine regulating it with \mathbb{Z}_K for $K \gg 1$; the topological phase is reached when the coupling is much smaller than $1/K$ [32], which is not a limit one usually considers in $U(1)$ theories.

the conformal group (with the space of boundary conditions spanned by Ishibashi states). Linear combinations of Ishibashi states, chosen in order to have nice modular properties, are called Cardy states [38]. There is one Cardy state for each primary operator. In a study of the Ising CFT [25], Cardy states at the entangling edges in the path integral construction were connected with choices of algebras.²¹ In this reference it was also found that the leading universal term does not depend on these choices, which appears to be a non-generic phenomenon: as we will review in more detail below, the universal term depends on the choice of entropy in $d = 3$ free theories.

Specifically, for an interval \mathbb{V} , consider the maximal subalgebra $\mathcal{A}_{\mathbb{V}}$ of the Ising model $H = \sum_i (X_i X_{i+1} + Z_i)$ and the corresponding Dirichlet-type algebras $\mathcal{A}_{\mathbb{V}}^{\text{D,D}}$, $\mathcal{A}_{\mathbb{V}}^{\text{D,O}}$, and $\mathcal{A}_{\mathbb{V}}^{\text{O,D}}$ defined by dropping Z generators from one or both of the edge sites. The algebra $\mathcal{A}_{\mathbb{V}}$ corresponds to open boundary conditions on the lattice. In terms of conformal boundary conditions, it corresponds to the Cardy states $|\sigma\rangle$ on both entangling cuts. An open boundary condition on one edge site corresponds to the $|\sigma\rangle$ state on that site in the CFT.

The Cardy states $|\mathbb{1}\rangle$ and $|\varepsilon\rangle$ are more interesting, and they arise when Dirichlet conditions are imposed. The edge modes here are labeled by the pair $k = (k_1, k_2)$, $k_i \in \{\pm 1\}$, of eigenstates of edge operators X_v . When $k_1 = k_2$, the Cardy state is $|\mathbb{1}, \mathbb{1}\rangle$ or $|\varepsilon, \varepsilon\rangle$. (The two choices give the same path integrals, and they calculate the individual sector entropies $S_{\mathbb{V}}^{(k)}$.) When $k_1 = -k_2$, the Cardy state is $|\mathbb{1}, \varepsilon\rangle$ (or $|\varepsilon, \mathbb{1}\rangle$). When one boundary condition is open and the other Dirichlet, there are two superselection sectors labeled by, say, k_1 , and the Cardy states on the edges are $|\mathbb{1}, \sigma\rangle$ or $|\varepsilon, \sigma\rangle$, depending on the choice of k_1 .

This is the only known example of a connection between conformal boundary conditions and algebras. It would be interesting to understand this link in more instances, and to check whether the universal leading-order behavior in other $d = 1$ CFTs does not depend on the choice of algebra, as was found to be the case in the Ising model.

7.4. Universal terms

In a CFT with a UV cutoff ϵ and in $d > 1$ space dimensions, it is common to say that the ground state entanglement entropy for a spherical region \mathbb{V} of size R takes the form

$$S_{\mathbb{V}} = a_{d-1} \frac{R^{d-1}}{\epsilon^{d-1}} + a_{d-3} \frac{R^{d-3}}{\epsilon^{d-3}} + \dots + C_d \begin{cases} -\gamma, & d \text{ even} \\ a \log \frac{R}{\epsilon}, & d \text{ odd} \end{cases} + O(\epsilon/R). \quad (46)$$

Here C_d is a theory-independent constant,²² and γ and a are universal (regularization-independent) and can often be interpreted as counting the number of degrees of freedom in the CFT [9,39]. Which of the entropies defined so far should be understood when writing such formulas?

This question is not trivial; different entropy choices lead to different values of γ and a . Each may have a counting interpretation and lead to c -theorems of interest [40,41]. To answer this question, we note that for odd d , the quantity a can be connected to the trace anomaly of the

²¹ Ref. [25] focused on the Rényi entropies $(1-n)^{-1} \log \text{Tr} \rho_{\mathbb{V}}^n$ in the limit $n \rightarrow \infty$. We here assume that the established connection between subalgebras and this particular limit of Rényi entropies can also extend all the way to $n = 1$, where the usual entanglement entropy is recovered.

²² Different authors have different conventions for this constant, and in section 8 we will simply absorb it into the definition of a and γ .

CFT [42].²³ Briefly, the trace anomaly is calculated by the one-point function of the trace of the stress-energy tensor, $T_\mu{}^\mu$, on a curved manifold, with the stress-energy tensor defined as the variation of the action under a Weyl rescaling of the metric.

Now, if an entropy is computed by a replica trick path integral whose only dependence on the metric enters through the Lagrangian L and the normalization of $[d\varphi]^n$, then integrating over a replicated spacetime will yield a universal contribution proportional to the integral of $\langle T_\mu{}^\mu \rangle$ over the entanglement edge $\partial\mathbb{V}$. The full entropy $S_{\mathbb{V}}$ from eq. (43) is computed this way, unlike $S_{\mathbb{V}}^{\text{dist}}$ or any of the $S_{\mathbb{V}}^{(k)}$'s, which all have additional edge mode actions with nontrivial metric dependence in the path integral. (Of course, in particular QFT states the edge mode action may be metric-independent, as discussed in section 7.2, but we have no reason to believe this happens in the ground state of any CFT.) On the other hand, the gauge-invariant entropy $S_{\mathbb{V}}^{\text{g-inv}}$ differs from $S_{\mathbb{V}}$ only by gauge-fixing at every point along the entanglement edge. Thus, at this stage, we claim that only $S_{\mathbb{V}}$ and $S_{\mathbb{V}}^{\text{g-inv}}$ can be entropies whose universal terms can generically match the trace anomaly.

We have also not specified *which* algebra we want to compute the full or gauge-invariant entropy of. Based on calculations in free theories [21,35], we conjecture that in all theories the trace anomaly matches the universal term in the entropy when $\mathcal{A}_{\mathbb{V}}$ is the *maximal* algebra on region \mathbb{V} . Moreover, based on calculations in topological and confined phases of lattice theories [14,32], we also conjecture that the quantity γ (sometimes also denoted F) that is customarily defined in even d also matches the universal piece of the entropy of a maximal algebra on \mathbb{V} .

In the context of gauge theories, we will provide more plausibility arguments for this conjecture in section 8. In particular, we will note that the topological entanglement entropy γ of Chern-Simons theory is reproduced only if no gauge-fixing is assumed at $\partial\mathbb{V}$. Thus, we can sharpen our conjecture and claim that it is always the full entropy $S_{\mathbb{V}}$ of the maximal $\mathcal{A}_{\mathbb{V}}$ that contains the conventionally defined universal terms. This claim was also made in section 5 of ref. [18].

