# Entanglement in Gauge Theories

A Ph.D. thesis

submitted to
**Tata Institute of Fundamental Research, Mumbai**

for the degree of

Doctor of Philosophy
in
Physics

by

Ronak M Soni

Department of Theoretical Physics
Tata Institute of Fundamental Research, Mumbai
Mumbai - 400005, India

June 2018
Final Version Submitted September 2018

# Declaration of Authorship

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Professor Sandip P. Trivedi, at the Tata Institute of Fundamental Research, Mumbai.

**Ronak M Soni**

In my capacity as the supervisor of the candidates thesis, I certify that the above statements are true to the best of my knowledge.

**Prof. Sandip P. Trivedi**
Date:

यारों बहाम गुंथे हुए कायनात के बिखरे टुकड़े,
एक फुल को जुंबिश दोगे तो तारा कांप उठेगा।

So entangled are all the scattered corners of creation
that
poke a flower, and a star shall quiver.

- रघुपति सहाय 'फ़िरक़ गोरखपुरी'
Raghupati Sahai 'Firaaq Gorakhpuri'

*What's my gauge? It is nor hand nor foot,*
*Nor arm, not face, nor any other part*
*Belonging to a state. O, be some other gauge!*
*What's in a gauge? That which we call a rose*
*By any other gauge would smell as sweet;*

- William Shakespeare[1]

[1] Citaton needed.

# *Abstract*

Tata Institute of Fundamental Research, Mumbai
Department of Theoretical Physics

Doctor of Philosophy

**Entanglement in Gauge Theories**

by Ronak M Soni

I discuss the problem of defining entanglement entropy between a region and its complement in a quantum field theory with gauge symmetry. There are two related approaches to defining it, one using the Hilbert space – the "extended Hilbert space" definition – and one by considerations of local subalgebras. I explain both and compare them in the case of an Abelian and a non-Abelian gauge group. Then, I interpret the entanglement found in this way in an information-theoretic manner. I then take the continuum limit and calculate it in the case $(3+1)$-d free $U(1)$ gauge theory.

# *Acknowledgements*

First and foremost, I'd like to thank everyone in the world who has directly and indirectly subsidised my life; the people whose taxes my salary comes out of, the rural and poor people on whose unconsenting backs a middle class urban life necessarily rests, the infrastructure of the government and the various institutes I've been in who have helped me personally and handled my file.

Then, of course, I'd like to thank my parents and the rest of my family for everything they've done for me, for giving birth to me, raising me, making me the person I am today, and (least importantly) for talking me out of my schoolboy whim that one time of doing commerce and following in their footsteps.

Another of the biggest influences on me and my life has been my school from seventh to twelfth, Rishi Valley School. It taught me nearly everything I know about the world beyond my own nose, thanks both to the incredibly knowledgeable and friendly faculty and my dizzyingly fast-paced peer group. Of particular note are SK, Sid Menon, Rajan, Radhika akka, Arya, Hari and Raghav.

My physics career has been possible thanks to the mentorship and training of many people, starting with my college, Chennai Mathematical Institute. CMI let me in to the physics program despite my not having got through, and I cannot thank them enough for that gesture of kindness. Even after getting into CMI, I would not have continued in physics, or at least not with the same vigour, if it hadn't been for my amazing junior Anirudh whose excitement for physics is what first woke it up in me. My B.Sc. thesis advisor Narayan was also expecially instrumental in my growth as a physicist, especially in teaching me how to make a thought precise.

But, to my mind, CMI is dwarfed a thousand times over by the mentorship I received, and continue to receive, at TIFR. My advisor Sandip gave me direction, tolerated my clueless tangents, and taught me to discard the unnecessary framework around any thought. Shiraz taught me around three-fourths of all the physics I know, and much of it without any writing technology; I learnt from him a lot about how to talk about science, and also by observing him important lessons about sociology of a scientific group. Gautam inspired me by knowing everything. Rajdeep patiently tolerated my blind flapping about for the period in which I worked on a problem with him, and one of my greatest regrets in my PhD has been not being able to drive that problem through. Rajdeep and Kedar both let me pretend to be a condensed-matter person when the whim came upon me, and I am significantly better as a physicist for it. And last: there are few joys while growing as a physicist than having a peer on your exact wavelength growing up with you, and I had the extreme good

fortune of having exactly that in Pranjal. And, of course, I would also like to thank all my collaborators, Sudip, Rohan, Vishal, Shukla, Miguel, Laurent and Florian, for putting up with me.

And now, finally, I can thank my awesome friends. Ananth for teaching me enthusiasm and universal love, Kamath for making me stay grounded, Anup for imparting to methe eternal chill. Som for the excitement in my college days, and Arpit for knowing everything and being my longest-running friend. Sarbajaya for being there for me when I needed someone, Sounak for explaining to me people and TeXing the Hindi version of the shayari and Kocherla for his criticisms and people-attracting properties. Pavithra and Kocherla for socialising me. Manoj for keeping my rebellious spark alive and giving it more kindling. Pranjal for being there and also for providing me the shayari at the front of my thesis *two years* in advance, and Rohan for his bottomless fount of folksy wisdom. Shiva, Rahul, Alok and Madhu for many many lessons about politics and the situation of the world. And of course, KA, NV and most of all SU.

# Contents

*To that rabbit I saw one time in Canada*
*and the universe reflected in its eyes.*

# List of Publications

Those that are presently relevant:

1. **On the Entanglement Entropy for Gauge Theories**,
   Sudip Ghosh, Ronak M Soni and Sandip P. Trivedi,
   JHEP **1509** 069 (2015), arXiv:1501.02593.

2. **Aspects of Entanglement Entropy for Gauge Theories**,
   Ronak M Soni and Sandip P. Trivedi,
   JHEP **1601** 136 (2016), arXiv:1510.07455.

3. **Entanglement Entropy in Free** $(3+1)$**-d Free** $U(1)$ **Gauge Theory**,
   Ronak M Soni and Sandip P. Trivedi,
   JHEP **1702** 101 (2017), arXiv:1608.00353.

Others:

1. **A Note on S-Matrix Bootstrap for Amplitudes with Linear Spectrum**,
   Pranjal Nayak, Rohan R. Poojary and Ronak M Soni,
   arXiv:1707.08135.

2. **On the Dynamics of Near-Extremal Black Holes**,
   Pranjal Nayak, Ashish Shukla, Ronak M Soni, Sandip P. Trivedi and V. Vishal,
   arXiv:1802.09547.

# Chapter 1

# Introduction

Entanglement, the fact that a state of two subsystems can't be written as independent states on each subsystem, is one of the simplest signatures of non-classicality in quantum systems, and its study in both the information theory and field theory communities have been fruitful, leading to for example new diagnostics of certain phase transitions, important advances in understanding the emergence of the bulk in AdS/CFT, the development of a numerical technique called tensor networks, and the study of quantum thermodynamics beyond the second law.

In quantum field theory, we're interested in the entanglement between a region and its complement on a spatial slice; regions are natural objects in field theory since the degrees of freedom are local in space. In gauge theories, however, the purely local description has a gauge redundancy and the physical degrees of freedom are not point-like, but are loops of electric and magnetic flux. Because of this, subregions in a gauge theory don't really correspond to subsystems, and there is a real question of whether there is a sensible entanglement-type question we can ask between subregions.

In this thesis, we show that there is indeed such a question, examine its properties and meaning, and calculate the answer in some systems. The basic physical point will be that the lack of factoriability can be thought to come purely from the Gauss' law constraint $\nabla \cdot E = 0$, or its analogue, applied as an operator equation. In particular, the problem will be caused by these constraints at the boundary of the region of interest, since because of these constraints the electric flux entering the region and leaving it at a point on the boundary are exactly the same variable, even though one should count as an inside variable and the other the outside.

Because of this relationship, the density matrix has a block-diagonal structure, with each block corresponding to a *boundary electric flux configuration*, or just *edge mode* for short,

denoted throughout by $\mathbf{k}$, and the entanglement entropy can be written as

$$S_E = -\sum_{\mathbf{k}} p_{\mathbf{k}} \log p_{\mathbf{k}} - \sum_{\mathbf{k}} p_{\mathbf{k}} \rho_{in}^{\mathbf{k}} \log \rho_{in}^{\mathbf{k}} + \sum_{\mathbf{k}} p_{\mathbf{k}} \log d_{\mathbf{k}} \equiv S_c + S_q + S_f, \tag{1.1}$$

where

1. $p_{\mathbf{k}}$ is the probability of measuring the boundary electric flux configuration to be $\mathbf{k}$,
2. $\rho_{in}^{\mathbf{k}}$ is the density matrix restricted to the sector given by $\mathbf{k}$, normalised so that $\operatorname{tr} \rho_{in}^{\mathbf{k}} = 1$, and
3. $d_{\mathbf{k}}$ is the dimension of the representation $\mathbf{k}$, and is always 1 for Abelian gauge groups.

We call the three terms $S_c, S_q, S_f$ the classical, quantum, and fusion terms respectively.

After giving a short introduction to entanglement as well as gauge theories reviewing relevant facts, we build up the framework required to arrive at the formula (1.1) in chapter 2. In it, after discussing the problems with gauge-fixing, we define EE for a gauge-invariant state — defined as being invariant under gauge transformations — on the lattice in two ways. The first is the *extended Hilbert space* (EHS) definition, in which we embed the non-factorisable gauge-invariant Hilbert space in a larger Hilbert space that factorises because it has both gauge-invariant and non-gauge-invariant states and then define the reduced density matrix as being that in this larger factorisable Hilbert space. This definition has the decomposition (1.1). Second, we introduce the algebraic *electric centre* definition, where the reduced density matrix can be thought of as a linear combination of all the gauge-invariant operators that act within the region, and show that it lacks the fusion term in the EE. We end chapter 2 by showing that the EHS definition agrees with a lattice replica trick, and show that the continuum limit of this lattice replica trick gives the same answer as the naive continuum replica trick. Since the continuum replica trick is one of the major tools used to calculate EE in QFT, this means that the EHS definition is the one of most direct physical interest.

Then, in chapter 3, we try to understand the meaning of this EE. We use the non-appearance of the fusion term in the electric centre definition to argue that it measures correlations between non-gauge-invariant operators. Using the operational definition of entanglement as being the number of Bell pairs a state can be converted to by local operations and classical communication, we show that only the quantum term $S_q$ should be thought of as real entanglement and the classical term $S_c$ should be thought of as classical correlation.

Finally, because concreteness is a virtue and also because it is useful to illustrate the fact that all three terms, despite their differing operational interpretations, contain physical information, we calculate the EE and its decomposition into the three pieces for three different theories in chapter 4.

The first system is $2+1$-dimensional lattice $SU(2)$ theory to first non-trivial order in the strong coupling expansion, which is a confining theory. Here, the EE is exactly proportional to the number of boundary vertices, and comes entirely from the classical and fusion terms. The first fact is related to the fact that the theory is confining, and the second fact comes mostly because of the simplicity of the unperturbed state.

The second is the $2+1$-dimensional toric code for a general discrete gauge group. This theory is the extreme weak-coupling limit, which is singular for continuous gauge groups. In this theory, we show that the EHS entanglement entropy is proportional not to the number of boundary vertices but one less than it. This correction is called the topological entanglement entropy; its existence is related to the theory being deconfined, and its value is a quantity known as the *total quantum dimension* of the spectrum of anyonic excitations of this theory. We also show that the value gets a non-zero contribution from the fusion piece, showing that all three terms have physical content.

Finally, we consider a ball-shaped region in free continuum Maxwell theory in $3+1$ dimensions. In this theory, conformal field theory arguments and an explicit calculation had given different answers for the universal term in the entanglement. We show that the replica trick, and therefore the EHS, answer agrees with the CFT answer, and further that the discrepancy is exactly the classical piece. Finally, we show that even in this theory there is a negative constant correction to the area law similar to the topological entanglement entropy, but that it is not topologically protected or universal any more.

We end with a summary of results and a discussion of open questions in chapter 5.

## 1.1 Entanglement Basics

Entanglement is the fact that a state on a composite system can't be written as independent states on subsystems. In the case where the Hilbert space of a full system can be written as the tensor product of the Hilbert spaces of two subsystems,

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2, \tag{1.2}$$

a *separable state* is, sticking to the pure state case, one that can be written as

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle\,, \tag{1.3}$$

and any non-separable state is called entangled.

The *entanglement entropy*(EE) is a simple measure of how bad this failure to factorise is. First, we define a reduced density matrix, defined to be a density matrix on one of the

subsystems, say $\mathcal{H}_1$, that reproduces all the expectation values of operators acting only on that subsystem,

$$\mathrm{tr}(\rho_1 O_1) = \langle\psi|O_1 \otimes \mathbb{1}_2|\psi\rangle\,. \tag{1.4}$$

$\rho_2$ is similarly defined. A more explicit expression for the reduced density matrix is

$$\left\langle i\middle|\rho_1\middle|i'\right\rangle = \sum_j \langle i,j|\psi\rangle\left\langle\psi\middle|i',j\right\rangle \equiv \left\langle i\middle|\mathrm{tr}_2\left(|\psi\rangle\langle\psi|\right)\middle|i'\right\rangle, \tag{1.5}$$

where $i$ and $j$ label bases for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively.

When the state is separable, the reduced density matrix is also pure, and when it is entangled the reduced density matrix is that of a mixed state in the first subsystem *despite the fact that the full state was pure.* Thus, the amount of entropy in the reduced density matrix is a measure of how entangled the two subsystems are, since the full state is pure and all the uncertainty comes from the entanglement. Thus, we define the EE as the von Neumann entropy of the reduced density matrix,

$$S_E(|\psi\rangle) = -\,\mathrm{tr}_1\,\rho_1 \log \rho_1. \tag{1.6}$$

While this definition may seem needlessly asymmetrical, it so happens that $\rho_1$ and $\rho_2$ have the same eigenvalues and therefore the same EE, as can be seen from the fact that any entangled state can be written in the *Schmidt decomposition*,

$$|\psi\rangle = \sum_{i,j}\psi_{ij}\,|i\rangle_1\,|j\rangle_2 = \sum_{I=1}^{\min(d_1,d_2)}\alpha_I\,|I\rangle_1\,|I\rangle_2 \quad\Rightarrow\quad \rho_i = \sum_I |\alpha_I|^2\,|I\rangle_i\,\langle I|_i\,. \tag{1.7}$$

There is also a more operational way to understand entanglement and EE. A prototypical example of an entangled state is the Bell pair, $|00\rangle + |11\rangle$.[1] One might naively think that the defining property of this state is that measurements of $\sigma^z$ on both qubits are perfectly correlated, but if one creates a pair of qubits that are with probability $\frac{1}{2}$ both 0 or both 1 the state one gets is actually $|00\rangle\langle 00| + |11\rangle\langle 11|$ instead of the Bell pair. What, one might ask, is precisely the difference? The answer was found by Bell [4], who showed that there are certain inequalities satisfied by all classically correlated systems and violated by some entangled systems, called *Bell inequalities*. A particularly convenient version, found by Clauser et al [5] is

$$\left\langle \sigma_1^z\frac{\sigma_2^z+\sigma_2^x}{\sqrt{2}} + \sigma_1^z\frac{\sigma_2^z-\sigma_2^x}{\sqrt{2}} + \sigma_1^x\frac{\sigma_2^z+\sigma_2^x}{\sqrt{2}} - \sigma_1^x\frac{\sigma_2^z-\sigma_2^x}{\sqrt{2}}\right\rangle \leq 2. \tag{1.8}$$

When evaluated in the Bell state, this expectation value is $2\sqrt{2}$.

[1] We will be cavalier about normalisation in this section.

An important question we'd like to ask about entanglement is how to measure it. If we had a class of operations that was guaranteed not to increase entanglement, we could compare the entanglement in two different states by which one can be converted into the other using operations from only that class. The standard class of operations defined for this purpose is *local operations and classical communication* (LOCC), defined as unitaries/measurements that are performed on only one subsystem and classical communication. For example, $U_1 \otimes \mathbb{1}_2, \mathbb{1}_1 \otimes U_2, U_1 \otimes U_2$ are all LOCC, but $U_1 \otimes U_2 + V_1 \otimes V_2$ isn't. Similarly, measurements have to involve degrees of freedom on only one subsystem.

Given $n$ copies of the state, if $k_{max}$ is the maximum number of Bell pairs that can be created from it using LOCC, the *distillable* or *extractable entanglement* is $\lim_{n\to\infty} \frac{k_{max}}{n}$. Similarly, if $k_{min}$ is the minimum number of shared Bell pairs required to create $n$ copies of the state using only LOCC, the *entanglement of dilution* is $\lim_{n\to\infty} \frac{k_{min}}{n}$. When the full state is pure, the two quantities coincide and are equal to the entanglement entropy.

## 1.2 Entanglement in Scalar Quantum Field Theory

States in a quantum system are usually given by wavefunctions, which are functions from a configuration space to complex numbers, where the configuration space is the analogue of the set of allowed 'positions' in the classical theory. For example, the configuration space for a qubit is two points, and that for a particle living in $\mathbb{R}^n$ is $\mathbb{R}^n$. A quantum field theory is a theory whose configuration space is a set of configurations of some field living on a particular space, so that the configuration space of a $d+1$-dimensional real/complex scalar field theory is the set of functions from $\mathbb{R}^d$ to real/complex numbers satisfying some boundary conditions.

It is conventional to associate to each configuration a basis element of the Hilbert space, so that a state can be written as

$$|\psi\rangle = \int D\phi(x)\psi[\phi(x)]\,|\phi(x)\rangle\,, \quad \big\langle \phi(x)\big|\phi'(x)\big\rangle = \delta(\phi - \phi'), \tag{1.9}$$

where $D\phi(x) = \prod_{x\in\mathbb{R}^d} \phi(x)$ integrates over all configurations, $\psi[\phi]$ is called the wavefunctional, and the $\delta$-function is defined according to $\int D\phi\delta(\phi-\phi')f[\phi] = f[\phi']$. However, as is the case in text book quantum mechanics the unnormalisable basis states $|\phi\rangle$ themselves aren't in the Hilbert space; we don't let this worry us.

Since two field configurations can differ at any point, we may write the basis element $|\phi\rangle$ as

$$|\phi\rangle = \bigotimes_x |\phi(x)\rangle_x\,, \tag{1.10}$$

and therefore the entire Hilbert space has the structure

$$\mathcal{H} = \bigotimes_x \mathcal{H}_x. \tag{1.11}$$

Making these two statements precise requires some regularisation; it is simplest to imagine all the statements in this section as being on a lattice. While we will continue to speak in continuum language in this section, we will be very explicit about the lattice in the discussion of gauge theories. [6] has a rigorous description directly in the continuum, but we will not go into those details here.

Since the natural subsystems of the full Hilbert space live on points, it is natural to define regions as subsystems in QFT. We will always discuss the entanglement between a region and its complement, which we will call "in" and "out" respectively. Here, the words "in" and "out" hold no special significance except to the human mind; they come from the picture in which "in" is a particular region in space that we have access to, for example a lab, and "out" is everywhere else.

Thus, we define a reduced density matrix for the inside

$$\rho_{in} = \int D\phi_{out} D\phi_{in} D\tilde{\phi}_{in} \, \langle \phi_{in}, \phi_{out} | \psi \rangle \left\langle \psi \middle| \tilde{\phi}_{in}, \phi_{out} \right\rangle \left| \phi_{in} \middle\rangle \middle\langle \tilde{\phi}_{in} \right|. \tag{1.12}$$

Clearly this expression only makes sense on the lattice, and consequently we find a UV-divergent answer. In four space-time dimensions, it has the form

$$S_E = c_2 \frac{A}{\epsilon^2} + C \log \frac{A}{\epsilon} + c_0 + O(\epsilon), \tag{1.13}$$

where $A$ is the area of the boundary and $\epsilon$ is the UV cutoff — for example, the lattice spacing. See [7] for a discussion of the source of this structure and its generalisation to other dimensions.

If we're interested in the continuum theory, the regulator $\epsilon$ is an extra variable whose value is up to us. So, we're interested in only universal or regularisation-independent quantities, ie those independent of $\epsilon$. Consider taking $\epsilon \to 2\epsilon$; then we find that the form (1.13) is preserved but

$$\begin{aligned} c_2 &\to \frac{c_2}{4} \\ c_0 &\to c_0 - C \log 2. \end{aligned} \tag{1.14}$$

Thus, we see that only the coefficient of the log is universal. Other example of universal quantities are relative entropies etc, but we will only be concerned with the coefficient of the log here.

## 1.3 Gauge Theories

A gauge theory is a quantum field theory with a redundancy in the description, the most famous example being electrodynamics where $A_\mu$ and $A_\mu + \partial_\mu \lambda$ correspond to the same electric and magnetic fields and therefore denote the same state. This section is a short introduction to gauge theories with a focus on Hamiltonian lattice gauge theory. In section 1.3.1, we remind the reader of some salient facts about continuum gauge theories without matter. Then, in section 1.3.2, we introduce the Wilson path integral on the lattice, and the limit in which the Hamiltonian formulation comes out of it, and then give a general description of Hamiltonian lattice gauge theory with examples. Details for this section can be found in [8–10].

### 1.3.1 A Short Review of Continuum Gauge Theory

Let us begin with a short review of free Maxwell theory in four dimensions, since it is the most well-known example of a gauge theory. It consists of a gauge field, $A_\mu$, and the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.15}$$

This is clearly invariant under the gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \lambda, \tag{1.16}$$

since the field strength $F_{\mu\nu}$ is.

To see why we can't just call this a symmetry and have to call it a redundancy, it is useful to consider the Hilbert space on which this symmetry is represented.[2] The momenta conjugate to the $A_\mu$s, which we call $\Pi^\mu$, are

$$\Pi^i = E_i = \partial_0 A_i - \partial_i A_0, \quad \Pi^0 = 0, \tag{1.17}$$

since the Lagrangian has no time-derivatives of $A_0$. However, since the momentum operator $\Pi^0$ has to be $-i\partial_{A_0}$, $\Pi^0 = 0$ is only consistent if we demand that all allowed wavefunctionals are independent of $A_0$. Secondly, if we work out the Hamiltonian of the theory, we find that it is

$$H = \frac{1}{2} E^2 + \frac{1}{2} B^2 + A_0 \nabla \cdot E, \quad B_i = \frac{1}{2} \varepsilon_{ijk} \partial_j A_k. \tag{1.18}$$

[2] This paragraph is a dumbed-down version of the statement that the phase space has first-class constraints. For more details, see [11].

Clearly, if the wavefunction has support on a non-zero eigenspace of $\nabla \cdot E$, evolution with this Hamiltonian generates dependence on $A_0$. A more direct way of seeing this is that

$$\dot{\Pi}^0 = [\Pi^0, H] = \nabla \cdot E, \tag{1.19}$$

so that $\Pi^0 = 0$ is only conserved if $\nabla \cdot E = 0$ as well. Finally, notice that a gauge transformation by the function $\lambda(x,t)$ is implemented on the Hilbert space by the operator

$$Q[\lambda] = e^{i \int d^3x \ \dot{\lambda}\Pi^0 - \lambda \nabla \cdot E}, \tag{1.20}$$

where we have assumed that $\lambda$ dies off on the boundaries of space-time so that we can perform an integration by parts. Since we already argued that the physical Hilbert space is defined by $\Pi^0 = \nabla \cdot E = 0$, we find that the gauge transformation acts as the identity on it, showing that it is a redundancy.[3]

In practice, however, this is not a good way to quantise the theory. Since the theory is most easily formulated in terms of the gauge fields,[4] the subspace on which the gauge transformations act trivially is the subspace in which all the wavefunctions are symmetric superpositions over *gauge orbits*, the sets of all gauge-related configurations. This makes the state hard to work with for various reasons we will not go into here. The way to make things simpler is to *fix gauge*, to pick out either one state or a localised superposition of states in a gauge orbit. Gauge-fixed Hilbert spaces are unitarily equivalent to the gauge-invariant Hilbert space,[5] and so we can equivalently formulate the theory in a gauge-fixed Hilbert space.

We will work always in the temporal gauge $A_0 = 0$. The state in this gauge corresponding to the fully gauge-invariant state $\delta[\Pi^0]\psi[A_i]$ is simply $\delta[A_0]\psi[A_i]$, and the unitary nature of this map is clear. This is an example of an incomplete gauge-fixing, however, since gauge parameters independent of time cause non-trivial transformations of the $A_i$s while preserving the $A_0 = 0$ condition. These *residual gauge transformations* are generated, then, by $\nabla \cdot E$, and we can see that the $A_0 = 0$ state we wrote down is still annihilated by this Gauss' law operator. Here onwards, gauge transformations will always refer to those generated by the Gauss' law. We call the Hilbert space of all states, gauge-invariant states and not, in the $A_0 = 0$ gauge the *extended Hilbert space* $\mathcal{H}$ and the space of gauge-invariant states the *gauge-invariant Hilbert space* $\mathcal{H}_{ginv}$.

[3] A more direct way of getting this would have been to take the commutator of the generator (1.20) with the Hamiltonian (1.18); we would have seen that it was only 0 on the subspace $\nabla \cdot E = 0$. However, we would still have had to demand $\Pi^0 = 0$, making it somewhat aesthetically unappealing as a proof.

[4] This is not completely obvious in the free case, but for example interaction terms are usually of the form $J^\mu A_\mu$.

[5] Ignoring subtleties caused by incomplete gauge-fixing and Yoneya-Wadia-Gribov ambiguities and so on.

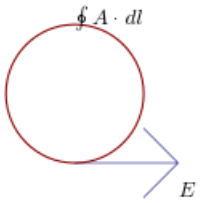


FIGURE 1.1: The non-zero commutation relations are between Wilson loops and electirc fields tangent to them.

In temporal gauge, in terms of gauge-invariant variables, the non-trivial commutation relations are between Wilson loops $\oint A \cdot dl$ and electric fields tangent to them. The reason the gauge-invariant variable is the integral around a loop is that, under gauge transformations,

$$\delta \int_x^y A \cdot dl = \lambda(y) - \lambda(x). \tag{1.21}$$

This is only 0 for arbitary $\lambda$ if $x = y$. Of course, by Ampere's law, this relarted to the amount of magnetic flux through the loop, and so this encodes the full set of non-trivial commutations.

Another important point is that the set of gauge transformations form a group. Naively, the set of transformations (1.16) in electrodynamics form the group $\mathbb{R}$, but it is also consistent with the group being $U(1)$ if the quantum fluctuations of the gauge field are much smaller than the effective radius of the $U(1)$ circle, since the only difference between these cases are periodicity conditions and in this limit the fluctuations aren't big enough to see them. In section 2.4.2, we will show that the continuum $U(1)$ gauge theory as usually formulated is in this limit. However, the compact nature of the group often turns up when there are non-contractible cycles around which the group can wind, and also in the quantisation of charge. In the real world, we see that charge is quantised and therefore the theory of electrodynamics has a $U(1)$ symmetry.

The final concept that will be useful is that of a *large gauge transformation*. If the theory lives on a spatial region with boundary, then gauge transformations where the gauge parameter $\lambda$ is non-zero on the boundary, if they are allowed by the boundary conditions, are different from usual gauge transformations in that they are real transformations of the physical variables rather than redundancies.

We can also formulate gauge theory for other continuous Lie groups. Here, the gauge fields $\mathbb{A}_\mu$ are valued in the adjoint representation and the gauge transformations are

$$\mathbb{A}_\mu \to U(x)^{-1}\mathbb{A}(x)U(x) + U(x)^{-1}\partial_\mu U. \tag{1.22}$$

The field strength, ignoring some relative factors, is

$$\mathbb{F}_{\mu\nu} = \partial_\mu \mathbb{A}_\nu - \partial_\nu \mathbb{A}_\mu + [\mathbb{A}_\mu, \mathbb{A}_\nu], \tag{1.23}$$

and the Yang-Mills Lagrangian that generalises the Maxwell Lagrangian is

$$L = -\frac{1}{4} \operatorname{tr} \mathbb{F}_{\mu\nu} \mathbb{F}^{\mu\nu}. \tag{1.24}$$

We will always consider the case of a compact gauge group. We also note that the analog of a Wilson loop in a non-Abelian gauge theory is

$$\operatorname{tr} P e^{i \oint \mathbb{A} \cdot dl}, \tag{1.25}$$

where $P$ is path-ordering.

### 1.3.2 An Introduction to Hamiltonian Lattice Gauge Theory

A lattice gauge theory is a lattice quantum system locally invariant under the actions of a gauge group $G$, which we will take to be compact, The gauge field becomes a $G$-valued variable living on the links of the lattice, and the value of the link variable is analogous to the Wilson line on the lattice $U_{ij} \sim e^{i \int_i^j A}$; we denote the identity element by 1, clarifying that it is a grou element and not a number when there is possibility of confusion. In this section, we will introduce the Hamiltonian formulation of lattice gauge theory [9, 10] in the temporal gauge $U_{i,i+\hat{t}} = 1$, where 1 denotes the identity element of $G$. We will first introduce the formulation for a general compact group to emphasise the general structure, and then illustrate it in the cases of the groups $\mathbb{Z}_N, U(1), SU(2)$.

We approach the Hamiltonian formulation from the Euclidean path integral. We consider a square lattice in $d+1$ dimensions. We label vertices of the lattice by small letters $i$, $j$, etc. and links by the endpoints as $(ij)$, with $(ij)$ and $(ji)$ denoting the same link. The set of vertices is $V$, the set of links $L$ and the set of plaquettes, the elementary square of the lattice, $P$. The variables are elements of the group on every link, and we use the Haar measure, defined by

$$dU = d(VU) = d(UV) \quad \text{and} \quad \int_G dU = 1. \tag{1.26}$$

to integrate over them. In the case of a discrete group, the Haar measure is just a sum over all the group elements, with normalisation given by (1.26). The Wilson action, which for a continuous Lie group gives the Yang-Mills action (1.24) in the continuum, is

$$S_W = \sum_{p \in P} \operatorname{tr} (U_{ij} U_{jk} U_{kl} U_{li}) + cc, \quad ijkl \in p. \tag{1.27}$$

This action has a gauge-invariance given by a group element $V_i$ at every vertex,

$$U_{ij} \to V_i U_{ij} V_j^{-1}. \tag{1.28}$$

The path integral

$$\int DU e^{-S_W}, \quad DU = \prod_i dU_i, \tag{1.29}$$

automatically sums over gauge orbits, and the normalisation of the Haar measure means that this sum doesn't diverge in an uncontrollable manner; thus, it is well-defined without gauge-fixing, a property that will not continue to be true in the continuum, see section 2.4.2 for details.

The Hamiltonian formulation that we will use is given by fixing temporal gauge

$$U_{i,i+\hat{t}} = 1, \tag{1.30}$$

and taking the limit in which time is continuous, see [8] for details. In more general gauges, we have to include ghosts and negative norm states in the Hamiltonian formulation, and they have not yet been satisfactorily incorporated into the entanglement story. Without describing in detail how it comes about, we will now describe the Hilbert space.

The Hilbert space $\mathcal{H}_{ij}$ on a link $(ij)$ is spanned by a set of basis vectors $\{|U_{ij}\rangle\}$, one for every element of the group $G$,

$$\mathcal{H}_{ij} = \mathcal{H}_{ji} = \text{span}\left(\left\{ |U\rangle_{ij} \middle| U \in G \right\}\right) \tag{1.31}$$

with the identification

$$|U\rangle_{ij} = \left|U^{-1}\right\rangle_{ji}; \tag{1.32}$$

this is the analogue of the fact that $A \cdot \hat{x} = -A \cdot (-\hat{x})$. We define the (matrix-valued) 'field' operator $\hat{U}^r_{ij}$ by

$$\hat{U}^r_{ij} |U\rangle_{ij} = D^r(U) |U\rangle_{ij} \tag{1.33}$$

with the convention that $\hat{U}_{ji} = \hat{U}_{ij}^{-1}$ to agree with (1.32). Here, $D^r$ is the Wigner matrix or representation matrix of the irrep $r$, and the operator $U^r_{ij}$ without a superscript means that $r$ is the fundamental irrep.[6]

There are two electric operators that don't commute with $\hat{U}_{ij}$ for every element of $G$,

$$\begin{aligned} \hat{L}_{ij}(V) |U\rangle_{ij} &= |VU\rangle_{ij} \\ \hat{L}_{ji}(V) |U\rangle_{ij} &= \left|UV^{-1}\right\rangle_{ij}. \end{aligned} \tag{1.34}$$

[6]The representation matrices of the fundamental irrep themselves transform in the adjoint irrep.

From the definition, it is clear that

$$[\hat{L}_{ij}, \hat{L}_{ji}] = 0. \tag{1.35}$$

There is a theorem, the Peter-Weyl theorem, see [12, 13] for an introduction, that is the non-Abelian version of a Fourier transform, according to which we can write

$$|U\rangle_{ij} = \sum_{r,\alpha,\beta} \sqrt{d_r} D^r_{\alpha\beta}(U) \, |r, \alpha, \beta\rangle_{ij} \, , \tag{1.36}$$

where $r$ denotes the irrep, $d_r$ its dimension, and $D^r_{\alpha\beta}$ is the $\alpha, \beta$ element of the representation matrix of the irrep $r$ and the 'electric basis' states $|r, \alpha, \beta\rangle$ have the property that

$$\begin{aligned} \hat{L}_{ij}(V) \, |r, \alpha, \beta\rangle_{ij} &= |r, \gamma, \beta\rangle_{ij} \, D^r_{\gamma\alpha}(V), \\ \hat{L}_{ji}(V) \, |r, \alpha, \beta\rangle_{ij} &= D^r_{\beta\delta}(V) \, |r, \alpha, \delta\rangle_{ij} \, . \end{aligned} \tag{1.37}$$

We can convert between the two bases using the completeness relations

$$\begin{aligned} \int dU \, D^{r'}_{\alpha'\beta'}(U)^* D^r_{\alpha\beta}(U) = \int dU \, D^{r'}_{\beta'\alpha'}(U^{-1}) D^r_{\alpha\beta}(U) = \frac{1}{d_r} \delta_{r,r'} \delta_{\alpha,\alpha'} \delta_{\beta,\beta'}, \\ \sum_r d_r \sum_{\alpha,\beta} D^r_{\alpha\beta}(U)^* D^r_{\alpha\beta}(V) = \delta(U^{-1}V). \end{aligned} \tag{1.38}$$

Naively, the Hilbert space of the entire lattice should be the tensor product of those on the different spatial links,

$$\mathcal{H} = \bigotimes_{(ij)\in \text{links}} \mathcal{H}_{ij}. \tag{1.39}$$

Gauge transformations are implemented at every vertex by the operator

$$\hat{G}_i(U) = \prod_{j|(ij)\in L} \hat{L}_{ij}(U). \tag{1.40}$$

Since gauge-invariance should be thought of not as invariance under a physical symmetry but a redundancy, the real Hilbert space is the *gauge-invariant Hilbert space* $\mathcal{H}_{ginv}$ defined as the subspace invariant under these transformations,

$$|\psi\rangle \in \mathcal{H}_{ginv} \quad \Leftrightarrow \quad \forall i, U, \; \hat{G}_i(U) \, |\psi\rangle = |\psi\rangle \, . \tag{1.41}$$

We call the bigger Hilbert space defined in (1.39) the *extended Hilbert space*; it is the natural Hilbert space in theories where the gauge symmetry is only a low-energy property.