7.5. Holographic entanglement entropy

If a theory is holographic, the replica trick path integral can be calculated using a replicated bulk geometry, and the entropy associated to this bulk partition function is approximated by the area of the Ryu-Takayanagi surface in the classical limit [5,43]. An important aspect of this story is that the Ryu-Takayanagi prescription does not involve a choice of subalgebra or of a type of boundary condition. Which of the many entropies that can be associated to a region \mathbb{V} does holography compute?

This question reduces to the one discussed in the previous subsection. The holographic entropy is known to reproduce the correct trace anomaly (or the γ quantity that features in contact terms) in its universal terms when the bulk contains Einstein gravity [39,44]. Moreover, in higher derivative theories of gravity, the Ryu-Takayanagi prescription is modified so that the holographic entanglement entropy still computes the correct trace anomalies [45]. Building on the conjectures from the previous subsection, we propose that, in all dimensions, it is the full entropy of the maximal algebra on \mathbb{V} that is computed by the holographic entanglement entropy, defined via

²³ For even d , the trace anomaly vanishes and we do not have an analogous argument. However, the quantity γ can be connected to contact terms in correlation functions that play roles similar to perturbative anomalies in odd d .

Ryu-Takayanagi or its generalizations. It is not known what gravity constructions would calculate any of the other entropies that can be associated to \mathbb{V} .

We emphasize that the arguments used to buttress our conjectures all relied on universal terms in entanglement entropies, in both odd and even dimensions. It may sound surprising that universal terms can be influenced by $p_R \log \dim R$ terms in the full entropy, eq. (29), because in the path integral language these terms arise from a local gauge-fixing on each point at the entanglement edge. The situation is most clearly assessed by recalling the expressions for p_k in a free theory, derived in subsection 7.2; the probabilities p_R are functions of p_k and hence they are generically nonlocal, too. In other words, the degrees of freedom at the edge that we exclude by gauge-fixing will generically exhibit nonlocal correlations that contain universal information. A very concrete example of this is the fact that in a gauge theory there must exist an overall constraint on all the fluxes piercing the entanglement edge, due to the Gauss law. This constraint must be felt by both $p_k \log p_k$ and $p_R \log \dim R$ terms. In $d = 2$, it is precisely this overall constraint that induces the topological entanglement entropy in a lattice gauge theory (albeit there it comes from the $p_k \log p_k$ term). In a nonabelian Chern-Simons theory, however, the topological entanglement entropy instead arises from the $p_R \log \dim R$ term, as we will review in the next section.

8. Local extensions of the Hilbert space in gauge theories

In this section, we discuss the second universality class of approaches to define the entanglement entropy of a lattice gauge theory. In contrast to the previous sections, here we will take the perspective that the Hilbert space of the gauge theory contains only the gauge-invariant states that faithfully represent the gauge-invariant operator algebra. We then *locally* extend this Hilbert space along the entangling cut by adding the minimal number of degrees of freedom at the cut so that it splits into tensor factors, and so that the time-independent gauge symmetry at the boundary is promoted to be a faithfully acting global symmetry.²⁴ Although this operation may seem quite arbitrary, the two main points in this section are that (i) it is a physically meaningful procedure when the gauge theory emerges as an effective low-energy description of the physics, and (ii) it coincides with an especially natural set of boundary conditions in the replica trick, that roughly amount to smoothing out the conical singularity. (In section 7.4 we arrived at an analogous requirement: that the replicated path integral depends on the metric only through the action and the usual measure.) Much of this section serves as a review of a cross-section of the literature on entanglement entropy in gauge theories in which these points were crystalized.

In cases where we can put the gauge theory on a lattice, the results obtained in this section coincide with the full entanglement entropy of the maximal algebra, (29), in the conventions of the previous sections. This is because previously, we were already using the convention that the algebraic entanglement entropy for lattice gauge theories was defined w.r.t. a (globally) extended Hilbert space, as described in footnote 6 (and the extension of the Hilbert space away from the entangling surface does not affect the partial trace calculation).²⁵

Organizationally, we will start by introducing the extended Hilbert space approach both on the lattice and in the continuum. We first do this at a formal level, then provide a physical justification

²⁴ Without this last requirement, the minimal extension of the Hilbert space would not have any degrees of freedom directly on the edge; this is tantamount to removing the $\log \dim R$ terms by gauge-fixing all the edge degrees of freedom.

²⁵ In particular, earlier references to the “full entropy of the maximal algebra” in a gauge theory include both the Shannon and $\log \dim R$ edge terms in the notation below.

for it. We then discuss the replica trick boundary conditions, and finally conclude with assorted comments.

8.1. Definition

8.1.1. On the lattice

The extended Hilbert space definition on the lattice is easiest to introduce by example. Consider Yang-Mills theory on a $d = 1$ spatial lattice with two sites connected by two links, and suppose that we want to assign a reduced density matrix to one of the links. This is the problem of computing the entanglement across an interval in $d = 1$ Yang-Mills theory on a spatial circle [16].

Yang-Mills in two spacetime dimensions is solvable, see e.g. [46]. It's known for instance that the Hilbert space on the circle is the space of square-integrable class functions on the gauge group G (L^2 -functions on the group manifold s.t. $\psi(gug^{-1}) = \psi(u)$ for all $g \in G$), for which a convenient basis is furnished by the characters,

$$\mathcal{H}_{\text{phys}} = \text{class functions in } L^2(G) = \text{span}\{|R\rangle\}. \quad (47)$$

The gauge-invariant operators in the theory are the Casimirs built out of the electric fields, which are diagonal in the representation basis, and the Wilson loop operators in all the representations of G . The Hilbert space (47) doesn't refer to the underlying spatial manifold and so clearly doesn't factorize. In the extended Hilbert space prescription, we're instructed to embed it into the minimal larger one that does.

Intuitively, the presence of extended operators that cannot be split into local constituents (i.e. the Wilson loops) is what prevents the Hilbert space that faithfully represents the operator algebra from factorizing. We can get around this by adding pairs of non-dynamical (infinitely massive) surface charges in all the representations of the gauge group at the ends of our entangling interval, which allow us to cut the Wilson loops into Wilson lines. This is tantamount to lifting the Gauss constraint at the lattice sites, so the extended Hilbert space is the tensor product of two link Hilbert spaces, which are spanned by the matrix elements of the group representations:

$$\mathcal{H}_{\text{ext.}} = \mathcal{H}_{\text{left}} \otimes \mathcal{H}_{\text{right}}, \quad \mathcal{H}_{\text{left/right}} = L^2(G) = \text{span}\{|R_{\alpha\beta}\rangle_{\text{left/right}}\}. \quad (48)$$