The gauge-invariance constraint actually means that the irreps $r_{ij}$ fuse to the identity irrep, since the product of representation matrices must act as the identity for every group element

according to (1.41). So, the gauge-invariant Hilbert space is spanned by a set of states called *spin networks*, labelled by an irrep $r_{ij}$ on every link, and an *intertwiner* $\iota_i$ at every vertex labelling how the irreps on all the links fuse to the identity. The intertwiner is necessary in non-Abelian groups when there are multiple ways the irreps can fuse to the identity; for example, labelling irreps of $SU(2)$ by the value of $j$,

$$\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 0 \oplus 1 \oplus 2 \oplus 1 \oplus 1, \tag{1.42}$$

One 0 comes from the doublets pairing up and forming singlets, whereas the other 0 comes from two pairs of doublets both forming triplets and the two triplets fusing to a singlet. The intertwiner labels which of these fusions actually happens at the vertex where these four links meet.

### $\mathbb{Z}_N$ Theory

$\mathbb{Z}_N$ is a discrete Abelian group consisting of $N$ elements $0, 1 \cdots N-1$, with the group multiplication being addition modulo $N$. Its irreducible representations are labelled by an integer $k \in \{0, 1 \cdots N-1\}$, and the Wigner matrices — which are complex numbers, since this is an Abelian group — are

$$D^k(n) = e^{i\frac{2\pi}{N}kn}. \tag{1.43}$$

The basis change between electric and magnetic bases is just the discrete Fourier transform

$$|k\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{n-1} e^{i\frac{2\pi}{N}kn} |n\rangle . \tag{1.44}$$

This is why we stated that the Peter-Weyl theorem was the non-Abelian analogue of the Fourier transform.

Since the group is Abelian, the electric basis is labelled just by the $k_{ij}$s and a basis for the gauge-invariant Hilbert space is given by the subset of these configurations that satisfy Gauss' law,

$$\left\{ |k_{ij}\rangle \,\middle|\, \forall i, \sum_{j,(ij)\in L} k_{ij} \mod N = 0 \right\}. \tag{1.45}$$

### $U(1)$ Theory

$U(1) = \lim_{N\to\infty} \mathbb{Z}_N$ is the group whose defining representation is the unit circle in the complex plane,

$$U(1) \cong \left\{ e^{i\theta} \,\middle|\, \theta \in [0, 2\pi) \right\}. \tag{1.46}$$

Its representations are labelled by integers, which we again call $k$, with Wigner matrices

$$D^k(\theta) = e^{ik\theta}. \tag{1.47}$$

The $k^{th}$ irrep corresponds to the circle winding around itself $k$ times.

The basis change is again a Fourier transform

$$|k\rangle = \int \frac{d\theta}{2\pi} e^{ik\theta} |\theta\rangle , \tag{1.48}$$

and the gauge-invariant Hilbert space is again spanned by a set of the form (1.45) except with $N$ set to $\infty$.

The new element from the $\mathbb{Z}_N$ case is that we can define an electric field operator $E_{ij}$ by

$$L_{ij}(\theta) = e^{i\theta E_{ij}}. \tag{1.49}$$

The eigenvalues of the $E_{ij}$s are actually the $k_{ij}$s defined previously, and the gauge-invariance constraint becomes equivalent to the lattice Gauss' law

$$\sum_{j,(ij)\in L} E_{ij} |\psi\rangle = 0. \tag{1.50}$$

### $SU(2)$ Theory

$SU(2)$ is the group of $2 \times 2$ unitary matrices with determinant 1, familiar to physicists from the study of the angular momentum algebra. It is a three-dimensional Lie group, whose generators we denote by $T^a$, $a = x, y, z$. This means that every element can be written as $e^{i\epsilon_a T^a}$. Further, the $T^a$s span a vector space, known as the Lie algebra, and have standard angular momentum commutation relations

$$[T^a, T^b] = i\varepsilon_{abc} T^c. \tag{1.51}$$

The irreps are labelled by a number $j \in \left\{0, \frac{1}{2}, 1 \cdots\right\}$, with dimension

$$d_j = 2j + 1. \tag{1.52}$$

The irrep is completely specified the quadratic Casimir

$$\sum_a D^j(T^a) D^j(T^a) = j(j+1)\mathbb{1}_{2j+1} \tag{1.53}$$

This is, of course, just a fancy way of writing the usual $J^2$ operator that measures the total angular momentum.

As in the case of $U(1)$, we prefer electric field operators in the Lie algebra rather than Lie group; meaning that we replace $L_{ij}$ by $J_{ij}$, defined as

$$\begin{aligned} e^{i\epsilon_a \hat{J}^a_{ij}} |U\rangle_{ij} &= \left|e^{i\epsilon_a T^a} U\right\rangle_{ij} \\ e^{i\epsilon_a \hat{J}^a_{ji}} |U\rangle_{ij} &= \left|U e^{-i\epsilon_a T^a}\right\rangle_{ij} . \end{aligned} \tag{1.54}$$

From this definition, it is apparent that $[\hat{J}_{ij}, \hat{J}_{ji}] = 0$. Since a link is specified by only one irrep,

$$\hat{J}^2_{ij} = \hat{J}^2_{ji}. \tag{1.55}$$

Using this, we define an 'electric basis' $\{|j_{ij}, m_{ij}.m_{ji}\rangle\}$, with the property that

$$\begin{aligned} \hat{J}^2_{ij} |j_{ij}, m_{ij}, m_{ji}\rangle = \hat{J}^2_{ji} |j_{ij}, m_{ij}, m_{ji}\rangle &= j_{ij}(j_{ij}+1) |j_{ij}, m_{ij}, m_{ji}\rangle , \\ \hat{J}^z_{ij} |j_{ij}, m_{ij}, m_{ji}\rangle &= m_{ij} |j_{ij}, m_{ij}, m_{ji}\rangle \quad \text{and} \\ \hat{J}^z_{ji} |j_{ij}, m_{ij}, m_{ji}\rangle &= m_{ji} |j_{ij}, m_{ij}, m_{ji}\rangle . \end{aligned} \tag{1.56}$$

The transformation between the two bases is

$$|j_{ij}, m_{ij}, m_{ji}\rangle = \sqrt{2j_{ij}+1} \int dU D^{j_{ij}}_{-m_{ij}, m_{ji}}(U) |U\rangle . \tag{1.57}$$

Gauge transformations are implemented by operators at every vertex

$$\hat{G}^a_i(\theta_a) = e^{i\theta_a \mathcal{G}^a_i} \quad \Big/ \mathcal{G}^a_i = \sum_{j,(ij)\in L} \hat{J}^a_{ij}. \tag{1.58}$$

We note for later reference that (1.41) also means that

$$\mathcal{G}^a_i \big|_{\mathcal{H}_{ginv}} = 0. \tag{1.59}$$

# Chapter 2

# Definition

Now that we have discussed the basics of entanglement and the quantisation of gauge theories, we are ready to define entanglement in gauge theories. The reason there is a problem is that there are no truly local excitations in a gauge theory; the physical excitations are loops of magnetic and electric flux.

More mathematically, the problem is that the gauge-invariant Hilbert space doesn't factorise into inside and outside Hilbert spaces because of the Gauss' law constraint, $\nabla \cdot E \left|\psi\right\rangle = 0$ (or its discrete group analogues). Since the Gauss' law operator has a derivative, it involves operators both inside and outside the region. On the lattice, this can be made precise, since $\nabla \cdot E \sim \sum_{in} E + \sum_{out} E$, which is clearly a sum of inside and outside operators. We will not try to formulate it precisely in the continuum in this thesis.

Since $\mathcal{H}_{ginv}$ can't be written as $\mathcal{H}_{in} \otimes \mathcal{H}_{out}$, subregions are not subsystems and therefore there is no simple notion of the reduced density matrix of a subregion. One could try and change the question to being about subsystems, but not only can defining a correct subsystem be complicated — especially in non-Abelian theories — but also, in field theory, subregions are more natural and interesting objects to thing about.

In this chapter, we will stick to the lattice, define a region as a set of links as in figure (2.1), and discuss how to define entanglement despite the Gauss' law constraint. It is only in the end that, using a path integral, we will take the continuum limit.

The naive solution to the non-factorisation issue would be to gauge-fix, but as we show using both a specific example and general arguments in section 2.1, this procedure generically rearranges the degrees of freedom in a gauge-dependent way and therefore the answer is gauge-dependent. More sophisticated solutions involve sticking with the gauge-invariant

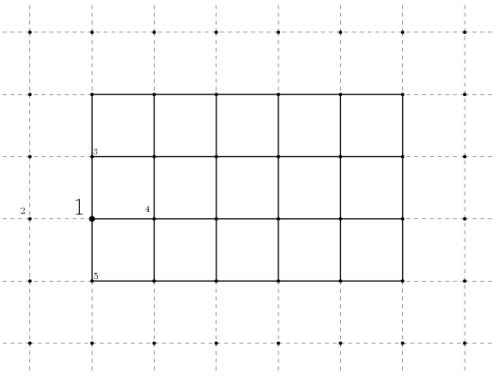


FIGURE 2.1: We define our region as a set of links. Here, the solid lines denote the inside region, and the dotted lines denote the outside region.

Hilbert space, and fall into two classes. The first, based on a Hilbert-space-oriented perspective, is called the *extended Hilbert space* definition [1, 14–16] and defined in section 2.2. The second, based on an operator-algebraic perspective [17–19], is defined in section 2.3.

We end this chapter by describing the relation of the extended Hilbert space definition to a lattice replica trick [1] and taking the continuum limit of this path integral to show that it (almost) gives the continuum replica trick path integral [3] in section 2.4.

## 2.1 The Follies of Gauge-Fixing

In this section, we show that gauge-fixing beyond temporal gauge makes the entanglement depend on gauge. We will stick to the gauge group $\mathbb{Z}_2$ for simplicity and definiteness. First, in a lattice consisting of only one plaquette (and therefore only one qubit worth of degrees of freedom), we will show that different gauges have different entanglement entropies. Then, we give some general arguments that this is generic.

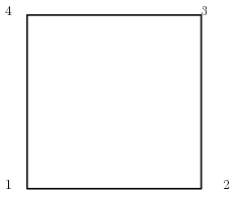


FIGURE 2.2: An example lattice for understanding the problems with gauge-fixing for calculating the EE. We calculate the EE between link (12) and the rest.

The lattice we consider for the example is shown in figure (2.2). We consider a state

$$|\psi\rangle = \cos\theta\,|0\rangle + e^{i\phi}\sin\theta\,|1\rangle\,, \quad \Big/ \prod \sigma^z\,|i\rangle = (-1)^i\,|i\rangle \tag{2.1}$$

and consider entanglement between the link (12) and the rest.

We can take the gauge

$$\sigma^z_{14} = \sigma^z_{23} = \sigma^z_{34} = 1. \tag{2.2}$$

Clearly, in this gauge, the entire information of the state is localised on the link (12) and therefore there is no entanglement.

On the other hand, considering the gauge

$$\sigma^z_{12} = \sigma^z_{23} = \sigma^z_{34} = \sigma^z_{41}, \tag{2.3}$$

the state can be written as

$$|\psi\rangle = \cos\theta\,|0000\rangle + e^{i\phi}\sin\theta\,|1111\rangle\,, \tag{2.4}$$

which clearly has a non-zero entanglement.

The basic reason for this is that the Gauss' law constraint on the boundary points (here 1 and 2) involves both inside and outside links, and so there are degrees of freedom that are neither/both inside nor/and outside, and gauge-fixing involves making a choice to put it somewhere. Consider, for example, fixing Coulomb gauge $\nabla \cdot A = 0$. Once this gauge is fixed, the gauge potential $A_i$ is a *different*, gauge-invariant operator than the $A_i$ that we have spoken about till now,

$$A_i^{Coulomb} = \left(\delta_{ij} - \partial_i \frac{1}{\nabla^2}\partial_j\right) A_j^{EHS}. \tag{2.5}$$

Clearly, in fixing gauge, we have rearranged degrees of freedom, in the sense that operators that restricted to one region in one gauge need not be so in another. Further, fixing a different gauge would correspond to a different rearrangement of the degrees of freedom, and this is why if we naively gauge-fix the state and try to calculate the EE the answer will be gauge-dependent.

This also gives us a perspective on why we were allowed to fix temporal gauge. The fully gauge-invariant state has the form $\otimes_i \left|\Pi^0 = 0\right\rangle \otimes |\psi\rangle$, and therefore the $A_0$ sector doesn't have any entanglement. We can project the $A_0$ sector to for example $\frac{1}{\sqrt{2}}\left(\otimes_i |A_0 = 0\rangle + \otimes_i |A_0 = 1\rangle\right)$ and get an extra contribution to the entanglement. Projecting onto $A_0 = 0$ at every vertex, however, clearly leaves the $A_0$ sector unentangled, and so it doesn't change the total entanglement.

## 2.2 Extended Hilbert Space Definition

The first class of definitions, which is known as the extended Hilbert space (EHS) definition [1, 14–16], follows directly from the observation that while $\mathcal{H}_{ginv}$ doesn't have a tensor product structure, the EHS $\mathcal{H}$ by definition (1.39) does. Since $\mathcal{H}_{ginv} \subset \mathcal{H}$,

$$|\psi\rangle \in \mathcal{H}_{ginv} \subset \mathcal{H} = \mathcal{H}_{in} \otimes \mathcal{H}_{out} \tag{2.6}$$

and use this tensor product factorisation to define the reduced density matrix and therefore entanglement.

As an example, we can consider the state (2.1) from section 2.1. The basis states, in the basis of the individual $\sigma^z$s, are

$$\begin{aligned} |0\rangle &= \frac{|0000\rangle + (|0011\rangle + \text{permutations}) + |1111\rangle}{\sqrt{8}} \\ |1\rangle &= \frac{(|0001\rangle + \text{permutations}) + (|0111\rangle + \text{permutations})}{\sqrt{8}}. \end{aligned} \tag{2.7}$$

The reason this is so complicated is that these are symmetric superpositions over gauge orbits, and the gauge orbit in this system has eight states. To simplify the calculation, we go the electric basis, in which the $\sigma^x$s are diagonal instead. The usefulness of this basis is that the Gauss' law operators are diagonal in this basis, and so Gauss' law becomes a constraint on allowed basis elements instead of something that picks out particular symmetric combinations of them, as show explicitly in (1.45) for example. In the electric basis, the basis states above become

$$|i\rangle = \frac{|0000\rangle + (-1)^i |1111\rangle}{\sqrt{2}}. \tag{2.8}$$

Gauss' law sets all the $\sigma^x$s equal to each other, and so there are only the two basis elements we see on the right. In this basis, the state is

$$|\psi\rangle = \frac{\cos\theta + e^{i\phi}\sin\theta}{\sqrt{2}} |0000\rangle + \frac{\cos\theta - e^{i\phi}\sin\theta}{\sqrt{2}} |1111\rangle, \tag{2.9}$$

giving for the density matrix of the link (12)

$$\rho = \frac{(1 + \sin 2\theta \cos\phi) |0\rangle\langle 0| + (1 - \sin 2\theta \cos\phi) |1\rangle\langle 1|}{2}. \tag{2.10}$$

Note, in particular, that this reduced density matrix has different eigenvalues from either on in section 2.1.

One might wonder, given that we ignored the Gauss' law constraint to define the reduced

density matrix, whether gauge-invariance is completely irrelevant and therefore there's nothing more to be said. But, one must remember that even though we embedded the state into a bigger Hilbert space, it was still a gauge-invariant state and therefore the state itself has some structure that will manifest in the density matrix. To understand this structure, we will first consider Abelian gauge groups and then deal with the complications that arise from non-Abelian gauge groups.

We are particularly interested in vertices which are connected to both inside and outside links, since it is those Gauss' law constraints which will cause the interesting structure to turn up. Taking $i$ to be a boundary vertex, we therefore define the quantities

$$\begin{aligned} G_i^{in} &= \prod_{j\in in} L_{ij} \sim e^{-iE\cdot\hat{n}}, \\ G_i^{out} &= \prod_{j\in out} L_{ij} \sim e^{iE\cdot\hat{n}}, \end{aligned} \tag{2.11}$$

where $\hat{n}$ is the outward-pointing normal and we have written the analogous operator in continuum $U(1)$ gauge theory to build intuition. Since these operators would generate large gauge transformations if the theory was truly restricted to the inside/outside, we abuse nomenclature and call these operators also the generators of large gauge transformations. Since, because of Gauss' law,

$$G_i^{in} = \left(G_i^{out}\right)^{-1} \tag{2.12}$$

at every boundary vertex $i$, a gauge-invariant state has the schematic form

$$\begin{aligned} |\psi\rangle = \sum_{\mathbf{k}} \sqrt{p_{\mathbf{k}}}\,|\psi_{\mathbf{k}}\rangle = \sum_{\mathbf{k}} \sqrt{p_{\mathbf{k}}}\psi_{\mathbf{k}}(\text{bulk variables})\,|\text{inside variables}, \mathbf{k}\rangle_{in}\,|\mathbf{k}, \text{outside variables}\rangle_{out}\,, \\ \Big/\ \langle\psi_{\mathbf{k}}|\psi_{\mathbf{k}}\rangle = 1, \end{aligned} \tag{2.13}$$

where $\mathbf{k}$ is a collective variable for the eigenvalues of all the $G_i^{in}$s and $\mathbf{k}$ in the outside ket means that $G_i^{out}$s have the corresponding value.

Because $\mathbf{k}$ is both an inside and an outside variable, so that

$$\mathrm{tr}_{out}\,|\psi_{\mathbf{k}}\rangle\,\langle\psi_{\mathbf{k}'}| \propto \left\langle \mathbf{k}\middle|\mathbf{k}'\right\rangle_{out} = \delta_{\mathbf{k},\mathbf{k}'}, \tag{2.14}$$

tracing over the outside region with the state (2.13) necessarily gives us

$$\rho_{in} = \bigoplus_{\mathbf{k}} p_{\mathbf{k}}\rho_{in}^{\mathbf{k}}, \quad \rho_{in}^{\mathbf{k}} = \mathrm{tr}_{out}\,|\psi_{\mathbf{k}}\rangle\,\langle\psi_{\mathbf{k}}|\,, \tag{2.15}$$

Because the reduced density matrix is diagonal in a subset of all the variables, it is *block-diagonal*.

There is a more abstract derivation of this block-diagonal structure, which highlights its connection with the gauge-invariance structure of the reduced density matrix. Taking $i$ to be a boundary vertex and $U$ to be a particular group element,

$$\begin{aligned}\rho_{in} &= \mathrm{tr}_{out}\left(|\psi\rangle\langle\psi|\right) \\ &= \mathrm{tr}_{out}\left(G_i(U)\,|\psi\rangle\langle\psi|\,G_i^\dagger(U)\right) \\ &= G_i^{in}(U)\rho_{in}(G_i^{in}(U))^\dagger,\end{aligned} \tag{2.16}$$

where in the last step we have written $G_i = G_i^{in}G_i^{out}$ and used cyclicity of trace to cancel the $G_i^{out}$s. This manipulation implies that the reduced density matrix commutes with the $G_i^{in}$s and therefore must be block-diagonal in that basis. Further, it shows the important difference between an inside gauge transformation and a boundary gauge transformation, since for $i \in in$, $G_i\rho_{in} = \rho_{in}$ whereas for the boundary operator we have (2.16).

In the non-Abelian case, the $L_{ij}$s on a link don't commute with each other, and therefore can't be simultaneously diagonalised. In particular, the $G_i^{in}$s are also not gauge-invariant and can't be simultaneously diagonalised. However, since the $G_i^{in}$s form a representation, we can go to a basis in which each basis element is in a particular irrep at every boundary point. Then, the analogue of $\mathbf{k}$ is a list of all the irreps at all the points; it can, for example, be the eigenvalues of all the Casimirs. Since at every point the representation of $G_i^{in}$ and $G_i^{out}$ has to fuse to a singlet, the two representations must be conjugate representations in each branch of the wavefunction, giving a decomposition of the form (2.13) except here $\mathbf{k}$ in the outside ket means the conjugate representation.

There is, however, an extra piece of structure: not only do the two representations have to be the same, but they also have to fuse to a singlet. Two irreps that fuse to a singlet are necessarily maximally entangled, and so the reduced density matrix is of the form

$$\rho_{in} = \oplus_{\mathbf{k}} p_{\mathbf{k}}\rho_{\mathbf{k}} \bigotimes \otimes_i \frac{\mathbb{1}_{d_{k_i}}}{d_{k_i}}, \tag{2.17}$$

where $d_{k_i}$ is the dimension of the irrep $k_i$.

In the example $G = SU(2)$, $k_i = J_{i,in}^2$ is the quadratic Casimir, every representation is conjugate to itself, and the singlet state in the representation $k_i = j_i(j_i+1)$ is

$$\tfrac{1}{\sqrt{2j+1}} \sum_{m_i=-j_i}^{j_i} \left|J_{i,in}^z = m_i, J_{i,out}^z = -m_i\right\rangle. \tag{2.18}$$

From here, the form (2.17) follows.

Taking the density matrix (2.17), we see that the EE has the decomposition

$$S_E = -\sum p_{\mathbf{k}} \log p_{\mathbf{k}} - \sum p_{\mathbf{k}} \rho_{\mathbf{k}} \log \rho_{\mathbf{k}} + \sum p_{\mathbf{k}} \log d_{\mathbf{k}} \equiv S_c + S_q + s_f. \tag{2.19}$$

The third term is automatically 0 for Abelian gauge theories for the simple reason that all the irreps are one-dimensional. This decomposition can be thought of as

1. The classical term $S_c$: A classical Shannon entropy for the probability distribution of the edge modes.
2. The quantum term $S_q$: A quantum piece which is the average of the EE in all the different edge mode sectors.
3. The fusion term $S_f$: A contribution that turns up due to the requirement of fusion into singlets.

This decomposition, apart from the definition, is the main point of this section.

To end this section, we note some important properties of this definition

1. It is unambiguous, because there is a canonical embedding of $\mathcal{H}_{ginv}$ in the EHS $\mathcal{H}$.
2. The reduced density matrix gives the correct expectation values for all the gauge-invariant operators with support on only the inside links, because that is literally how it's defined, (1.4).
3. Because $\mathcal{H}$ has a tensor product decomposition, it satisfies standard properties of EE like strong subadditivity. For three non-intersecting regions $A, B, C$, the strong subadditivity inequality is

   $$S_{ABC} + S_B \leq S_{AB} + S_{BC}. \tag{2.20}$$

   We mention it mainly because it is a very important property of EE.
4. It is the natural definition in cases where the gauge symmetry is only a symmetry of the low-energy subspace, since for those theories the EHS $\mathcal{H}$ is the real Hilbert space.
5. It generalises simply to gauge theories coupled to matter.
6. It naturally admits a lattice replica trick calculation, as we will show in section 2.4.

## 2.3 Electric Centre Definition

There is another point-of-view on defining EE [17–19], which begins with the observation that it measures correlations between inside and outside operators, and so we must investigate not Hilbert spaces but algebras of operators. We will see that the failure of factorisation manifests here in the existence of a non-trivial *centre* in the subalgebra corresponding to a region, and then that a particular choice of subalgebra, called the *electric centre* choice, gives an EE equal to the first two terms of (2.19) — that is, it is equivalent to the EHS definition in the Abelian case but not in the non-Abelian case.

An algebra $\mathcal{A}$ is a set of operators that are closed under multiplication and addition, so that if $O_1$ and $O_2$ are elements of the algebra so are $O_1O_2$ and $O_1 + O_2$. A subalgebra $\mathcal{B} \subset \mathcal{A}$ is a subset that is also an algebra. A centre $\mathcal{Z}$ of $\mathcal{A}$ is a subalgebra that commutes with every element of $\mathcal{A}$.

For us, the elements of the algebras will always be operators. A state, abstractly, is a function that takes every operator to its expectation value. Since the algebra is closed under addition, it forms a vector space and we can always write the state as a sum of basis elements with coefficients given by their expectation values, so that every state translates uniquely to an element $\rho_{\mathcal{A}} \in \mathcal{A}$. In other words, we can think of a state as an element of the algebra that gives all the correct expectation values.

In factorisable systems, where $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, we can associate subalgebras to the factors,

$$\mathcal{A}_1 = \{O_1 \otimes \mathbb{1}_2\} \quad \text{and} \quad \mathcal{A}_2 = \{\mathbb{1}_1 \otimes O_2\}. \tag{2.21}$$

We note for future reference that the centres of both the algebras are the same, and given by multiples of the identity; we call this case the *trivial centre*. The reduced density matrix $\rho_1$ is clearly an operator in $\mathcal{A}_1$ that gives all the correct expectation values, and therefore is the density matrix reduced to this algebra.

Note, however, that both $\rho_1$ and $\rho_1 \otimes \mathbb{1}_2$ are the correct density matrix! This really comes from the fact that $\mathcal{A}_1$ has both the representation in (2.21) as well as another representation with the $\mathbb{1}_2$s stripped off. This is an example of the general fact that algebras don't have unique representations as operators, and there may in general be *representation ambiguities*. We will always take the representation to be the smallest faithful representation, and we will see that the EHS density matrix in the non-Abelian case differs from the electric centre one due solely to such a representation ambiguity.

Now that we have covered the basics of operator algebras, we can consider the gauge theory case. Consider a set of links and the maximal subalgebra we can associate to that set and

its complement; we will call these $\mathcal{A}$ and $\mathcal{A}'$. We only consider these operators restricted to $\mathcal{H}_{ginv}$, so that two operators that differ only on non-gauge-invariant states are the same element of the algebra; this means that we don't need the crutch of the EHS for these considerations.

The basic point is that the subalgebra $\mathcal{A}$ has a non-trivial centre. A simple proof of this follows from (2.12), $G_i^{in} = \left(G_i^{out}\right)^{-1}$. Since $G_i^{out}$ is supported entirely on the outside links and two operators on different links necessarily commute, this means that $G_i^{in}$ necessarily commutes with every inside operator. Thus, the set of these operators (or the corresponding Casimirs in the non-Abelian case) is in the centre of $\mathcal{A}$. Since this centre consists only of electric operators, it is called the electric centre algebra.

More prosaically, while $G_i^{in}$ has non-trivial commutator with $U_{ij}, j \in in$, this operator is not gauge-invariant. Gauge-invariant operators are necessarily of the form of Wilson loops, which look like

$$\mathrm{tr}(\cdots U_{j_1 i} U_{i j_2} \cdots). \tag{2.22}$$

Conjugation with $G_i^{in}(V)$ converts this operator, by definition, to

$$G_i^{in}(V)\,\mathrm{tr}(\cdots U_{j_1 i} U_{i j_2} \cdots)\left(G_i^{in}(V)\right)^\dagger = \mathrm{tr}\Big(\cdots U_{j_1 i} V^\dagger V U_{i j_2}\Big), \tag{2.23}$$

and thus it is invariant. Thus, we see again that the $G_i^{in}$s (or the corresponding Casimirs) form the centre.

Since every operator in $\mathcal{A}$ commutes with the centre and $\rho_{\mathcal{A}} \in \mathcal{A}$, $\rho_{\mathcal{A}}$ must also commute with the centre. This means that it must be block-diagonal in the eigenbasis of the centre variables, giving again the form (2.15) when restricting to the minimal representation of the algebra. This means that the block-diagonal form (2.15) is in the subalgebra of the inside links; by uniqueness, it must be the density matrix reduced to the algebra.

Further, we see that every operator in the inside algebra must also have the same block-diagonal structure. Since, because of this, no operator can change the value of $\mathbf{k}$, each sector of definite $\mathbf{k}$ is called a *superselection sector*.[1]

In the non-Abelian case, including the fusion spaces leads to a non-minimal representation, since all gauge-invariant operators act as multiples of the identity on them; this is clearly true of operators on links not involved in that fusion space, and also true of the Casimirs of the $G_i^{in}$ representation because these are irreps. Thus, if we restrict to a minimal representation, we don't find the extra piece of the density matrix (2.17) and even in the non-Abelian case the density matrix looks like (2.15).

[1] A selection sector is a subspace that the Hamiltonian doesn't evolve the state out of, like a sector of fixed angular momentum; superselection sectors are more strongly protected than selection sectors, hence the name.

Thus, we see that in the electric centre definition the EE has the form (2.19), except without the last term. While this may seem like semantics, since we purposefully restricted to the minimal representation, restricting to the minimal representation has the direct physical representation of not including entropy of degrees of freedom on which one can't act in any non-trivial way; in particular, there are no operators on the inside and outside whose correlation gives you the fusion piece of the EE.

Finally, we note that electric centre isn't the only choice of subalgebra; one can remove some boundary operators and end up with subalgebras with different centres. It is even possible to make the centre completely trivial or completely magnetic. We will here take the view that we prefer to work with the maximal subalgebra and therefore not consider any of these. For more details on these different choices and their meaning, see [17].

## 2.4 Replica Trick and Continuum Limit

EE in field theory is usually calculated not by Hamiltonian methods but by using the replica trick path integral, see [7] for details. This is a Euclidean path integral on a manifold with a conical singularity at the boundary of the region, such that we have to go around an angle $2\pi n$ instead of $2\pi$ around the boundary to come back to the same place. We show that the lattice EHS definition is calculated by a lattice version of the replica trick, and then take the continuum limit (when it exists) to show that it gives the same answer as the continuum replica trick up to uninteresting terms.

In section 2.4.1, following [1], we describe the basic idea of the replica trick and construct the appropriate replica trick for the EHS definition. Then, in section 2.4.2, followng [3], we take the continuum limit in the case of four-dimensional $U(1)$ gauge theory, a case in which we know the continuum limit exists; we expect the details to work out similarly in every case where the continuum limit exists, but have only worked it out in this case.

### 2.4.1 Lattice Replica Trick for the EHS Definition

The replica trick path integral is constructed to carry out the calculation for $\operatorname{tr} \rho_{in}^n$. From here, we analytically continue to non-integer $n$ and use the fact that

$$S_E = -\operatorname{tr} \rho_{in} \log \rho_{in} = -\lim_{n\to 1} \partial_n \operatorname{tr} \rho_{in}^n \tag{2.24}$$

to calculate the EE. This analytic continuation need not work, but it is customary to ignore this fact [20]. A pictorial representation of the construction of the replica trick is given in the figures.

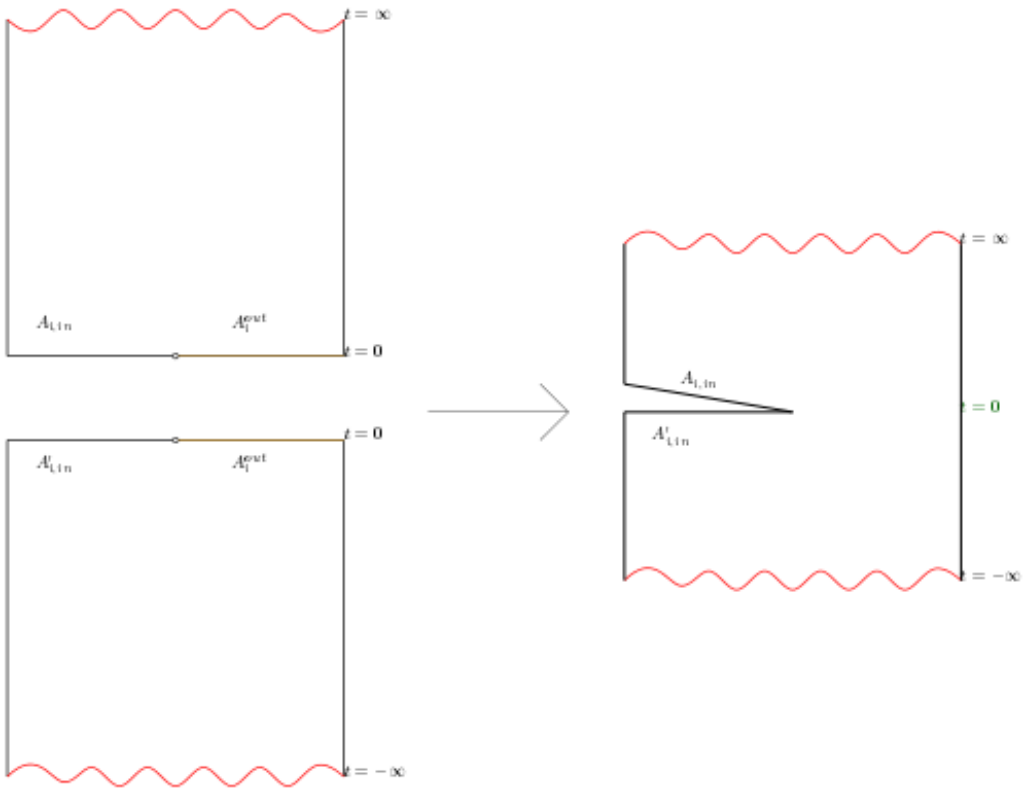


FIGURE 2.3: The replica trick path integral for the reduced density matrix is constructed by taking two half-space-time path integrals, one for the ket and one for the bra, and sewing them up along the outside. $\operatorname{tr}\rho_{in}^n$ is constructed by taking $n$ copies of this cut sheet and sewing them up according to the rules of matrix multiplication and tracing.

For definiteness, let us take the simplest non-trivial case, $n = 2$, and expand $\operatorname{tr}\rho_{in}^2$.

$$\operatorname{tr}\rho_{in}^2 = \int DU_{in}^1 DU_{in}^2 DU_{out}^1 DU_{out}^2 \ \langle U_{in}^1, U_{out}^1|\psi\rangle \langle\psi|U_{in}^2, U_{out}^1\rangle \langle U_{in}^2, U_{out}^2|\psi\rangle \langle\psi|U_{in}^1, U_{out}^2\rangle \tag{2.25}$$

This expression, though seemingly complicated, follows simply from the definitions of partial trace and matrix multiplication. Here, $DU$ is the product of Haar measures on every link for continuous gauge groups and just a sum for discrete groups.

The path integral comes in when we restrict to the case where $|\psi\rangle$ is the vacuum $|0\rangle$. This is because the vacuum can be created by an infinite Euclidean time-evolution,

$$|0\rangle = \lim_{T\to\infty} \frac{e^{-HT}|arb.\rangle}{\sqrt{\langle arb.|e^{-2HT}|arb.\rangle}}, \tag{2.26}$$

and this time evolution can be written as a path integral,

$$\langle\{U_\alpha(\mathbf{n})\}|0\rangle = \int^{U_\alpha(\mathbf{n},t=0)=U_\alpha(\mathbf{n})} \prod_{\mathbf{n},t,\alpha} dU_\alpha(\mathbf{n},t) e^{-S}, \quad \alpha = 1,2,3, \tag{2.27}$$

where we have fixed temporal gauge $U_0(\mathbf{n}) = 1$ and taken time to be continuous, and $dU$ is the Haar measure.

Plugging this into (2.25), we find that $\operatorname{tr} \rho_{in}^2$ is calculated by the path integral

$$\operatorname{tr} \rho_{in}^2 = \frac{Z_2}{Z_1^2}, \tag{2.28}$$

where $Z_1$ is the path integral corresponding to $\langle arb.|e^{-2HT}|arb.\rangle$, and $Z_2$ is the sewing up of the four half-space-time path integrals with identifications as shown in figure (2.3).