(See section 3.4 for more details.) Under a gauge transformation $\Lambda \in G$ on the first site, these states transform as

$$|R_{\alpha\beta}\rangle_{\text{left}} \mapsto \sum_{\alpha'} R_{\alpha\alpha'}^{-1}(\Lambda) |R_{\alpha'\beta}\rangle_{\text{left}}, \quad |R_{\alpha\beta}\rangle_{\text{right}} \mapsto \sum_{\alpha'} R_{\alpha\alpha'}(\Lambda) |R_{\alpha'\beta}\rangle_{\text{right}}, \quad (49)$$

and under a gauge transformation on the second site they transform as

$$|R_{\alpha\beta}\rangle_{\text{left}} \mapsto \sum_{\beta'} R_{\beta\beta'}(\Lambda) |R_{\alpha\beta'}\rangle_{\text{left}}, \quad |R_{\alpha\beta}\rangle_{\text{right}} \mapsto \sum_{\beta'} R_{\beta\beta'}^{-1}(\Lambda) |R_{\alpha\beta'}\rangle_{\text{right}}. \quad (50)$$

The physical Hilbert space $\mathcal{H}_{\text{phys}}$ is embedded into $\mathcal{H}_{\text{ext.}}$ as

$$|R\rangle \rightarrow \frac{1}{\dim R} \sum_{\alpha, \beta \in 1, \dots, \dim R} |R_{\alpha\beta}\rangle_{\text{left}} \otimes |R_{\beta\alpha}\rangle_{\text{right}}. \quad (51)$$

It is easy to see that the state above is indeed gauge-invariant. The physical picture is that, with the matrix indices labeling the infinitely massive surface charges, we want pairs of charges across

the entangling cut to transform oppositely under gauge transformations in order to make a gauge-invariant state, so we assign them the same index.

We can now take the most general state $|\Psi\rangle = \sum_R \psi_R |R\rangle \in \mathcal{H}_{\text{phys}}$, embed it in the extended Hilbert space as $\sum_{R\alpha\beta} \frac{\psi_R}{\dim R} |R\alpha\beta\rangle_{\text{left}} \otimes |R\beta\alpha\rangle_{\text{right}}$, and compute the reduced density matrix by tracing out, say, the right tensor factor in the extended Hilbert space. Its von Neumann entropy is

$$S_{\text{left}} = - \sum_R p_R \log p_R + 2 \sum_R p_R \log \dim R \quad (52)$$

where $p_R = |\psi_R|^2$. The first term is a probability distribution over the states; we'll call this the Shannon entropy. The second, "log dim R " term, comes from the statistics of fusing the non-dynamical surface charges to make a singlet. (The factor of 2 comes from the entangling interval having two ends.) This concludes our review of the $d = 2$ example.

In higher-dimensional lattice gauge theories, the extended Hilbert space idea was discussed in refs. [12,13,17].²⁶ There we find the same types of edge terms as in (52) when we promote the R 's that label the edge states $\{|R\rangle\}$ to instead label superselection sectors specified by the eigenvalues of the Casimirs at all the boundary lattice sites (as mentioned in section 4, where these were included in the general sector labels k). Then the entropy associated to a region \mathbb{V} is

$$S_{\mathbb{V}} = - \sum_R p_R \log p_R + \sum_R p_R \log \dim R - \sum_R p_R \text{Tr}(\rho_R \log \rho_R), \quad (53)$$

ρ_R being the normalized density matrix in the superselection sector R (see eq. (27)). The first two terms are the edge terms from each lattice site along the boundary, and the third captures the distillable [17] entanglement of interior degrees of freedom, that the $d = 2$ example was too simple to support.

In short, we see that the extended Hilbert space prescription on the lattice agrees with the full entanglement entropy, (29), w.r.t. the globally extended conventions of the previous sections. This was first pointed out in [17].

It's interesting to compare this result to the entropy $S_{\mathbb{V}}^{\text{g-inv}}$ of the maximal gauge-invariant subalgebra/electric algebra *with respect to the gauge-invariant Hilbert space*. As discussed around (29), (30),

$$S_{\mathbb{V}} = S_{\mathbb{V}}^{\text{g-inv}} + \sum_R p_R \log \dim R; \quad (54)$$

the entropy of the maximal gauge-invariant subalgebra in the region \mathbb{V} contains the Shannon edge term and distillable entanglement in (53), but not the log dim R edge term. From the point of view of the infinitely massive charges that we introduced in this section, this isn't surprising. The edge term basically measures the correlations of the charges, that aren't part of our gauge-invariant operator algebra.

Note that the expression "log dim R " is the expectation value of an IR operator which is a function of the gauge group Casimirs. Hence, the result from the extended Hilbert space prescription is gauge-invariant despite that we got it from carrying around gauge-variant data.

²⁶ The conventions in these references differ by some cosmetic details. In [13], the entangling cut intersects a collection of boundary links instead of boundary sites, and one extends the Hilbert space by adding a new lattice site at each such intersection where one doesn't impose the Gauss law. In [17], the extended Hilbert space on the lattice is taken to be the "globally extended" one, $\mathcal{H}_{\mathbb{V}} = \otimes \mathcal{H}_{\ell}$, instead of just lifting the Gauss law at the boundary sites. But as explained in section 4, an extension of the Hilbert space away from the entangling cut doesn't affect the entanglement entropy of states in the gauge-invariant subspace.

8.1.2. In the continuum

As mentioned above, one way to define the extended Hilbert space on the lattice is to lift the Gauss constraint at boundary lattice sites. The naively analogous operation in the continuum would be to quantize the phase space after partial gauge-fixing in the time direction (i.e. $A_0 = 0$), but without then imposing the residual gauge symmetry at the entangling surface $\partial\mathbb{V}$, instead promoting the time-independent gauge symmetry to a global symmetry along the entangling surface. (See the discussion below eq. (44).) The embedding $\mathcal{H}_{\text{phys}} \subset \mathcal{H}_{\text{ext}}$ is then fixed by demanding that the states in $\mathcal{H}_{\text{phys}}$ are singlets w.r.t. this new global symmetry.

We will say more about the formal justification for this operation below, but for now let us put it to use in an example. By following these steps we can recover the topological entanglement entropy (TEE) in Chern-Simons (CS) theory in $d = 2$, as recently discussed in [47,48]. Consider the CS theory on a spacetime $\mathbb{M} \times \mathbb{R}$. Its action is

$$S_{\text{CS}} = \frac{k}{4\pi} \int_{\mathbb{M} \times \mathbb{R}} \text{Tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right). \quad (55)$$

On a compact \mathbb{M} , its Hilbert space is the space of conformal blocks of the chiral WZW model on \mathbb{M} with gauge group G and level k [49,50]. In particular, the Hilbert space of the CS theory on $\mathbb{M} = \mathbb{S}^2$ is one-dimensional and can't be meaningfully partitioned. On the other hand, the TEE of the Abelian CS theory across a disk $\mathbb{V} \subset \mathbb{M}$ is [4]

$$\gamma_{\text{CS}} = \frac{1}{2} \log k. \quad (56)$$

This is the universal piece of the entanglement entropy $S_{\mathbb{V}}$ (see eq. (46)), which is clearly nonzero for $k > 1$. This raises the question of what it is that we are counting.