More generally, the procedure for constructing $Z_n$ can be described as follows. Take $2n$ half-space-time path integrals, $n$ extending from $t = -\infty$ to $t = 0$ (corresponding to the kets) and $n$ extending from $t = 0$ to $t = \infty$ (corresponding to the bras). Pair them up and sew the outside variables in each pair, where sewing corresponds to identifying the variables and integrating over them; this gives us $n$ sheets with cuts along the inside region at $t = 0$. These cuts correspond to the $U_{in}$s at $t = 0^-$ (the value on the ket path integral) and that at $t = 0^+$ not being identified with each other. Finally, sew the $t = 0^-$ inside boundary of sheet $i$ to the $t = 0^+$ inside boundary of sheet $i + 1$, with sheet $n + 1$ being the same as sheet 1. Sewing up the outside on the same sheet corresponds to partial tracing, and the sewing up across sheets of the inside variables corresponds to matrix multiplication.

Finally, noting that the smooth full-space-time path integral $Z_1$ is exactly what one obtains from the above procedure for $n = 1$, we see that

$$\operatorname{tr} \rho_{in}^n = \frac{Z_n}{Z_1^n}, \tag{2.29}$$

and the EE is

$$S_E = \lim_{n\to 1}(1 - \partial_n) \log Z_n. \tag{2.30}$$

A popular way to do this continuation to non-integer $n$ is to note that the boundary conditions on the $n$-fold cover are such that one needs to go an angle $2\pi n$ around the boundary to reach the same sheet, which is the general form of a conical singularity,[2] and consequently consider manifolds with similar boundary conditions but for non-integer $n$. On the lattice, of course, the conical singularity corresponds to the boundary points having a different coordination number than the rest of the vertices, and so this procedure isn't possible; we don't worry about this.

Coming back to the gauge theory case, we note that the natural path integral to use for translating (2.26) is the continuous-time temporal gauge path integral, ie the path integral used to derive the Hamiltonian formulation in the first place. Because we are considering

[2] Cones as usually thought of have $n < 1$.

compact gauge groups, we do not need to perform any further gauge-fixing; thus, we will end up with the un-gauge-fixed vacuum state, which is the state we would like to use for the EHS definition. Secondly, we note that the full replica trick, the way we have defined it, naturally computes the analogue of (2.25), which is exactly the expression for $\mathrm{tr}\,\rho_{in}^n$ in the EHS definition. Thus, we see that very generally, independent of gauge group, the EHS definition can be calculated by a lattice replica trick path integral.

### 2.4.2 The Continuum Limit in Four-Dimensional $U(1)$ Gauge Theory

For $G = U(1)$, the Haar measure is

$$dU = \frac{d\theta}{2\pi}. \tag{2.31}$$

The continuous-time lattice action for $U(1)$ theories is [8]

$$S = \int dt \sum_{\mathbf{n}} \left\{ \frac{g^2}{2} \sum_{\alpha} \dot{\theta}_\alpha(\mathbf{n})^2 - \frac{1}{g^2} \sum_{\alpha,\beta} \cos\Big(\theta_\alpha(\mathbf{n}) + \theta_\beta(\mathbf{n}+\hat{\alpha}) - \theta_\beta(\mathbf{n}) - \theta_\alpha(\mathbf{n}+\hat{\beta})\Big) \right\}. \tag{2.32}$$

This goes over into the continuum $F^2$ action in the weak-coupling limit (when the spread of the $\theta$s is much smaller than the size of the $U(1)$ so that we can treat the variable as non-compact), with the variable redefinition

$$A_\alpha(\mathbf{n}) = \frac{\theta_\alpha(\mathbf{n})}{g\varepsilon}. \tag{2.33}$$

There are important constraints on how small $g$ has to be for this limit to exist [8]. At the moment, we just assume this limit exists and go on.

The above variable redefinition has an effect on the Haar measure, which is now

$$dU = \frac{g\varepsilon}{2\pi} dA \equiv \frac{dA}{V_G}. \tag{2.34}$$

Here, we call $V_G = 2\pi/(g\varepsilon)$ the volume of the gauge group.

We gauge-fix the path integral in Coulomb gauge. To do so, we insert a factor of 1, in the form

$$V_G^{N_V} \Delta_{FP} \prod_{\mathbf{n}} \delta\left( \frac{1}{\varepsilon} \sum_{\alpha} \Delta_\alpha A_\alpha(\mathbf{n}) \right). \tag{2.35}$$

Here, $N_V$ is the number of vertices. The factor of $V_G^{N_V}$ is really the integration over all the gauge transformations, which factorises out, as shown in any standard treatment of the Faddeev-Popov procedure. $\Delta_{FP}$ is the Faddeev-Popov determinant, which we will now proceed to calculate.

To calculate the Faddeev-Popov determinant, it is more useful to gauge-fix in momentum space. First, we decompose the vector potential as

$$A_\alpha(\mathbf{n}) = \frac{1}{V}\sum_k e^{i\mathbf{k}\cdot\mathbf{n}}(A^L_\alpha(\mathbf{k}) + A^{T1}_\alpha(\mathbf{k}) + A^{T2}_\alpha(\mathbf{k})), \tag{2.36}$$

where $A^L$ is the longitudinal mode, and $A^{T1}, A^{T2}$ are the transverse modes. For explicitness, the longitudinal mode on the lattice is of the form

$$A^L_\alpha(\mathbf{k}) = (e^{ik_\alpha\varepsilon} - 1)a_L(\mathbf{k}) \tag{2.37}$$

and the transverse modes are similarly formed out of a scalar field multiplied by two vectors orthogonal to the momentum. The coulomb gauge condition, then, is

$$\frac{1 - e^{-ik_\alpha\varepsilon}}{\varepsilon}A^L_\alpha(\mathbf{k}) = 0, \tag{2.38}$$

since the transverse modes have no divergence.

Now consider a gauge-transformation given by the scalar field

$$\chi(\mathbf{n}) = \frac{1}{V}\sum_{\mathbf{k}\neq 0} e^{i\mathbf{k}\cdot\mathbf{n}}\chi(\mathbf{k}). \tag{2.39}$$

We have excluded the 0-mode because it doesn't change the vector potential and therefore isn't a real gauge transformation. Under this transformation, the longitudinal mode changes as

$$A^L_\alpha(\mathbf{k}) \to A^L_\alpha(\mathbf{k}) + \left(e^{ik_\alpha\varepsilon} - 1\right)\chi(\mathbf{k}) \tag{2.40}$$

and the other two modes are unaffected. The Faddeev-Popov determinant is the determinant of the matrix

$$\frac{\partial\sum_\alpha \frac{1-e^{-ik_\alpha\varepsilon}}{\varepsilon}A^L_\alpha(\mathbf{k})}{\partial\chi(\mathbf{k}')} = \delta_{\mathbf{k},\mathbf{k}'}\sum_\alpha\left|\frac{e^{ik_\alpha\varepsilon}-1}{\varepsilon}\right|^2. \tag{2.41}$$

Therefore, the Faddeev-Popov determinant is

$$\Delta_{FP} = \prod_{\mathbf{k}\neq 0}\sum_\alpha\left|\frac{e^{ik_\alpha\varepsilon}-1}{\varepsilon}\right|^2. \tag{2.42}$$

Thus, (2.35) now reads

$$1 = V_G^{N_V}\prod_{t<0}\prod_{\mathbf{k}\neq 0}\left\{\sum_\alpha\left|\frac{e^{ik_\alpha\varepsilon}-1}{\varepsilon}\right|^2\right\}\delta\left(\frac{1-e^{-ik_\alpha\varepsilon}}{\varepsilon}A^L_\alpha(\mathbf{k})\right), \tag{2.43}$$

where the $\prod_{t<0}$ turn up because we need to fix gauge at every time-slice and the reason we don't include $t = 0$ is simply that we aren't integrating over the fields in that time-slice, and the gauge-fixed path integral is

$$\langle A_\alpha(\mathbf{n})|0\rangle \propto V_G^{N_V-N_L-N_t} \prod_{t<0}\prod_{\mathbf{k}\neq 0}\left\{\sum_\alpha \left|\frac{e^{ik_\alpha\varepsilon}-1}{\varepsilon}\right|^2\right\} \int^{A_\alpha(\mathbf{n},t=0)=A_\alpha(\mathbf{n})} DA\; e^{-S}\prod_{t<0}\prod_{\mathbf{k}\neq 0}\delta\left(\frac{1-e^{-ik_\alpha\varepsilon}}{\varepsilon}A_\alpha^L(\mathbf{k})\right), \tag{2.44}$$

where $N_t$ is the number of time-slices on which we are integrating over the fields.

The extra $-N_t$ in the exponent of $V_G$ comes from the fact that the gauge-fixing is not entirely local, since whether a gauge transformation is constant or not is a global property; thus, there is one less gauge trnasformation than there are numbers of points on each time-slice. This change in the exponent can be also interpreted in a manner allowing for gauge transformations to be local. If we allow constant gauge transformations (and therefore independent gauge transformations at every vertex), the integral over the gauge transformations gives $V_G^{N_V}$. However, the Faddeev-Popov determinant for the zero-mode is different from the ones for the non-zero modes: it is just the volume of the gauge group, and therefore

$$\Delta_{FP}^0 = V_G^{-N_t}. \tag{2.45}$$

Now, since $e^{ik_\alpha\varepsilon}-1$ in the continuum limit is $ik_\alpha\varepsilon$, the continuum limit of this path integral, up to factors of volume of gauge group which we'll deal with separately, is

$$\langle A_\alpha(\mathbf{x})|0\rangle \propto \prod_{t<0}\det\left(\boldsymbol{\nabla}^2\right)\int^{A_\alpha(\mathbf{x},t=0)=A_\alpha(\mathbf{x})} DA\; e^{-S}\delta\left(\vec{\boldsymbol{\nabla}}\cdot\vec{A}(\mathbf{x},t)\right), \tag{2.46}$$

which is just the continuum path integral.

Now, we need to sew up these half-space path integrals to form the ones used in the replica trick. We continue working with the lattice versions, and take the continuum limit at the last possible moment.

There are two distinct sewn-up path integral we need to form, $Z_1$ and $Z_n$. Creating $Z_1$ is absolutely straightforward. We just take the two half-space path integrals, and integrate over the field configurations at $t = 0$ with a gauge-fixing delta function and a Faddeev-Popov determinant in this time-slice (recall that the half-space path integrals didn't involve gauge-fixing in the $t = 0$ slice) to get a path integral that looks very much like that in (2.44), with the exponent of $V_G$ becoming $N_V^{tot} - N_L^{tot} - N_t^{tot}$, where these are the number of vertices, links and time-slices in the whole space-time.

Creating $Z_n$ is a little more subtle. To see why, consider (2.25), the expansion of $\operatorname{tr}\rho_{in}^2$ into wavefunctionals. Since $|0\rangle$ is a gauge-invariant state, $|0\rangle = G|0\rangle$. Let us take $G$ to be a gauge transformation on the boundary, so that we can decompose its action into an inside action and an outside action, $G = G_{in}G_{out}$, and see what we need to do to get the integral back into its original form.

$$\begin{aligned}
\operatorname{tr}\rho_{in}^2 &= \int DU_{out,1}DU_{out,2}DU_{in,1}DU_{in,2} \\
&\qquad \langle U_{out,1}, U_{in,1}|G|0\rangle \ \langle 0|G^\dagger|U_{out,1}, U_{in,2}\rangle \ \langle U_{out,2}, U_{in,2}|0\rangle \ \langle 0|U_{out,2}, U_{in,1}\rangle \\
&= \int DU_{out,1}DU_{out,2}DU_{in,1}DU_{in,2} \\
&\qquad \left\langle U^G_{out,1}, U^G_{in,1}\middle|0\right\rangle \left\langle 0\middle|U^G_{out,1}, U^G_{in,2}\right\rangle \langle U_{out,2}, U_{in,2}|0\rangle \ \langle 0|U_{out,2}, U_{in,1}\rangle \\
&= \int D\tilde{U}_{out,1}DU_{out,2}D\tilde{U}_{in,1}D\tilde{U}_{in,2} \\
&\qquad \left\langle \tilde{U}_{out,1}, \tilde{U}_{in,1}\middle|0\right\rangle \left\langle 0\middle|\tilde{U}_{out,1}, \tilde{U}_{in,2}\right\rangle \left\langle U_{out,2}, (\tilde{U}_{in,2})^{G^\dagger}\middle|0\right\rangle \left\langle 0\middle|U_{out,2}, (\tilde{U}_{in,1})^{G^\dagger}\right\rangle \\
&= \int D\tilde{U}_{out,1}D\tilde{U}_{out,2}D\tilde{U}_{in,1}D\tilde{U}_{in,2} \left\langle \tilde{U}_{out,1}, \tilde{U}_{in,1}\middle|0\right\rangle \\
&\qquad \left\langle 0\middle|\tilde{U}_{out,1}, \tilde{U}_{in,2}\right\rangle \left\langle \tilde{U}_{out,2}, \tilde{U}_{in,2}\middle|G^\dagger\middle|0\right\rangle \left\langle 0\middle|G\middle|\tilde{U}_{out,2}, \tilde{U}_{in,1}\right\rangle \\
&= \int D\tilde{U}_{out,1}D\tilde{U}_{out,2}D\tilde{U}_{in,1}D\tilde{U}_{in,2} \left\langle \tilde{U}_{out,1}, \tilde{U}_{in,1}\middle|0\right\rangle \\
&\qquad \left\langle 0\middle|\tilde{U}_{out,1}, \tilde{U}_{in,2}\right\rangle \left\langle \tilde{U}_{out,2}, \tilde{U}_{in,2}\middle|0\right\rangle \left\langle 0\middle|\tilde{U}_{out,2}, \tilde{U}_{in,1}\right\rangle .
\end{aligned} \tag{2.47}$$

Above, we have used the invariance of the Haar measure in several steps. The important observation from these manipulations is the following: that, for all four subregions, $\tilde{U} = U^G$. Therefore, the consistent gauge transformations on the boundary vertices act on all four subregions together.[3] This means that, in $Z_n$, there is only one gauge transformation on each boundary vertex, as opposed to $n$ in $Z_1^n$.

The first effect of this is that, when we take the ratio $Z_n/Z_1^n$, the factors of $V_G$ imperfectly cancel out to give

$$V_G^{N_{V,bd}(1-n)}, \tag{2.48}$$

where $N_{V,bd}$ is the number of boundary vertices, and so it has a contribution to the entropy

$$N_{V,bd}\log V_G = N_{V,bd}\log\frac{2\pi}{g\varepsilon}. \tag{2.49}$$

The dimension of the argument of the log will be cancelled by a compensating term in the rest of the entropy, as we'll come to later; the dimension comes from the dimension of the argument of the gauge-fixing $\delta$ function. This is a contribution to the area law,

[3] Another way to see this is that, because of the four inner products in (2.25), the Gauss' law at the boundary vertices relates electric fields of both inside regions to both outside regions (in the extended Hilbert space view, Gauss' law is imposed by the vanishing of certain combinations of such inner products).

and therefore non-universal. We will henceforth ignore this part of the path integral, and therefore entropy.

The second effect of this turns up in gauge-fixing. Recall that, in defining the half-space-time path integrals, we gauge-fixed at every time-slice except $t = 0$. Now, if we would like to gauge-fix here, we can only impose one gauge condition at every boundary vertex. The gauge-fixing we can do is complicated, so we forget about Coulomb gauge altogether; the preceding discussion served mainly to clarify the conceptual issues. In fact, it will be convenient to even lift the temporal gauge that we assumed while constructing the path integral.

Now, when we sew up the integrals to from the replica lattice, there is no gauge-fixing $\delta$ function or Faddeev-Popov determinant, and therefore there are no complications to worry about. This is a straightforward Wilson path integral on a lattice with a peculiar coordination number. Then, before changing variables from $\theta$ to $A$, we gauge-fix everything to Lorentz gauge, where on the boundary vertices this means that the sum of all the $\theta$s leaving each link is fixed to be 0. In these variables, the volume of the gauge group is just $2\pi$, and thus we see that the fact that there was a dimensional quantity inside the log in (2.49) must resolve itself.

We are left with a lattice path integral in Lorenz gauge on the replica lattice, up to the volume of gauge group factors already discussed. The continuum limit can now be taken in a manner parallel to the above discussion, provided we are willing to tolerate the existence of a conical singularity in the manifold – which we are. This completes the proof that $Z_n$ is just the continuum gauge-fixed path integral, up to the volume of gauge group factors already dealt with.

So, finally we find that

$$S_{EHS} \xrightarrow{\text{continuum}} \frac{A}{\varepsilon^2} \log \frac{2\pi}{g} + S_{\text{continuum replica trick}}. \tag{2.50}$$

Since the extra term is a state-independent pure area law contribution, it will not contribute to any universal quantities and we don't need to worry about the difference. Thus, we've shown for all practical purposes that the continuum limit of the EHS replica trick is the continuum replica trick.

#### 2.4.2.1 Discussion

Anyone familiar with the literature would be confused by the statement that the continuum replica trick definition reproduces the EHS answer, because the first sign that there was an issue with understanding entanglement in gauge theories came from an explicit calculation of

the continuum replica trick [21], which found a negative answer for the EE. The explanation proposed for this was that the continuum replica trick definition did not really have an interpretation as EE. Subsequent investigations [22, 23] found that this negative contribution was very sensitive to regularisation and sometimes even gauge-dependent, making it seem even more unlikely that the answer could be the same as a manifestly gauge-independent quantity like the EHS EE. However, it was shown in [24] that these peculiarities were pathologies of the regularisation procedure used, namely heat kernel regularisation. As we will see in section 4.3, the edge modes have the odd property of having larger fluctuations at smaller wavelengths, and the heat-kernel regulator just gives a negative answer for this sort of a probability distribution; in particular, in 4.3.2 we will see that even in a Hamiltonian calculation of the non-extractable piece, using heat-kernel regularisation will give a negative leading answer.

Finally, we note an important connection to AdS/CFT. The Ryu-Takayanagi formula in AdS/CFT [25, 26], which relates boundary entanglement to the area of a minimal surface in the bulk, was proved using the continuum replica trick in the boundary theory [27]. This means that the Ryu-Takayanagi answer calculates the EHS answer, thus telling us that the EHS answer is the interesting one to consider in that context.

## 2.5 Other Perspectives on Edge Modes

Entanglement edge modes have been understood in different ways over the years, and it is worth mentioning the different perspectives. The basic point in all these approaches is that edge modes are just large gauge transformations of the inside theory, and the block-diagonal structure is caused by the fact that these large gauge transformations are really just normal gauge transformations of the full theory. There are two different variations on this theme: first, some imagine them as a charged gas confined to the boundary that absorb the outgoing electric flux, and secondly some just deal with the large gauge transformations as large gauge transformations. These are, roughly, two conjugate bases for the same modes.

The first picture seems to have first been constructed in [28]. In this paper, the authors considered a gauge theory on a finite region, and introduced a charged scalar field on the boundary so that the large gauge transformations could also be considered as redundancies. Because of this, the momentum conjugate to the scalar field, $\Pi_\phi$, is perfectly correlated with the electric field leaving the region by the Gauss' law $E \cdot \hat{n} = \Pi_\phi$, and therefore $\phi$ is a canonically conjugate way of talking about the edge modes.

This picture was advocated by the first paper to understand the EHS definition [14], and a version of it appeared in the first replica trick calculation [21] as well. It was rediscovered

in [29] from the point of view of the $U(1)$ theory being emergent at low energies; they imagined an infinitely massive charged gas playing the role of transforming under the large gauge transformations of each region. There are claims [30], however, that these are not completely equivalent to edge modes as we have defined them. Finally, [31] introduced a similar field in the symplectic product to perform the same job, with the advantage that they were able to talk about edge modes in the phase space of gravity.

There are two instances where they were discovered as large gauge transformations directly, to my knowledge. [32] was able to reduce the replica trick path integral for Chern-Simons theory to the high-temperature partition function of a boundary Wess-Zumino-Witten theory on a small cylinder surrounding the boundary. Since the WZW theory modes are the large gauge transformations of the Chern-Simons theory [33], this calculation basically amounts to finding the edge modes as large gauge transformations on the boundary.

Finally, in a beautiful paper, [34] considered the entanglement of half of space with the other half. First, they found by quantising carefully that large gauge transformations on the two sides have to be entangled, thus achieving the block-diagonal structure. Secondly, using the fact that for the half-space the time-evolution generated by the *modular Hamiltonian* $H_{mod} = -\log \rho_{in}$ is just evolution in Rindler time and the fact that $\rho_{in}$ commutes with the edge modes, they identified the edge modes with the zero-modes of Rindler evolution. The fact that edge modes are zero-modes of the modular Hamiltonian also turns up indirectly in the calculation of the classical piece in [35].

# Chapter 3

# Interpretation

We now turn to the question of interpretation of the decomposition (2.19). We will show, following [2, 36], that the three terms correspond respectively to

1. The classical term $S_c = -\sum p_{\mathbf{k}} \log p_{\mathbf{k}}$: the entropy related to classical correlations of measurements of $\mathbf{k}$ on the inside and outside.
2. The quantum term $S_q = -\sum p_{\mathbf{k}} \rho_{in}^{\mathbf{k}} \log \rho_{in}^{\mathbf{k}}$: the actual entanglement in the state.
3. The fusion term $S_f = \sum p_{\mathbf{k}} \log d_{\mathbf{k}}$: Extra entropy that comes from correlations between non-gauge-invariant operators.

The interpretation of the fusion term was already explained in the previous chapter. In the case of $SU(2)$, this term comes from the fact that $J^z_{in,i}$ and $J^z_{out,i}$ are perfectly anti-correlated in any irrep. However, since these are not gauge-invariant operators, this correlation can't actually be measured. Mathematically, the fact that this entropy comes purely from such correlation is expressed most precisely in the fact that the minimal representation electric centre definition doesn't have this piece, since the minimal representation electric centre definition only knows about statistics of gauge-invariant operators.

Most of this chapter is about proving the interpretation of the first two terms in a very precise manner. However, we can also show it in a heuristic sense, by counting Bell inequalities. Bell inequalities are important, because their violation universally distinguishes quantum entanglement and classical correlation. In both the Bell pair $|00\rangle + |11\rangle$, which is the canonical example of an entangled state, and the classical ensemble $|00\rangle\langle 00| + |11\rangle\langle 11|$, the measurements of $\sigma^z$ on the two qubits are perfectly correlated. Entanglement is different in that it has a particular combination of correlations for non-commuting operators that the

classically corrlated state doesn't have, made precise in the fact that the Bell inequality[1]

$$\left\langle \sigma_1^z \frac{\sigma_2^z + \sigma_2^x}{\sqrt{2}} + \sigma_1^z \frac{\sigma_2^z - \sigma_2^x}{\sqrt{2}} + \sigma_1^x \frac{\sigma_2^z + \sigma_2^x}{\sqrt{2}} - \sigma_1^x \frac{\sigma_2^z - \sigma_2^x}{\sqrt{2}} \right\rangle \leq 2 \tag{3.1}$$

is violated by the Bell pair but not by the classical ensemble. The main point is that in the case of the edge modes there are no operators restricted to one side that fail to commute with the edge mode operators that we called $G_i^{in}, G_i^{out}$ in the previous chapter, and therefore there is no sensible way to even formulate a Bell inequality! For the bulk modes, however, there is no such obstruction. This in itself suggests the interpretation as given above.

To actually prove it precisely, however, we show that, by the usual operational interpretation of EE, it is really only the quantum term that should be called the EE. The operational interpretation of EE is that, for pure states, it can be interpreted as the number of Bell pairs 'stored' in the state, because $n$ copies of a pure bipartite state $|\psi\rangle$ can, in the $n \to \infty$ limit, be reversibly converted into exactly $nS_E(|\psi\rangle)$ Bell pairs using only local operations and classical communication (LOCC), as elucidated in [37–40].

These two inverse processes are called entanglement distillation and dilution. The set-up for distillation is that Alice and Bob share $n$ copies of the state $|\psi\rangle$ and each has $k$ unentangled reference qubits; distillation is the process of converting the qubits into Bell pairs by using up the entanglement in $|\psi\rangle$. The *distillable entanglement* is the maximum ratio $k/n$ achievable in the asymptotic limit. The set-up for dilution is that Alice and Bob share $k'$ Bell pairs and $n$ unentangled copies of the system that $|\psi\rangle$ lives on; dilution is the process of converting this into $n$ copies of $|\psi\rangle$ and $k'$ unentangled qubits on each side. The entanglement of dilution is minimum ratio $k'/n$ achievable in the asymptotic limit. For pure states in factorisable systems,

$$\lim_{n\to\infty} \left(\frac{k}{n}\right)_{max} = \lim_{n\to\infty} \left(\frac{k'}{n}\right)_{min} = S_E(|\psi\rangle). \tag{3.2}$$

In the case of a gauge theory, as we have already explained, we don't expect this to remain true. So, we may ask what the maximum number of Bell pairs that we may reversibly extract from a gauge theory is .

We first precisely state the problem in section 3.1. Before moving on to the protocols, we address the question of what conversions are possible, using results from SVC [41, 42], in section 3.2. Finally, we explicitly work out the distillation and diltuion protocols in sections 3.3 and 3.4 respectively.

[1] This particular Bell inequality is known as the CHSH inequality.

## 3.1 The Distillation and Dilution Problems in an Algebra with Centre

Here we precisely state the distillation and dilution problems in the case of interest. Rather than tackling it in the full complexity of a gauge theory, we specialise to a simple system and argue that it generalises to more complicated cases.

The simple system is a four-qubit system, two of which are with Alice and two with Bob, in a state of the form

$$|\psi\rangle = \sum_{i,j,k=0}^{1} \sqrt{p_j} a_{ijk} |i,j\rangle_A |j,k\rangle_B = \sum_j \sqrt{p_j} |\psi_j\rangle_{AB} \quad / \langle\psi_j|\psi_j\rangle = 1. \tag{3.3}$$

The fact that the middle two qubits are constrained to be the same is the analogue of the fact that the electric flux leaving a region is the same as that entering its complement. Further, we also impose the constraint that these two qubits can only be flipped together. That is, that the full algebra of operators is generated by

$$\mathcal{G} = \{\sigma_1^z, \sigma_1^x, \sigma_2^z, \sigma_2^x \otimes \sigma_3^x, \sigma_3^z, \sigma_4^z, \sigma_4^x\}\,. \tag{3.4}$$

This is analogous to the fact that incoming electric flux can only be changed by Wilson loops crossing the boundary, which are not part of either local algebra. The centre of either subalgebra is generated by $\{\sigma_2^z = \sigma_3^z\}$. And since the centre acts only on the middle two qubits, we abuse nomenclature by referring to the effectively two-dimensional Hilbert space of these qubits as the centre.

An important comment here is that, if one considers $n$ copies of a state, every copy has its own centre, and therefore the size of the central Hilbert space grows as $2^n$. This is different from the case considered by for example SVC, for whom the centre arose from a global conservation law like particle number, such that the number of operators in the centre stayed the same no matter how many copies were considered.

Another important comment is that the full Hilbert space here is the analogue of the gauge-invariant Hilbert space in a $\mathbb{Z}_2$ theory. The physical system it corresponds to is four links attached to the same vertex, with gauge-invariance imposed only at the common vertex, with two of the links being the system $A$ and the other two being the system $B$, see figure (3.1). An example of a mapping between the gauge theory operators and the spin system operators is: $\sigma_1^z$ is the electric operator on one of the inside links, $\sigma_2^z$ is the product of the two electric operators on the two inside links, and similarly on the outside.

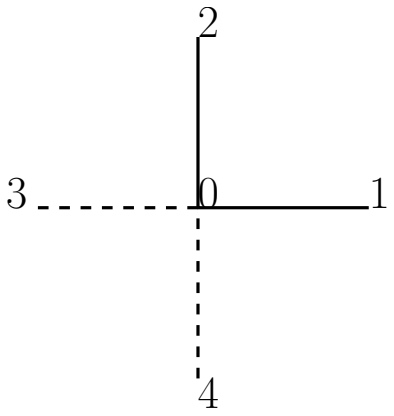


FIGURE 3.1: The example system is $\mathbb{Z}_2$ theory on this lattice, with the systems $A$ and $B$ being the solid and dashed lines respectively.

Now, we state the precise distillation and dilution questions that we will prove in sections 3.3 and 3.4.

*Distillation*: Given a state $|\psi\rangle$ and a small number $\epsilon$, what is the maximum conversion ratio $\lim_{n\to\infty} \frac{k}{n}$ such that we can convert $n$ copies of the state $|\psi\rangle$ into $k$ Bell pairs using only LOCC, so that the protocol succeeds with probability at least $1-\epsilon$?

*Dilution*: Given a small number $\epsilon$, what is the minimum conversion ratio $\lim_{n\to\infty} \frac{k'}{n}$ such that $k'$ Bell pairs are required to create $|\psi\rangle^{\otimes n}$ with fidelity greater than $1-\epsilon$ using only LOCC? Here, we assume that Alice and Bob also share the residual state

$$\left(\sum_j \sqrt{p_j}\,|0,j\rangle\,|j,0\rangle\right)^{\otimes n}. \tag{3.5}$$

We'll show that both these numbers are $S_q$.

## 3.2 The Set of Possible Conversions

Before moving on to the distillation protocol, we examine the set of possible conversions. Since no operation can mix the sectors, the relevant unitaries and measurements are the ones that act sector by sector; in particular, this means that any unitary or projection that is diagonal in the centre is an allowed operation. SVC, in fact, prove that a conversion $|\phi\rangle \to |\psi\rangle$ is deterministically possible if and only if $p_j = p'_j$ and the conversions $|\phi_j\rangle \to |\psi_j\rangle$ are deterministically possible. More generally, the conversion of $|\phi\rangle$ to $|\psi_a\rangle$ with probability $q_a$ is possible, if the average weight of each sector is preserved, and the unnormalised conversions within the sectors are possible.

A simple corollary of this is that a state that is pure centre,

$$\frac{|0,0,0,0\rangle + |0,1,1,0\rangle}{\sqrt{2}}, \tag{3.6}$$

has distillable entanglement exactly 0. This is because, no matter how many copies one takes or how many extra un-entangled qubits one considers, the state in each sector is still unentangled. A more explicit way to see this is to note that any swap gate that Alice may construct to try and extract entanglement from this must involve a $\sigma_2^x$ on the central registers, but that operator is not in the local algebra.[2]

Another corollary of this is that the reversible conversion

$$\begin{aligned}\left(p_0 \frac{|0,0,0,0\rangle + |1,0,0,1\rangle}{\sqrt{2}} + p_1 \frac{|0,1,1,0\rangle + |1,1,1,1\rangle}{\sqrt{2}}\right) \otimes |0,0\rangle \\ \longleftrightarrow (p_0 |0,0,0,0\rangle + p_1 |0,1,1,0\rangle) \otimes |\Phi_+\rangle ,\end{aligned} \tag{3.8}$$

where $|\Phi_+\rangle$ is a Bell pair, is possible by LOCC. The reason this is possible is that the centre is not entangled with the rest of the system.

## 3.3 The Distillation Protocol

Now we explicitly show the distillation protocol and prove its reversibility. For the distillation protocol we first write the state $|\psi\rangle$ in its Schmidt basis in every sector,

$$|\psi_j\rangle = \sum_I \alpha_{I,j} \left|I^j, j\right\rangle_A \left|j, I^j\right\rangle_B , \tag{3.9}$$

where the $I$s have a superscript whenever they appear as labels for states because the Schmidt basis is independent in every sector so that $\left|0^0\right\rangle \neq \left|0^1\right\rangle$.

We consider $n$ copies, and express the state of the $n$ copies as a cascading series of symmetric superposition within outer and inner 'type classes.' The $m^{th}$ outer type class is a subspace all of whose basis elements have exactly $m$ occurrences of $j = 1$. Inner type classes are labelled by a pair $(l_0, l_1)$, where $l_0$ and $l_1$ are the number of occurrences of $(I = 1, j = 0)$

[2] The swap gate on two qubits has the form

$$SWAP = \frac{1+\sigma^z}{2} \otimes \frac{1+\sigma^z}{2} + \frac{1-\sigma^z}{2} \otimes \frac{1-\sigma^z}{2} + \sigma^- \otimes \sigma^+ + \sigma^+ \otimes \sigma^-. \tag{3.7}$$

and $(I = 1, j = 1)$ respectively. First, we write it as a sum over outer type classes,

$$|\psi\rangle^{\otimes n} = \sum_{(j_r)} \sqrt{\prod_r p_{j_r}} \, |(\psi_{j_r})\rangle_{AB} \quad / \, |(\psi_{j_r})\rangle = \bigotimes_r |\psi_{j_r}\rangle_{AB}$$
$$= \sum_{m=0}^{n} \sqrt{p_0^{n-m} p_1^m \binom{n}{m}} \left( \frac{1}{\sqrt{\binom{n}{m}}} \sum_{(j_r)|m} |(\psi_{j_r})\rangle \right), \tag{3.10}$$

where $\sum_{(j_r)|m}$ means a sum over all lists such that $\sum j_r = m$. We consider the decomposition into inner type classes within each superselection sector (the list $(j_r)$). The decomposition is

$$|(\psi_{j_r})\rangle = \sum_{l_1=0}^{m} \sum_{l_0=0}^{n-m} \alpha_{1,1}^{l_1} \alpha_{0,1}^{m-l_1} \alpha_{1,0}^{l_0} \alpha_{0,0}^{n-m-l_0} \sqrt{\binom{m}{l_1}\binom{n-m}{l_0}}$$
$$\left( \frac{1}{\sqrt{\binom{m}{l_1}\binom{n-m}{l_0}}} \sum_{(I_r)|l_0,l_1} \left|(I_r^{j_r}, j_r)\right\rangle_A \left|(j_r, I_r^{j_r})\right\rangle_B \right). \tag{3.11}$$

A great simplification has occurred here; the coefficients depend only on the type class and not the superselection sector. Thus, we can pull the other summations out and write the state in the form

$$|\psi\rangle^{\otimes n} = \sum_m \sum_{l_0,l_1} \sqrt{p_0^{n-m} p_1^m \binom{n}{m}} \alpha_{1,1}^{l_1} \alpha_{0,1}^{m-l_1} \alpha_{1,0}^{l_0} \alpha_{0,0}^{n-m-l_0} \sqrt{\binom{m}{l_1}\binom{n-m}{l_0}}$$
$$\times \left( \frac{1}{\sqrt{\binom{m}{l_1}\binom{n-m}{l_0}}} \sum_{(I_r)|l_0,l_1} \left[ \frac{1}{\sqrt{\binom{n}{m}}} \sum_{(j_r)|m} \left|(I_r^{j_r}, j_r)\right\rangle_A \left|(j_r, I_r^{j_r})\right\rangle_B \right] \right). \tag{3.12}$$

The distillation protocol is simple. Alice measures $m$, and then $l_0, l_1$, obtaining the state in parentheses with probability given by the first line. While the resulting state is a symmetric superposition and therefore maximally entangled, it has an arbitrary size and therefore any transfer onto qubits is bound to be inefficient. So, she repeats the procedure $h$ times, till the product

$$D_h = \prod_{\eta=1}^{h} \binom{m_\eta}{l_{0,\eta}} \binom{n-m_\eta}{l_{1,\eta}} \tag{3.13}$$

satisfies, for some $s$,

$$2^s \leq D_h \leq 2^s(1+\epsilon), \tag{3.14}$$

with $\epsilon << \prod_\eta \binom{n}{m_\eta}^{-1}$. The final joint state is also a symmetric superposition of the same form, except that the outer superposition has $D_h$ states and the inner one has $\prod_\eta \binom{n}{m_\eta}$ states. Then, she performs a measurement within every sector that projects onto a subspace of size $2^s$ (on each side) with probability $1-\epsilon'$ and size $\epsilon 2^s$ with probability $\epsilon'$. Because of the

relative size of the subspaces, $\epsilon' < \epsilon$. If the measurement projects onto the smaller subspace (for any sector), the protocol fails; otherwise, she performs swap gates on the non-central qubits and the un-entangled qubits to transfer the entanglement onto the new qubits. Bob makes the same type class measurements on his side – he's guaranteed to get the same answer – and performs similar swaps.