We can answer this by looking at the canonical quantization of CS theory on a disk, $\mathbb{M} = \mathbf{D}^2$ [50]. Let us briefly review the steps in that paper. We first chose the gauge $A_0 = 0$, so that the action becomes

$$S_{\text{CS}} = \frac{k}{4\pi} \int_{\mathbb{M} \times \mathbb{R}} \text{Tr}(\tilde{A} \partial_t \tilde{A}), \quad (57)$$

where the tilde means that we are restricting to the space components. We should also keep track of the Gauss law $F_{12} = 0$ and the residual time-independent gauge symmetry. The Gauss law can be solved on the disk with $\tilde{A} = \tilde{d}U U^{-1}$, for \tilde{d} the spatial exterior derivative, and $U \in G$. Plugging this into (57), we find

$$S_{\text{CS}} = S_{\text{WZW}} = \frac{k}{4\pi} \int_{\partial\mathbb{M}} \text{Tr}(U^{-1} \partial_\phi U U^{-1} \partial_t U) d\phi dt + \frac{k}{12\pi} \int_{\mathbb{M}} \text{Tr}(U^{-1} dU)^3, \quad (58)$$

where ϕ is the angular coordinate on $\partial\mathbb{M}$. This is the chiral WZW action. At this point, the authors of [50] conclude that the CS Hilbert space on the disk is the Hilbert space of the WZW theory on the boundary.

According to the bookkeeping so far, we seem to have decided by fiat to forget about the residual time-independent gauge symmetry at the boundary of the disk, promoting it to be a global symmetry instead. Historically, the argument was that the action (55) is not gauge-invariant in the presence of a boundary, and indeed it was observed early on that this choice is needed to consistently glue Euclidean path integrals along open cuts (i.e. to get $Z_M = \int [d\psi] Z_{(M_1, \Sigma)}[\psi] Z_{(M_2, \Sigma)}[\psi]$ when we cut a closed manifold M into M_1, M_2 along Σ) [51]. In

our present context though, this construction of $\mathcal{H}_{\mathbf{D}^2}$ allows us to utilize the extended Hilbert space prescription, with

$$\mathcal{H}_{\text{ext.}} = \mathcal{H}_{\mathbf{D}^2} \otimes \mathcal{H}_{\mathbf{D}^2}. \quad (59)$$

To finish computing the entanglement entropy of the disk \mathbb{V} , all that remains is for us to decide how the unique state on \mathbf{S}^2 is embedded in the Hilbert space of the two disks. The defining property is that it should transform oppositely under the global symmetries of the two WZW CFT's in order to form a singlet, as the continuum analog of pairing up the matrix indices in (51). This uniquely specifies the embedding of the state on \mathbf{S}^2 in $\mathcal{H}_{\text{ext.}}$ to be

$$|I\rangle = \sum_n |\mathbb{1}, \bar{n}\rangle \otimes |\mathbb{1}, n\rangle, \quad (60)$$

where the sum runs over the (infinitely many) descendant states in the conformal module of the primary field associated with the identity representation of G .

Finally, tracing out one of the tensor factors in (60) with a thermal regulator $e^{-\epsilon H}$ [52,53], we find

$$S_{\text{CS}} = (\text{non-universal terms}) - \frac{1}{2} \log k. \quad (61)$$

In this way we see that the prescription described above is able to correctly compute the TEE.²⁷

The origin of the continuum extended Hilbert space prescription can be understood more rigorously as follows.²⁸ For a gauge theory on a space with a boundary, requiring the boundary part of the variation of the action to vanish usually leads to a *set* of allowed types of boundary conditions that we can pick from to define the theory on that space.

For instance, consider the generator of time-independent gauge transformations $Q = \int_{\mathbb{M}} \vec{E} \cdot \vec{\nabla} \lambda$ in a Maxwell theory, with λ the gauge parameter. (The generator for time-dependent gauge transformations doesn't have a boundary term on a manifold with spatial boundaries.) Integrating by parts, we see that on a closed manifold this operator vanishes by the Gauss law, but in the presence of a boundary,

$$Q = \int_{\partial \mathbb{M}} \lambda E_{\perp}, \quad (62)$$

where E_{\perp} is the component of the electric field perpendicular to $\partial \mathbb{M}$. As long as Q isn't set to zero by the choice of boundary conditions, it will act nontrivially on the states satisfying the Gauss law, and can be seen as the generator of a large gauge transformation, i.e., a global symmetry that acts on the boundary. (More precisely, in this case the boundary gauge transformations are not zero modes of the symplectic form [54].) In the Maxwell theory, we can set $Q = 0$ by choosing the electric conditions $E_{\perp} = 0$.

²⁷ One somewhat unsatisfactory aspect of this calculation is that the factor of k was buried in the thermal regulator, so although we are counting edge modes of the topological medium, we can't point to " \sqrt{k} microscopic entangled degrees of freedom". A related situation where we can be somewhat more precise is if we quantize Chern-Simons theory on \mathbf{S}^2 with a Wilson line ending on static non-dynamical charges. In this case, repeating the above steps, we would find eq. (60) for the analogous state corresponding to the representation assigned to the Wilson line, and a TEE differing from the identity representation by $\log \dim R$. This is the $\log \dim R$ type edge term of the Wilson line, quantifying the maximal entanglement of the static charges.

²⁸ We thank Ronak Soni for discussions on these points.

In Chern-Simons theory, like in any other theory with symmetry-protected edge states, there is no choice of boundary conditions which can kill off its version of Q . A straightforward way to see this is to consider the gauge variation of the Abelian Chern-Simons action in the presence of the boundary,

$$S_{\text{CS}} \mapsto S_{\text{CS}} + \frac{k}{4\pi} \int_{\partial\mathbb{M} \times \mathbb{R}} \lambda E_{\parallel}. \quad (63)$$

If the gauge constraint isn't relaxed at the boundary, the electric fields along it will have to vanish to make $Q = 0$, but this is impossible because the CS action requires there to be no magnetic flux, even right by the boundary, and it is impossible to enforce both $E_{\parallel} = 0$ and $B = 0$ at $\partial\mathbb{M}$, as these variables don't commute.

This analysis shows that both Chern-Simons and Maxwell theory admit consistent boundary conditions in which $A_0|_{\partial\mathbb{M}} = 0$ is fixed using time-dependent gauge transformations, no additional constraint is imposed on the gauge-invariant fields at $\tau = 0$,²⁹ and the appropriate generator Q of time-independent gauge transformations is nonvanishing on the boundary. Choosing these boundary conditions on $\partial\mathbb{M}$ and its complement allows us to construct the continuum extended Hilbert space with Q having a nontrivial action on the edge modes.