The conversion thus achieved is

$$|\psi\rangle_{AB}^{\otimes nh} \bigotimes |0\rangle_A^{\otimes s} \otimes |0\rangle_B^{\otimes s} \to \frac{1}{\sqrt{\prod_\eta \binom{n}{m_\eta}}} \sum_{(j_r)|\sum j_r = \sum m_\eta} \left|(0^{j_r}, j_r)\right\rangle_A \otimes \left|(j_r, 0^{j_r})\right\rangle_B \bigotimes |\Phi_+\rangle^{\otimes s}, \tag{3.15}$$

with failure probability $\epsilon' \prod_\eta \binom{n}{m_\eta} << 1$. The extra factor comes from the number of ways it can fail (at first order).

For large values of $n$, from Stirling's formula,

$$p_1^m p_0^{n-m} \binom{n}{m} = \sqrt{\frac{n}{m(n-m)}} 2^{-n(H(m/n)-H(p_0))}. \tag{3.16}$$

The variation in this is dominated by the exponential term, and the exponential term is sharply peaked at $m = np_1$ in the limit $n \to \infty$. From applying a similar limit to the other probabilities, we get that

$$\begin{aligned} &\sum m_\eta \approx nhp_1, \\ &\sum l_{1,\eta} \approx nhp_1 |\alpha_{1,1}|^2, \\ &\sum l_{1,\eta} \approx nhp_0 |\alpha_{1,0}|^2, \\ &\Rightarrow D_h \approx 2^{nh \sum_j p_j H(|\alpha_{1,j}|^2)}. \end{aligned} \tag{3.17}$$

This, then, means that $s$ approaches $nh \sum_j p_j H(|\alpha_{1,j}|^2)$ from below in the $n \to \infty$ limit, and the number of Bell pairs extracted by the protocol per copy of the state is $S_q$.

## 3.4 The Dilution Protocol

For any two small numbers $\delta_1$ and $\delta_2$ and number of copies $n$, let us define a subspace that is spanned by lists $(I_r, j_r)$ such that

$$2^{-n(S_c+\delta_1)} \leq \prod_r p_{j_r} \leq 2^{-n(S_c-\delta_1)} \quad \text{and} \quad 2^{-n(S_q+\delta_2)} \leq \prod_r |\alpha_{I_r j_r}|^2 \leq 2^{-n(S_q-\delta_2)}. \tag{3.18}$$

Here, $\delta_{1,2} \sim \frac{1}{\sqrt{n}}$. By the law of large numbers, it is always possible to pick an $n$ large enough such that this subspace has probability more than $1 - \epsilon$ (recall that $\epsilon$ has already

been specified, according to the statement of the problem). We pick such an $n$. This subspace is called a typical subspace, since the frequencies of the various indices are close to the mean predicted by probabilties.

We note that $\prod p_{j_r}$ depends only on the number of occurrences of $j = 1$ in the list; let us call this $m$. Similarly, the second product depends only on the number of occurrences of the pairs $(I = 1, j = 1)$ and $(I = 1, j = 0)$; let us call these $l_1$ and $l_0$ respectively. Then the inequalities can be rewritten

$$\begin{aligned} 2^{-n(S_c+\delta_1)} \leq & p_0^{n-m} p_1^m \leq 2^{-n(S_c-\delta_1)} \quad \text{and} \\ 2^{-n(S_q+\delta_2)} \leq & |\alpha_{00}|^{2(n-m-l_0)} |\alpha_{10}|^{2l_0} |\alpha_{01}|^{2(m-l_1)} |\alpha_{11}|^{2l_1} \leq 2^{-n(S_q-\delta_2)}. \end{aligned} \tag{3.19}$$

We also note that, within the typical subspace, the number of possible values of $m$ is $\frac{2n\delta_1}{|\log(p_0/p_1)|}$. Of course, when $p_0$ becomes really close to $\frac{1}{2}$, this value becomes bigger than $n$ itself — this comes from the fact that changing $m$ doesn't change the product in (3.18) very much — and so the actual number of allowed values of $m$ is actually $2^{\delta q}$ with

$$\delta q = \log \min \left( \frac{2n\delta_1}{|\log(p_0/p_1)|}, n+1 \right). \tag{3.20}$$

For dilution, Alice and Bob begin with the state

$$\left( \sum_j \sqrt{p_j} \, |0, j\rangle \, |j, 0\rangle \right)_{AB}^{\otimes n} \otimes |\Phi_+\rangle_{CD}^{\otimes k'} \otimes |\psi\rangle_{A'C'}^{\otimes n} \otimes |0\rangle_E^{\otimes k'}, \tag{3.21}$$

where the systems $B$ and $D$ can be accessed only by Bob and the rest can be accessed only by Alice, $E$ is a system of $k'$ qubits, and

$$k' = n(S_q + \delta_2) + \delta q. \tag{3.22}$$

Alice will use the $k'$ copies of the Bell pair to teleport Bob's half of the state to him, and the extra $\delta q$ pairs will be used to keep track of the superselection sector information.

Alice first performs a typical subspace measurement on $A'$, and also on $A$ (where this typical subspace is defined only by $\delta_1$). This projects onto the orthogonal complement with probability smaller than $\epsilon$; in this case the protocol has failed. In the other case, the state on $A'C'$ is now

$$N \sum_m{}' \sqrt{p_0^{n-m} p_1^m} \sum_{l_0, l_1}{}' A(m, l_0, l_1) \sum_{(I_r, j_r) | m, l_0, l_1} \left| (I_r^{j_r}, j_r) \right\rangle_{A'} \left| (j_r, I_r^{j_r}) \right\rangle_{C'}, \tag{3.23}$$

where $N$ is a normalisation constant whose value can be easily worked out but isn't important, $A(m, l_0, l_1) = \alpha_{00}^{(n-m-l_0)} \alpha_{10}^{l_0} \alpha_{01}^{(m-l_1)} \alpha_{11}^{l_1}$ is the probability amplitude of those sums and the primes on the sums indicate restriction to the typical subspace.

Without superselection sectors, Alice could have just used the Bell pairs to teleport the state of the $C'$ system to $D$. However, she can't do it in this case, since se can't make an off-diagonal measurement on the $j$ qubit of the $B$ system. This is where the system $E$, which has no analogue in normal discussions of distillation, comes in. Using $\delta q$ of the qubits to keep track of the superslection sector, she can swap the state of $A'$ in each sector to the corresponding subspace of $E$. Then, since $E$ doesn't have any gauge constraints but does have the information about what the state in each sector is, she can teleport the entire system to Bob, who swaps each sector of $E$ to the corresponding sector of $B$.

Thus, since $k' \sim nS_q + \sqrt{n} + \log n$, the conversion rate is asymptotically just $S_q$, and we have proved what we set out to.

## 3.5 Discussion

First, we note a serious potential problem with our analysis: the end-point of distillation (3.15) is not the same as the starting point of dilution, (3.21). This raises the question of whether we have really achieved the maximum possible rates of distillation and dilution. Note, however, that both these states have EHS entanglement per copy of $S_c$ in the $n \to \infty$ limit. And, we already mentioned in section 3.2 that this part of the entanglement can never be extracted, indicating that these two states are close enough.

Further, we can also argue that it is impossible to extract more than $S_q$ Bell pairs per copy. Suppose someone could do so. Then, allowing non-gauge-invariant operators, we can further extract $S_c$ Bell pairs per copy. Now, we have found a way to extract more than $S_E$ Bell pairs per copy in a normal qubit system without constraints, and this is a contradiction.

More generally, the various registers may take $N > 2$ values. In this case, instead of specifying the outer type class (say) by one number $m$, we specify it with a series of numbers $m_1, m_2, \ldots$ with $m_i$ being the number of occurrences of the label $i$. Similar modifications can be carried out for the other type class indices. The rest of the arguments carry through with these modifications, with $H(p_0)$ becoming the Shannon entropy of the whole distribution $H(\{p_i\})$ and similar changes.

To go to the gauge theory case, we note that gauge-invariance inside the region not posing any problems is ensured by the fact that we are working in the gauge-invariant subspace throughout; in particular, the measurements and swaps must be defined while restricted to

the gauge-invariant subspace. This correctly accounts for $S_c + S_q$, since we were working in the gauge-invariant Hilbert space. This protocol doesn't see $S_f$, since it can't be measured by any gauge-invariant operators. However, in the argument that $S_q$ is the maximum possible extractable entanglement from a few paragraphs ago we should really have replaced $S_c$ with $S_c + S_f$.

This completes the discussion for discrete gauge groups. In a continuous Lie group, the fluxes can take an infinite number of values, but we don't worry about those complications here.

# Chapter 4

# Calculations

## 4.1 EE in the Strong-Coupling Expansion

We will study the entanglement entropy of an $SU(2)$ gauge theory to first order in the strong-coupling expansion, in order to illustrate the three types of terms in (2.19).

The Kogut-Susskind Hamiltonian [9] is

$$H = \frac{g^2}{2a}\sum_{(ij)} \hat{J}_{ij}^2 + \frac{4}{ag^2}\,\mathrm{tr}\Big(\hat{P}+\hat{P}^\dagger\Big) \equiv \frac{g^2}{2a}\left[H_0 + \frac{3\epsilon}{2}V\right], \quad \epsilon = \frac{16}{3g^4}, \tag{4.1}$$

where $a$ is the lattice spacing, $\hat{P} = \hat{U}_{ij}\hat{U}_{jk}\hat{U}_{kl}\hat{U}_{li}$ and $V = \mathrm{tr}\Big(\hat{P}+\hat{P}^\dagger\Big)$. The strong-coupling limit is the limit where $g >> 1$, so that the unperturbed ground state is the one where all the links are in the state $j_{ij} = 0$.

First, we need the action of $V$ on the ground state. From the commutation relations

$$[(\hat{U}_{ij})_{\alpha\beta}, \hat{J}_{ij}^a] = \tau^a_{\alpha\gamma}\left(\hat{U}_{ij}\right)_{\gamma\beta} \quad \text{and} \quad [(\hat{U}_{ij})_{\alpha\beta}, \hat{J}_{ji}^a] = -(\hat{U}_{ij})_{\alpha\gamma}\tau^a_{\gamma\beta}, \tag{4.2}$$

where $\tau^a = \sigma^a/2$, and the form for the $j_{ij} = 0$ state

$$|0\rangle_{ij} = \int dU_{ij}\,|U\rangle_{ij}\,, \tag{4.3}$$

we can deduce that

$$\hat{U}_{\alpha\beta}\,|0\rangle_{ij} = \frac{s_\beta}{\sqrt{2}}\left|\frac{1}{2}, m_\alpha, -m_\beta\right\rangle_{ij} \quad /s_\alpha = (-1)^\alpha = 2m_\alpha. \tag{4.4}$$

Similarly we find

$$\left(\hat{U}_{ij}\right)^{\dagger}_{\alpha\beta}|0\rangle_{ij} = \frac{s_\alpha}{\sqrt{2}}\left|\frac{1}{2}, -m_\alpha, m_\beta\right\rangle_{ij}. \tag{4.5}$$

Considering a whole plaquette at once, we find that

$$\begin{aligned}\operatorname{tr}\hat{P}\,|0,0,0,0\rangle_{ijkl} &= \frac{1}{4}\sum_{\alpha\beta\gamma\delta} s_\alpha s_\beta s_\gamma s_\delta \left|\frac{1}{2}, m_\alpha, -m_\beta\right\rangle_{ij}\left|\frac{1}{2}, m_\beta, -m_\gamma\right\rangle_{jk}\left|\frac{1}{2}, m_\gamma, -m_\delta\right\rangle_{kl}\left|\frac{1}{2}, m_\delta, -m_\alpha\right\rangle_{li} \\ &= \left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle_{ijkl},\end{aligned} \tag{4.6}$$

where $\left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle$ is the unique gauge-invariant state on a plaquette with all links having $j = 1/2$. Similarly,

$$\operatorname{tr}\hat{P}^{\dagger}\,|0,0,0,0\rangle_{ijkl} = \left|\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right\rangle_{ijkl}. \tag{4.7}$$

Thus, denoting the unperturbed ground state by $\left|0^{(0)}\right\rangle$,

$$V\left|0^{(0)}\right\rangle = 2\sum_i |P_i\rangle\,, \tag{4.8}$$

where $|P_i\rangle$ is the state with the $i^{th}$ plaquette excited to $j = 1/2$.

BY the usual rules of perturbation theory, the first-order perturbed state is

$$\left|0^{(1)}\right\rangle = \frac{\Pi}{H_0 - E_0}(E_1 - \frac{3\epsilon}{2}V)\left|0^{(0)}\right\rangle = -\epsilon\sum_i |P_i\rangle\,. \tag{4.9}$$

Now, suppose some of the links have been defined as the inside region. We define inside plaquettes to be those which are constituted of only inside links, outside plaquettes as those which are constituted by only outside links, and boundary plaquettes to be those which are constituted by both. Pursuant to this classification, we can write

$$\left|0^{(1)}\right\rangle = -\epsilon\left(\sum_{i=1}^{n_{in}}\left|P_i^{in}\right\rangle + \sum_{i=1}^{n_{bd}}\left|P_i^{bd}\right\rangle + \sum_{i=1}^{n_{out}}\left|P_i^{out}\right\rangle\right) \tag{4.10}$$

and correspondingly the whole state is

$$|0\rangle = \frac{1}{\sqrt{1+N\epsilon^2}}\left(\left|0^{(0)}\right\rangle + \left|0^{(1)}\right\rangle\right), \tag{4.11}$$

where $N = n_{in} + n_{bd} + n_{out}$ is the total number of plaquettes.

The reduced density matrix, by the naive definition, is

$$\rho_{in} = \frac{1}{1+N\epsilon^2}\left[\left\{|0\rangle_{in} - \sqrt{n_{in}}\epsilon\left(\frac{1}{\sqrt{n_{in}}}\sum_{i=1}^{n_{in}}\left|P_i^{in}\right\rangle_{in}\right)\right\}\left\{\langle 0|_{in} - \sqrt{n_{in}}\epsilon\left(\frac{1}{\sqrt{n_{in}}}\sum_{i=1}^{n_{in}}\left\langle P_i^{in}\right|_{in}\right)\right\}\right.$$
$$\left. + \epsilon^2\sum_{i=1}^{n_{bd}}\mathrm{tr}_{\mathcal{H}_{out}}\left|P_i^{bd}\right\rangle\left\langle P_i^{bd}\right| + n_{out}\epsilon^2|0\rangle_{in}\langle 0|_{in}\right]. \tag{4.12}$$

It should be noted that we have ignored some of the $O(\epsilon^2)$ terms, namely the $\left|0^{(0)}\right\rangle\left\langle 0^{(2)}\right| + \left|0^{(2)}\right\rangle\left\langle 0^{(0)}\right|$ terms; we will justify in the end that it contributes only at subleading order. There are $n_{bd}+1$ superselection sectors; one, which we will call $\mathbf{j}_0$ is the one in which all the outgoing fluxes are 0, and the other $n_{bd}$, which we will call $\mathbf{j}_i$, are the ones in which one of the boundary plaquettes is excited and the corresponding outgoing fluxes are $1/2$.

To perform the trace in the $\mathbf{j}_i$ sectors, we note that any boundary plaquette has either $1, 2$ or 3 inside links. In the case of 1 link, the reduced density matrix is

$$\frac{1}{4}\sum_{\alpha,\beta}\left|\frac{1}{2}, m_\alpha, m_\beta\right\rangle\left\langle\frac{1}{2}, m_\alpha, m_\beta\right|, \tag{4.13}$$

where we have suppressed dependence on all other links. In the case of 3 links, similarly, it is

$$\frac{1}{16}\sum_{\alpha,\beta,\alpha',\beta',\gamma,\delta} s_\alpha s_\beta s_{\alpha'} s_{\beta'}\left|\frac{1}{2}, m_\delta, -m_\alpha; \frac{1}{2}, m_\alpha, -m_\beta; \frac{1}{2}, m_\beta, -m_\gamma\right\rangle$$
$$\left\langle\frac{1}{2}, m_\delta, -m_{\alpha'}; \frac{1}{2}, m_{\alpha'}, -m_{\beta'}; \frac{1}{2}, m_{\beta'}, -m_\gamma\right|. \tag{4.14}$$

There are two case with 2 inside links; one in which there are only two boundary vertices, and one in which there are four. In the case of two boundary vertices, we have

$$\frac{1}{8}\sum_{\alpha,\beta,\beta',\gamma} s_\beta s_{\beta'}\left|\frac{1}{2}, m_\alpha, -m_\beta; \frac{1}{2}, m_\beta, m_{-\gamma}\right\rangle\left\langle\frac{1}{2}, m_\alpha, -m_{\beta'}; \frac{1}{2}, m_{\beta'}, m_{-\gamma}\right|, \tag{4.15}$$

and in the case of four we have

$$\frac{1}{16}\sum_{\alpha,\beta,\gamma,\delta}\left|\frac{1}{2}, m_\alpha, m_\beta; \frac{1}{2}, m_\gamma, m_\delta\right\rangle\left\langle\frac{1}{2}, m_\alpha, m_\beta; \frac{1}{2}, m_\gamma, m_\delta\right|. \tag{4.16}$$

In all these, if we associate indices to vertices, we see that the state is pure on the indices corresponding to the inside vertices, so that the matrix form is always

$$\left(\frac{\mathbb{1}_2}{2}\right)^{\otimes v_i^{bd}}, \tag{4.17}$$

where $v_i^{bd}$ is the number of boundary vertices in plaquette $i$. Thus the EE in these sectors is

$$\epsilon^2 v_i^{bd} \log 2. \tag{4.18}$$

Further, since this reduced density matrix is formed from only one gauge-invariant state, we can see that in the electric centre definition these sectors would have entropy 0. Therefore, all these are contributions to the inaccessible piece $S_f$.