8.2. Path integral boundary conditions

In section 6, we explained how the entanglement calculation for different choices of gauge-invariant subalgebras on the lattice lead to different boundary conditions in the replica trick path integral. Which path integral boundary conditions does the extended Hilbert space prescription correspond to, even when applied to theories with no known lattice regulator? This question turns out to have the following nice answer in a wide range of compact gauge theories: *the extended Hilbert space prescription corresponds to open boundary conditions in the replica trick*.

For theories that we can discretize, this fact can be understood from the perspective of the earlier sections as follows. On the lattice, the extended Hilbert space definition equals the entropy of an extended, gauge-variant operator algebra that includes all the open Wilson lines ending on the entangling surface (see [17,21] for a related discussion). This larger algebra has no center (discounting the Gauss operators in the interior), so the path integral should have no constraints.

Several instances of this general claim (including also in Chern-Simons theories that we don't know how to discretize), were discussed piecewise in the literature. We finish this section by listing them here.

8.2.1. The replica trick in solvable gauge theories

In the Hartle-Hawking states of $d = 1$ Yang-Mills theory [16,56] and $d = 2$ CS theory [57] that are set up by a Euclidean path integral on a half-sphere, for any choice of entangling region, suppose we assume that the conical defect in the n -replicated manifold can be smoothed out without changing the topology of the manifold. E.g. for CS theory on a spatial \mathbb{S}^2 , suppose that the n -replicated manifold for the entanglement entropy of a region \mathbb{V} with disk topology has the

²⁹ Technically, this is true if we take the subregion to be a causal diamond instead of $\mathbb{V} \times \mathbb{R}$. E.g. for Maxwell theory, the boundary conditions that lead to nontrivial large gauge transformations are $A_0|_{\partial\mathbb{M}} = 0$ and $B_{\parallel} = 0$ on $\mathbb{V} \times \mathbb{R}$. The latter constrains a gauge-invariant operator so these BC's do not implement the extended Hilbert space prescription. But on a causal diamond, the second BC becomes ρB_{\parallel} for ρ the local Rindler radius, that vanishes on $\partial\mathbb{V}$ [55].

topology of a smooth \mathbf{S}^3 . Then the replica trick becomes quite straightforward to implement. E.g. for the CS theory,

$$\frac{Z_n}{Z_1^n} = \frac{Z(\mathbf{S}^3)}{Z(\mathbf{S}^3)^n} = (S_0^0)^{1-n} \implies S_V = -\partial_n \left. \frac{Z_n}{Z_1^n} \right|_{n=1} = \log S_0^0, \quad (64)$$

where we made use of formulas relating the CS partition function on three-manifolds to matrix elements of the modular S-matrix [58], e.g., $Z(\mathbf{S}^3) = S_0^0$.³⁰

A natural question is which Lorentzian prescription this “naive replica trick” corresponds to. In the examples that were studied in the literature, the answer is the extended Hilbert space prescription. The CS calculation (64) is one example, since both the “naive replica trick” and the extended Hilbert space prescription correctly reproduce the TEE. Also in $d = 1$ Yang-Mills theory, the “naive replica trick” includes the $\log \dim R$ -type edge term, (52), as an empirical fact (see ref. [16] for the calculation). In cases where the gauge theory can be put on a lattice (which is clear for Yang-Mills but not necessarily Chern-Simons theory), this result is merely a manifestation of our general statement above, where the extended Hilbert space corresponds to open boundary conditions at the entangling cut that then can be contracted to a point in the continuum limit. See [61–63] for related comments.

8.2.2. Conformal anomaly for a $U(1)$ gauge theory

In a continuum QFT, the entanglement entropy is dominated by UV-sensitive divergent terms. However, in even (odd) spacetime dimensions, the coefficient of the log-divergent (constant) term in an expansion in powers of the UV cutoff is expected to be universal, (46).

For a Maxwell theory in $d = 3$, the computation of the log coefficient in the entanglement entropy of a ball-shaped region was historically the subject of some dispute. This theory is conformal. In CFTs, the log coefficient of the entanglement entropy of a ball-shaped region is a (known) function of the central charges and the extrinsic curvature of the entangling boundary on symmetry grounds. In a $d = 3$ CFT, this universal term is proportional to the a -anomaly [39], and equals $\frac{31}{45} \log(R/\epsilon)$ (with R a length scale associated to the entangling region and ϵ the UV cutoff) for the $U(1)$ theory in particular.

However, a different answer was found in the literature [64] using what is now called the CHM map [65]. A general method to compute the entropy a ball-shaped region of a CFT, that relies on just the conformal symmetry, is to map the domain of dependence of the ball-shaped region to a hyperbolic or de Sitter spacetime, and the entanglement entropy to the thermal entropy on this space [65]. One can then find the thermal entropy by integrating the energy density with $dE = T dS$, or by finding the partition function on the dS /hyperbolic space with a periodic time β , and using the geometric entropy formula³¹

$$S = (1 - \beta \partial_\beta) \log Z(\beta)|_{\beta=2\pi}. \quad (65)$$

From the dS energy density, Dowker [64] found a different answer, $a_{\text{eff}} = 16/45$ (see also ref. [67]).

³⁰ See [59], [60] for replica trick calculations in the dual WZW theories.

³¹ Note that in the CHM map to hyperbolic space, $Z(\beta)$ refers to the partition function on $\mathbf{H}^{d-1} \times \mathbf{S}_\beta^1$ which is a smooth space without a conical singularity for all β . On the other hand, if we are doing the replica trick on a space where we use the geometry to provide the analytic continuation [66], as e.g. when computing the entropy in the Rindler wedge (section 8.2.3), $Z(\beta)$ refers to the partition function on a manifold which is singular for $\beta \neq 2\pi$.

The hyperbolic analog of this discrepancy was clearly explained in [68] (see also refs. [69] for the analogous explanation in dS space, which requires treating a conical singularity and is similar to section 8.2.3 below, as well as [21], which showed by directly analyzing the entanglement entropy of $U(1)$ gauge theory on the lattice that the Shannon edge term contributes the difference of $1/3$). The upshot is that the CHM prescription by itself is incomplete: in the thermal entropy problem, one also has to pick boundary conditions as the analog of the boundary conditions at the entangling surface, and different choices give different answers. In particular, if we replace the entanglement entropy problem with the problem of computing the partition function of a $U(1)$ gauge theory in hyperbolic space, we have to pick boundary conditions at the asymptotic boundary. Computing the thermal entropy with implicitly fixed asymptotic boundary conditions in a heat kernel approach [68], one finds $a_{\text{eff}} = 16/45$. However, one can account for the discrepancy by adding to the heat kernel calculation the partition function of a free scalar on S^2 , which is the amount of residual gauge symmetry on the asymptotic boundary after we fixed the time-dependent gauge symmetry. Hence, the extended Hilbert space prescription gives the correct conformal anomaly, and the difference of $1/3$ comes from the Shannon edge term.

Incidentally, this discussion is an explicit example of the principle that the choice of prescription for computing the entropy affects its universal term; see section 6.³²

8.2.3. Kabat contact term in the geometric entropy

For $d = 3$ Maxwell theory, another longstanding and closely related puzzle in the literature was the following. Consider the entanglement entropy of the half-space (Rindler wedge) \mathbb{V} in the vacuum of the Maxwell theory. If we apply the geometric entropy formula (65) to compute the entanglement entropy, where $Z(\beta)$ is the partition function on a manifold with a conical singularity at the origin for $\beta \neq 2\pi$, the entropy will acquire a contact term with the conical singularity [71], basically because the spin-1 Laplacian is $\square\delta_{\mu\nu} + R_{\mu\nu}$. For twenty years it was an open problem if this “Kabat contact term” had a state-counting interpretation.

This was recently answered in the affirmative by Donnelly and Wall [35,36] who explicitly showed the following: Kabat’s $Z(\beta)$ is equal to the partition function Z_{bulk} on the replicated space with a tubular region removed around the conical defect at $\partial\mathbb{V}$ and the boundary condition $E_{\perp} = 0$ placed there, multiplied by

$$Z_{\text{edge}} = \int [dE_{\perp}] e^{-I(E_{\perp})}, \quad (66)$$

where $I(E_{\perp})$ is the on-shell Euclidean action of the edge modes (the continuum version of $I(k)$ from eq. (40)), and the measure is obtained by taking the continuum limit of the discrete lattice measure. In short, we impose open boundary conditions around the defect. Here we note that we can use $E_{\perp} = 0$ to compute the entropy in each superselection sector; the distillable entropy is independent of E_{\perp} , since the theory is free. Thus the path integral factorizes into $Z = Z_{\text{bulk}} Z_{\text{edge}}$. Hence, the statistical interpretation of the Kabat contact term is that it counts the Shannon edge modes of the $U(1)$ gauge theory.

³² Relatedly, in case the Maxwell theory emerges from massive but dynamical charges, one finds a third value for the log coefficient, $a_{\text{eff}} = \frac{16}{45} + 1$ [24], [70]. In this case, the dynamical charges are screened with a thermal entropy in a layer of size m^{-1} around the entangling boundary.

8.3. Short remarks

8.3.1. Emergence

Until now, our discussion of the extended Hilbert space was formal, and perhaps in poor taste from the usual point of view that gauge symmetries are fictional. A physical justification for it is that if we replace the extended Hilbert space with a local UV Hilbert space in an emergent gauge theory, then (53) will hold up to a constant that is independent of the state in the IR. One situation where this is trivially true is the toric code and “globally extended Hilbert space” of section 4.

Ref. [72] argued that it is true more generally. The argument is the following. On the one hand, in a low-energy emergent gauge theory, the UV reconstructions of IR Wilson loops as composite operators are able to factorize along $\partial\mathbb{V}$ by assumption. On the other hand, if we think of the entanglement entropy as an entropy of the maximal UV operator algebra on the region, the explicit expansion of the density matrix, eq. (5), has zero support on UV operators that are not either the UV analogs of gauge-invariant operators or of Wilson lines ending on $\partial\mathbb{V}$. So the UV-exact entanglement entropy can only differ from the entropy of the “IR gauge-invariant algebra” by a constant related to the relative sizes of the Hilbert spaces in the UV and the IR.³³

8.3.2. Duality

An open question is how the edge terms map under duality. Euclidean replica trick results for (Seiberg-)dual nonabelian gauge theories appear to agree [73]. This result can be decomposed into the sum of distillable entropy and the entropy coming from edge terms. Given an exact duality map for all operators and states, the algebraically defined entropies $S_{\mathbb{V}}^{\text{g-inv}}$ of dual algebras must trivially agree, but the “log dim R ” edge terms are more mysterious, since they seem to depend explicitly on the gauge group. Do the “log dim R ”-type edge terms dualize to edge terms nonetheless? Preliminary evidence suggests yes [74], but it would be nice to clarify this in an analysis along the lines of ref. [75] for Abelian gauge theories.

9. Conclusion

We have given precise path integral prescriptions for calculating various entropies associated to a given subalgebra $\mathcal{A}_{\mathbb{V}}$. Our approach is elementary and applies to a wide array of theories, and we have shown how to connect it to various other recipes for calculating entropy in QFTs. While the results of this paper do not facilitate the computation of particular replica trick path integrals, they do show how they should be understood in any formal proof concerning properties of entanglement entropy. Several interesting questions remain open, however.

As emphasized in section 5, several different yet natural types of entanglement entropy exist for each choice of algebra on a given region. We have seen that certain special prescriptions for calculating entropy (holography and CHM-like maps in conformal theories with appropriate boundary conditions) correspond to the full entropy of the maximal algebra $\mathcal{A}_{\mathbb{V}}$ on \mathbb{V} in a small set of theories (holographic theories with Einstein gravity duals, the Ising CFT, free theories). An immediate goal is to fill in the blanks as much as possible. For instance, what bulk quantities calculate the full entropy of a non-maximal algebra? Answers to such questions may teach us more about the bulk reconstruction of operators.

³³ This constant arises from the different definitions of the trace used in calculating the von Neumann entropy, as discussed around eq. (54).

Another set of notions that we have not studied involves extending our discussion to other measures of entanglement, such as mutual information, entanglement negativity, covariant entanglement entropies, etc. In particular, it would be interesting to find the analogues of the various choices directly in the Tomita-Takesaki framework for continuum QFTs [11]. An important ingredient in any analysis of the monotonicity of entropic measures will be a careful definition of which representation of \mathcal{A}_V is appropriate for a given type of center.

Finally, we remark that it may be possible to associate entropies (or other measures of entanglement) to sets of operators that are not necessarily algebras [76]. This provides yet another direction in which our analysis of centers and boundary conditions could be generalized.

CRediT authorship contribution statement

Jennifer Lin: Conceptualization, Investigation, Writing - original draft. **Đorđe Radičević:** Conceptualization, Investigation, Writing - original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Rob Myers, Tatsuma Nishioka, and Kantaro Ohmori for useful conversations, and Lorenzo Di Pietro and Ronak Soni for discussions and comments on a draft of this paper. JL is supported by the William D. Loughlin Membership at the IAS and by The U.S. Department of Energy, and would also like to thank the Perimeter Institute and Galileo Galilei Institute for hospitality as this work was being completed. ĐR thanks the Galileo Galilei Institute and the ACRI fellowship, the University of Tokyo, and the Kavli Institute for the Physics and Mathematics of the Universe for hospitality and support while this work was in progress. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation.

References

- [1] M. Srednicki, Entropy and area, *Phys. Rev. Lett.* 71 (1993) 666–669, arXiv:hep-th/9303048.
- [2] G. Vidal, J.I. Latorre, E. Rico, A. Kitaev, Entanglement in quantum critical phenomena, *Phys. Rev. Lett.* 90 (2003) 227902, arXiv:quant-ph/0211074.
- [3] P. Calabrese, J.L. Cardy, Entanglement entropy and quantum field theory, *J. Stat. Mech.* 0406 (2004) P06002, arXiv:hep-th/0405152.
- [4] A. Kitaev, J. Preskill, Topological entanglement entropy, *Phys. Rev. Lett.* 96 (2006) 110404, arXiv:hep-th/0510092.
- [5] S. Ryu, T. Takayanagi, Holographic derivation of entanglement entropy from AdS/CFT, *Phys. Rev. Lett.* 96 (2006) 181602, arXiv:hep-th/0603001.
- [6] H. Li, F.D.M. Haldane, Entanglement spectrum as a generalization of entanglement entropy: identification of topological order in non-Abelian fractional quantum hall effect states, *Phys. Rev. Lett.* 101 (July 2008) 010504, arXiv:0805.0332.
- [7] M. Van Raamsdonk, Building up spacetime with quantum entanglement, *Gen. Relativ. Gravit.* 42 (2010) 2323–2329, arXiv:1005.3035, *Int. J. Mod. Phys. D* 19 (2010) 2429.
- [8] L. Amico, R. Fazio, A. Osterloh, V. Vedral, Entanglement in many-body systems, *Rev. Mod. Phys.* 80 (2008) 517–576, arXiv:quant-ph/0703044.
- [9] T. Nishioka, Entanglement entropy: holography and renormalization group, arXiv:1801.10352.

- [10] J. Preskill, Quantum information and physics: some future directions, *J. Mod. Opt.* 47 (2000) 127–137, arXiv: quant-ph/9904022.
- [11] E. Witten, Notes on some entanglement properties of quantum field theory, arXiv:1803.04993.
- [12] P.V. Buividovich, M.I. Polikarpov, Entanglement entropy in gauge theories and the holographic principle for electric strings, *Phys. Lett. B* 670 (2008) 141–145, arXiv:0806.3376.
- [13] W. Donnelly, Decomposition of entanglement entropy in lattice gauge theory, *Phys. Rev. D* 85 (2012) 085004, arXiv:1109.0036.
- [14] H. Casini, M. Huerta, J.A. Rosabal, Remarks on entanglement entropy for gauge fields, *Phys. Rev. D* 89 (8) (2014) 085012, arXiv:1312.1183.
- [15] D. Radičević, Notes on entanglement in Abelian gauge theories, arXiv:1404.1391.
- [16] W. Donnelly, Entanglement entropy and nonabelian gauge symmetry, *Class. Quantum Gravity* 31 (21) (2014) 214003, arXiv:1406.7304.
- [17] S. Ghosh, R.M. Soni, S.P. Trivedi, On the entanglement entropy for gauge theories, *J. High Energy Phys.* 09 (2015) 069, arXiv:1501.02593.
- [18] R.M. Soni, S.P. Trivedi, Aspects of entanglement entropy for gauge theories, *J. High Energy Phys.* 01 (2016) 136, arXiv:1510.07455.
- [19] P. Zanardi, D.A. Lidar, S. Lloyd, Quantum tensor product structures are observable induced, *Phys. Rev. Lett.* 92 (2004) 060402, arXiv:quant-ph/0308043.
- [20] M. Ohya, D. Petz, *Quantum Entropy and Its Use*, Springer Science & Business Media, 2004.
- [21] R.M. Soni, S.P. Trivedi, Entanglement entropy in $(3 + 1)$ -d free $U(1)$ gauge theory, *J. High Energy Phys.* 02 (2017) 101, arXiv:1608.00353.
- [22] D.L. Jafferis, A. Lewkowycz, J. Maldacena, S.J. Suh, Relative entropy equals bulk relative entropy, *J. High Energy Phys.* 06 (2016) 004, arXiv:1512.06431.
- [23] M. Levin, X.-G. Wen, Detecting topological order in a ground state wave function, *Phys. Rev. Lett.* 96 (2006) 110405.
- [24] M. Pretko, T. Senthil, Entanglement entropy of $U(1)$ quantum spin liquids, *Phys. Rev. B* 94 (12) (2016) 125112, arXiv:1510.03863.
- [25] K. Ohmori, Y. Tachikawa, Physics at the entangling surface, *J. Stat. Mech.* 1504 (2015) P04010, arXiv:1406.4167.
- [26] Y.-A. Chen, A. Kapustin, D. Radičević, Exact bosonization in two spatial dimensions and a new class of lattice gauge theories, *Ann. Phys.* 393 (2018) 234–253, arXiv:1711.00515.
- [27] Y.-A. Chen, A. Kapustin, Bosonization in three spatial dimensions and a 2-form gauge theory, arXiv:1807.07081.
- [28] S. Elitzur, Impossibility of spontaneously breaking local symmetries, *Phys. Rev. D* 12 (1975) 3978–3982.
- [29] A.Yu. Kitaev, Fault tolerant quantum computation by anyons, *Ann. Phys.* 303 (2003) 2–30, arXiv:quant-ph/9707021.
- [30] D. Radičević, Entanglement entropy and duality, *J. High Energy Phys.* 11 (2016) 130, arXiv:1605.09396.
- [31] U. Moitra, R.M. Soni, S.P. Trivedi, Entanglement entropy, relative entropy and duality, *J. High Energy Phys.* 08 (2019) 59, arXiv:1811.06986.
- [32] D. Radičević, Entanglement in weakly coupled lattice gauge theories, *J. High Energy Phys.* 04 (2016) 163, arXiv:1509.08478.
- [33] A. Altland, B.D. Simons, *Condensed Matter Field Theory*, Cambridge University Press, 2010.
- [34] J.B. Kogut, L. Susskind, Hamiltonian formulation of Wilson's lattice gauge theories, *Phys. Rev. D* 11 (1975) 395–408.
- [35] W. Donnelly, A.C. Wall, Entanglement entropy of electromagnetic edge modes, *Phys. Rev. Lett.* 114 (11) (2015) 111603, arXiv:1412.1895.
- [36] W. Donnelly, A.C. Wall, Geometric entropy and edge modes of the electromagnetic field, *Phys. Rev. D* 94 (10) (2016) 104053, arXiv:1506.05792.
- [37] P. Liendo, L. Rastelli, B.C. van Rees, The bootstrap program for boundary CFT_d , *J. High Energy Phys.* 07 (2013) 113, arXiv:1210.4258.
- [38] J.L. Cardy, Boundary conditions, fusion rules and the verlinde formula, *Nucl. Phys. B* 324 (1989) 581–596.
- [39] S. Ryu, T. Takayanagi, Aspects of holographic entanglement entropy, *J. High Energy Phys.* 08 (2006) 045, arXiv: hep-th/0605073.
- [40] H. Casini, M. Huerta, A finite entanglement entropy and the c-theorem, *Phys. Lett. B* 600 (2004) 142–150, arXiv: hep-th/0405111.
- [41] H. Casini, M. Huerta, On the RG running of the entanglement entropy of a circle, *Phys. Rev. D* 85 (2012) 125016, arXiv:1202.5650.
- [42] S.N. Solodukhin, Entanglement entropy, conformal invariance and extrinsic geometry, *Phys. Lett. B* 665 (2008) 305–309, arXiv:0802.3117.

- [43] A. Lewkowycz, J. Maldacena, Generalized gravitational entropy, *J. High Energy Phys.* 08 (2013) 090, arXiv:1304.4926.
- [44] H. Casini, M. Huerta, R.C. Myers, A. Yale, Mutual information and the F-theorem, *J. High Energy Phys.* 10 (2015) 003, arXiv:1506.06195.
- [45] L.-Y. Hung, R.C. Myers, M. Smolkin, On holographic entanglement entropy and higher curvature gravity, *J. High Energy Phys.* 04 (2011) 025, arXiv:1101.5813.
- [46] S. Cordes, G.W. Moore, S. Ramgoolam, Lectures on 2-d Yang-Mills theory, equivariant cohomology and topological field theories, *Nucl. Phys. Proc. Suppl.* 41 (1995) 184–244, arXiv:hep-th/9411210.
- [47] J.R. Fliss, X. Wen, O. Parrikar, C.-T. Hsieh, B. Han, T.L. Hughes, R.G. Leigh, Interface contributions to topological entanglement in Abelian Chern-Simons theory, *J. High Energy Phys.* 09 (2017) 056, arXiv:1705.09611.
- [48] G. Wong, A note on entanglement edge modes in Chern Simons theory, arXiv:1706.04666.
- [49] E. Witten, Quantum field theory and the Jones polynomial, *Commun. Math. Phys.* 121 (1989) 351–399.
- [50] S. Elitzur, G.W. Moore, A. Schwimmer, N. Seiberg, Remarks on the canonical quantization of the Chern-Simons-Witten theory, *Nucl. Phys. B* 326 (1989) 108–134.
- [51] E. Witten, On holomorphic factorization of WZW and coset models, *Commun. Math. Phys.* 144 (1992) 189–212.
- [52] X. Wen, S. Matsuura, S. Ryu, Edge theory approach to topological entanglement entropy, mutual information and entanglement negativity in Chern-Simons theories, *Phys. Rev. B* 93 (24) (2016) 245140, arXiv:1603.08534.
- [53] D. Das, S. Datta, Universal features of left-right entanglement entropy, *Phys. Rev. Lett.* 115 (13) (2015) 131602, arXiv:1504.02475.
- [54] W. Donnelly, L. Freidel, Local subsystems in gauge theory and gravity, *J. High Energy Phys.* 09 (2016) 102, arXiv:1601.04744.
- [55] A. Blommaert, T.G. Mertens, H. Verschelde, V.I. Zakharov, Edge state quantization: vector fields in rindler, arXiv:1801.09910.
- [56] A. Gromov, R.A. Santos, Entanglement entropy in 2D non-Abelian pure gauge theory, *Phys. Lett. B* 737 (2014) 60–64, arXiv:1403.5035.
- [57] S. Dong, E. Fradkin, R.G. Leigh, S. Nowling, Topological entanglement entropy in Chern-Simons theories and quantum Hall fluids, *J. High Energy Phys.* 05 (2008) 016, arXiv:0802.3231.
- [58] E. Witten, (2+1)-dimensional gravity as an exactly soluble system, *Nucl. Phys. B* 311 (1988) 46.
- [59] H.J. Schnitzer, Topological Rényi and entanglement entropy for a 2d q-deformed $U(N)$ Yang-Mills theory and its Chern-Simons dual, arXiv:1605.08075.
- [60] H.J. Schnitzer, Rényi entropy for a $2d$ CFT with a gauge field: $\widehat{SU}(N)_1$ WZW theory on a branched torus, arXiv:1611.03116.
- [61] D.L. Jafferis, Factorization in gravity, Talk at Bariloche IFQ meeting, <http://fisica.cab.cnea.gov.ar/itfromqubit/images/talks/jafferisBariloche.pdf>, 2018.
- [62] J. Cardy, E. Tonni, Entanglement Hamiltonians in two-dimensional conformal field theory, *J. Stat. Mech.* 1612 (12) (2016) 123103, arXiv:1608.01283.
- [63] W. Donnelly, G. Wong, Entanglement branes in a two-dimensional string theory, *J. High Energy Phys.* 09 (2017) 097, arXiv:1610.01719.
- [64] J.S. Dowker, Entanglement entropy for even spheres, arXiv:1009.3854.
- [65] H. Casini, M. Huerta, R.C. Myers, Towards a derivation of holographic entanglement entropy, *J. High Energy Phys.* 05 (2011) 036, arXiv:1102.0440.
- [66] C.G. Callan Jr., F. Wilczek, On geometric entropy, *Phys. Lett. B* 333 (1994) 55–61, arXiv:hep-th/9401072.
- [67] H. Casini, M. Huerta, Entanglement entropy of a Maxwell field on the sphere, *Phys. Rev. D* 93 (10) (2016) 105031, arXiv:1512.06182.
- [68] K.-W. Huang, Central charge and entangled gauge fields, *Phys. Rev. D* 92 (2) (2015) 025010, arXiv:1412.2730.
- [69] C. Eling, Y. Oz, S. Theisen, Entanglement and thermal entropy of gauge fields, *J. High Energy Phys.* 11 (2013) 019, arXiv:1308.4964.
- [70] M. Pretko, On the entanglement entropy of Maxwell theory: a condensed matter perspective, arXiv:1801.01158.
- [71] D.N. Kabat, Black hole entropy and entropy of entanglement, *Nucl. Phys. B* 453 (1995) 281–299, arXiv:hep-th/9503016.
- [72] J. Lin, Ryu-Takayanagi area as an entanglement edge term, arXiv:1704.07763.
- [73] T. Nishioka, I. Yaakov, Supersymmetric Renyi entropy, *J. High Energy Phys.* 10 (2013) 155, arXiv:1306.2958.
- [74] A. Lewkowycz, J. Maldacena, Exact results for the entanglement entropy and the energy radiated by a quark, *J. High Energy Phys.* 05 (2014) 025, arXiv:1312.5682.
- [75] W. Donnelly, B. Michel, A. Wall, Electromagnetic duality and entanglement anomalies, *Phys. Rev. D* 96 (4) (2017) 045008, arXiv:1611.05920.
- [76] S. Ghosh, S. Raju, Quantum information measures for restricted sets of observables, arXiv:1712.09365.