It is now easy to find $S_c$. The various superselection sector probabilities are

$$p_{\mathbf{j}_0} = \frac{1 + (n_{in} + n_{out})\epsilon^2}{1 + N\epsilon^2} \quad \text{and} \quad p_{\mathbf{j}_i} = \epsilon^2, \tag{4.19}$$

which gives us

$$S_c = n_{bd}\epsilon^2 - n_{bd}\epsilon^2 \log \epsilon^2 \tag{4.20}$$

to leading order in $\epsilon$.

The reduced density matrix in the $\mathbf{j}_0$ sector has the form

$$\frac{1}{1 + (n_{in} + n_{out})\epsilon^2} \begin{pmatrix} 1 + n_{out}\epsilon^2 & \sqrt{n_{in}}\epsilon \\ \sqrt{n_{in}}\epsilon & n_{in}\epsilon^2 \end{pmatrix} \tag{4.21}$$

in the basis $\left\{ |0\rangle_{in}, \frac{1}{\sqrt{n_{in}}} \sum |P_i^{in}\rangle \right\}$. This matrix has eigenvalues of order $\epsilon^4$, and therefore the entropy is negligible. Thus,

$$S_q = 0 + O(\epsilon^4) \tag{4.22}$$

and there is no entanglement at leading order.

We may now justify having neglected the terms $|0^0\rangle \langle 0^{(2)}| + |0^2\rangle \langle 0^{(0)}|$. It didn't make a difference because the only sector it would have contributed to in the reduced density matrix was $\mathbf{j}_0$, and it is easy to show that that contribution would also be of order $\epsilon^4$.

Thus, we have the result for the EE

$$\begin{aligned} S_c &= n_{bd}\epsilon^2(1 - \log \epsilon^2), \\ S_q &= 0 \quad \text{and} \\ S_f &= \epsilon^2 \left( \sum_i v_i^{bd} \right) \log 2. \end{aligned} \tag{4.23}$$

Specialising to a rectangular region, $\sum_i v_i^{bd} = 2n_{bd}$, and therefore the EE is proportional to $n_{bd}$. As we will see in the next two sections, its being proportional to $n_{bd}$ instead of $n_{bd} - 1$ is a manifestation of the fact that the theory is confining in this limit.

## 4.2 The Toric Code

Now that we have investigated the behaviour of the EE in the extreme strong coupling limit, it is worth going to the other end, the extreme weak coupling limit. For continuous gauge groups, the magnetic term in the Kogut-Susskind Hamiltonian (4.1), $\mathrm{tr}\left(P + P^\dagger\right)$ has a continuous spectrum and is not a good Hamiltonian on its own; thus, we have to keep both electric and magnetic terms at leading order and the calculation is not easy, except in the case where $G = U(1)$ and the Hamiltonian is free. We will investigate that theory in section 4.3. For discrete groups, however, we *can* go to the limit in which the electric term drops out; such models are called *toric codes* [43].

The usual formulation of the toric code involves the Gauss' law constraint being imposed energetically rather than at the level of the Hilbert space, the reason being that the interest in the model comes from implementing it in the lab on systems without a fundamental gauge redundancy. This is an example of an emergent gauge theory, where the gauge symmetry is a symmetry only of the low-energy subspace. For these systems with emergent gauge symmetry, the EHS definition is the natural one. Beyond this observation, however, we will treat it exactly as the extreme weak-coupling limit of a $G$ gauge theory.

The reason the toric code is interesting is that it is a topological field theory in the IR. This has two important consequences. First, it has a ground state degeneracy dependent on the topology of the spatial manifold, and the quantum numbers that span the degenerate subspace are given by Wilson loops around non-contractible cycles of the spatial manifold. Since it's a topological field theory, the value of the loop is invariant under small deformations and therefore robust to local perturbations, thus making it a good *error-correcting code* and giving rise to the name. Secondly, and more to the point for us, the excitation spectrum has anyons. Because anyons have non-trivial braiding, there is no Hilbert space of a single anyon; however, this doesn't faze physicists, and each anyon has a property that is the generalisation of the dimension of the internal Hilbert space, called the *quantum dimension*. If $d_a$s are the quantum dimensions of the various anyons, there is another interesting quantity called the *total quantum dimension* of the theory,

$$D = \sqrt{\sum_a d_a^2}. \tag{4.24}$$

In the models under consideration, it takes the value

$$D = |G|, \tag{4.25}$$

where $|G|$ is the number of elements of the group.

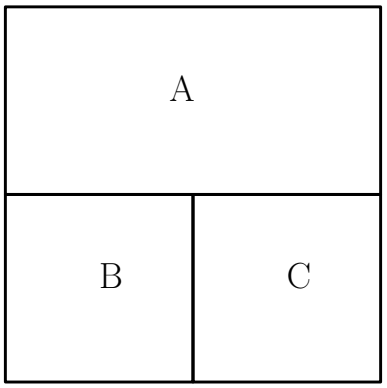


FIGURE 4.1: The three regions used to isolate the topological EE using the combination (4.27).

It was shown in [44, 45] that this total quantum dimension turns up as a correction to the area law in the EE for a region with simply connected boundary (so that the region is topologically a disk),

$$S_E = \alpha L - \log D. \tag{4.26}$$

This correction is called the *topological entanglement entropy* and is protected under both small deformations of the surface and the Hamiltonian. While we will not need it, it is often useful in calculations to isolate the topological contribution using a combination of the EEs of three regions as in figure (4.1),

$$S_{top} = -\log D = S_A + S_B + S_C - S_{AB} - S_{AC} - S_{BC} + S_{ABC}. \tag{4.27}$$

This is useful because in calculations one might worry that the constant piece is regularisation-dependent, but this combination neatly removes any such ambiguities. However, since we will work on the lattice we won't explicitly need to use this. Also, generalisations of this construction are useful to isolate the analogous topological EE in other dimensions, even when the constant term can be rendered non-universal by for example a log term [46].

With these introductory remarks, we can now introduce the model. First, we define a link operator that projects onto a particular group element,

$$T_{ij}(h)\,|g\rangle = \delta_{h,g}\,|g\rangle\,. \tag{4.28}$$

As usual, $T_{ji}(h) = T_{ij}(h^{-1})$. We then define plaquette and star operators,

$$\begin{aligned} A_i &= \sum_g \prod_{j,(ij)\in\text{links}} L_{ij}(g) \\ B_{ijkl} &= \sum_{g_1,g_2,g_3} T_{ij}(g_1)T_{jk}(g_2)T_{kl}(g_3)T_{li}(g_3^{-1}g_2^{-1}g_1^{-1}). \end{aligned} \tag{4.29}$$

The star operator $A_i$ implements a projection to the Gauss' law invariant subspace at the vertex $i$, and the plaquette operator $B_{ijkl}$ projects onto the subspace in which there is no magnetic flux through the plaquette. The Hamiltonian is

$$H = -J_e \sum_i A_i - J_m \sum_{\text{plaquettes}} B_{ijkl} \tag{4.30}$$

Notice that the Maxwell analogue of this Hamiltonian would be $\nabla \cdot E + B^2$; the lack of a term analogous to $E^2$ is why we said earlier that it is the extreme weak-coupling limit.

The other thing to notice about this is that all the operators in the Hamiltonian commute with each other, and so we can exactly diagonalise the Hamiltonian and find the ground state as the subspace in which all the $A_i$s and $B_{ijkl}$s take the value 1. That they commute with each other is the well-known fact that the value of a Wilson loop is gauge-invariant. When the spatial manifold is the plane, there is only one ground state, which can be created by taking the electric basis state in which all the links are in the trivial representation and acting with all the $B_{ijkl}$s to project onto the subspace where all the magnetic fluxes take value 1,

$$|0\rangle = \prod_{\text{plaquettes}} B_{ijkl} \, |00\cdots\rangle \, . \tag{4.31}$$

A useful picture for this arises from noting that the vacuum is a symmetric superposition of all spin networks, and thus we can think of it as a symmetric superposition of electric flux loop excitations (the electric fluxes being in loops ensures Gauss' law is satisfied on all the configurations).

We are interested in calculating the EE of a rectangular region, like figure (2.1), with the rest of the plane. We take the region to have $L$ boundary vertices and call them $i_1 \cdots i_L$ in a cyclic manner. To do this, we rewrite the vacuum in the form

$$|0\rangle = N \sum_{g_{bd}} T_{i_1 i_2}(g_{12}) T_{i_2 i_3}(g_{23}) \cdots T_{i_L i_1}(g_{L,L-1} \cdots g_{12}) B_{in}(g_{bd}) B_{out}(g_{bd}) \, |00\cdots\rangle \, , \tag{4.32}$$

where $g_{bd}$ stands for the entire boundary configuration $\{g_{12} \cdots g_{L-1,L}\}$ and $B_{in/out}(g_{bd})$ are projectors acting on the inside and outside links such that all plaquettes have no magnetic flux, and $N$ is a normalisation that we shall get to presently. Since all plaquettes inside the region have no magnetic flux, it stands to reason that the Wilson loop around the inside is also trivial, as is manifested in the fact that $g_{L1}$ is written as the inverse of $g_{12} \cdots g_{L-1,L}$.

The EE calculation is simply achieved once we make the observation that for two different boundary configurations $g_{bd}, \tilde{g}_{bd}$, $B_{out}(g_{bd})$ and $B_{out}(\tilde{g}_{bd})$ project onto orthogonal subspaces; and similarly with the $B_{in}$s. To see this, suppose that the two configurations differ on the link $(i_1 i_2)$. Then the products of the $g$s on the three links in the outside plaquette required

to make that plaquette 1 are different. Thus, there are no common configurations in the two subspaces and they are orthogonal.

This means that the reduced density matrix is diagonal in $g_{bd}$, and is

$$\rho_{in} = \tilde{N} \sum_{g_{bd}} B_{bd}(g_{bd}) B_{in}(g_{bd}) \left|00\cdots\right\rangle_{in} \left\langle 00\cdots\right| B_{in}(g_{bd}) B_{bd}(g_{bd}), \tag{4.33}$$

where $B_{bd}$ is shorthand for the product of the boundary $T$s in (4.32). Each summand here is a pure density matrix, since $B_{bd}B_{in}\left|00\cdots\right\rangle_{in}$ is a pure state on the inside. Further each summand is orthogonal, since the $B$s that correspond to different boundary configurations are orthogonal projectors. Thus, we see that the reduced density matrix is of the form

$$\rho_{in} = \frac{\mathbb{1}}{|G|^{L-1}}, \tag{4.34}$$

where we have fixed the normalisation using the fact that there are $|G|^{L-1}$ independent boundary configurations. Thus, the EE has the promised form

$$S_E = \log|G|(L-1). \tag{4.35}$$

The thing to note here is that this EE was calculated in the EHS definition and not the electric centre definition. One may ask whether the electric centre answer would have been different, or equivalently whether the fusion term contributes to this answer. It turns out that all three terms are required to give the correct topological EE, as we now show.

To show this, we need to decompose the reduced density matrix (4.33) into superselection sectors. This job is made a lot easier by the fact that the matrix is the identity on the subspace on which it has support, (4.34), and so we only need to figure out what electric configurations the matrix has non-zero support on.

The simplest way to do so, it turns out, is to relabel the subspace on which $\rho_{in}$ has support. Since each state in the subspace is uniquely fixed by the boundary configuration $g_{bd}$, we first define

$$\left|g_{bd}\right\rangle = B_{bd}(g_{bd}) B_{in}(g_{bd}) \left|00\cdots\right\rangle. \tag{4.36}$$

Now, consider the action of the inside large gauge transformation operator at one of the vertices, say $i_2$, $G^{i_2}_{in}(h) = \prod_{j|(i_2 j)\in in} L_{i_2 j}(h)$,

$$G^{i_2}_{in}(h) \left|g_{12}, g_{23}\cdots g_{L-2,L-1}\right\rangle = \left|g_{12}h^{-1}, h g_{23}\cdots g_{L-2,L-1}\right\rangle. \tag{4.37}$$

The trick turns out to be switching from link variables to vertex variables, such that each $G^i_{in}$ acts on only one vertex variable, by

$$g_{ij} = g_i g_j^{-1}, \quad g_L = 1. \tag{4.38}$$

Here, explicitly setting $g_L$ to the identity is something like a gauge-fixing, freezing an extra degree of freedom in the vertex description compared to the link description; this discrepancy occurs due to the fact that the product of all the link variables is 1. Thus, we get

$$g_1 = g_{1L} = g_{12} \cdots g_{L-1,L}, \quad g_2 = g_{21} g_1 = g_{23} \cdots g_{g_{L-1},L}, \quad \cdots \tag{4.39}$$

With this labelling, the action of, say, $G^{i_2}_{in}$ is

$$G^{i_2}_{in}(h) = |g_1, h g_2 \cdots g_{L-1}\rangle \,. \tag{4.40}$$

What then, one might ask, is the action of $G^{i_L}_{in}$? Going back to the link variables,

$$\begin{aligned} G^{i_L}_{in}(h)\, |g_{12} \cdots g_{L-1,L}, g_{L,L-1} \cdots g_{21}\rangle &= \left|g_{12} \cdots g_{L-1,L} h^{-1}, h g_{L,L-1} \cdots g_{21}\right\rangle \\ &= \left|g_{12} \cdots g_{L-1,L} h^{-1}, \left(g_{L-1,L} h^{-1}\right)^{-1} \cdots g_{21}\right\rangle \\ (4.39) \;\Rightarrow G^{i_L}_{in}(h)\, |g_1, g_2 \cdots g_{L-1}\rangle &= \left|g_1 h^{-1}, g_2 h^{-1} \cdots g_{L-1} h^{-1}\right\rangle . \end{aligned} \tag{4.41}$$

Thus, we see that this subspace is a product of $L-1$ regular representations, with the left actions given by the large gauge transformation at the first $L-1$ vertices and the joint right action given by the large gauge transformation at the $L^{th}$ vertex.

This means two things. First, since the large gauge transformations preserve the subspace, the subspace decomposes according to irreps of the operators $G^{i_1}_{in}, \cdots G^{i_{L-1}}_{in}$, let's call the irreps $r_1, \cdots r_{L-1}$. Secondly, it also means that $G^{i_L}_{in}$ acts in the reducible representation

$$r_1 \otimes \cdots \otimes r_{L-1} = \oplus_a r_a^{n_a}. \tag{4.42}$$

Here, $n_a$ is the number of times $r_a$ appears in the decomposition, and it is often bigger than one. Further, since a superselection sector is completely specified by the representations of $G^i_{in}$, all these $n_a$ subspaces are in the same superselection sector.

First, to see that this is the entire subspace that $\rho_{in}$ has non-trivial support on, let us compute its dimension,

$$\sum_{r_1 \ldots r_{L-1}} d_{r_1} \cdots d_{r_{L-1}} \sum_a n_a d_{r_a}. \tag{4.43}$$

But, from (4.42), $\sum_a n_a d_{r_a}$ is nothing but $d_{r_1} \cdots d_{r_{L-1}}$. Finally, using the fact that $\sum_r d_r^2 = |G|$, we see that the dimension is $|G|^{L-1}$ and thus this is the entire subspace.

To find the decomposition if the EE into the standard form (2.19), we need to now calculate the $p_{\mathbf{k}}$s.[1] Since $\rho_{in}$ is proportional to the identity on the subspace on which it lives, $p_{\mathbf{k}}$ is clearly the number of states in that superselection sector,

$$p_{\mathbf{k}} = \frac{d_{r_1} \cdots d_{r_L} N_{\mathbf{k}}}{|G|^{L-1}} \equiv \frac{d_{\mathbf{k}} N_{\mathbf{k}}}{|G|^{L-1}}, \tag{4.44}$$

where $N_{\mathbf{k}}$ is the number of times $r_L$ appears in $r_1 \otimes \cdots r_{L-1}$. Thus, we can write the density matrix in the form

$$\rho_{in} = \bigotimes_{\mathbf{k}} p_{\mathbf{k}} \rho_{in}^{\mathbf{k}} \otimes \frac{\mathbb{1}_{d_{\mathbf{k}}}}{d_{\mathbf{k}}}, \quad \rho_{in}^{\mathbf{k}} = \frac{\mathbb{1}_{N_{\mathbf{k}}}}{N_{\mathbf{k}}}, \tag{4.45}$$

giving the decomposition of the EE analogous to (2.19)

$$S_E = -\sum_{\mathbf{k}} p_{\mathbf{k}} \log p_{\mathbf{k}} + \sum_{\mathbf{k}} p_{\mathbf{k}} \log N_{\mathbf{k}} + \sum_{\mathbf{k}} p_{\mathbf{k}} \log d_{\mathbf{k}}. \tag{4.46}$$

Thus, we have found that the EHS definition gives an answer that agrees with the topological EE for the toric code, and also that all three terms in the decomposition (2.19) are required for this agreement to hold. This is an important piece of evidence that even the operationally empty fusion term carries physical information within it.

## 4.3 Four-Dimensional $U(1)$ Gauge Theory

Now that we have examined the weak coupling limit for discrete gauge groups, we turn to the simplest case of a continuous gauge group, $U(1)$. We will focus on the case of free Maxwell theory, with action

$$S = -\frac{1}{4} \int d^4x F_{\mu\nu} F^{\mu\nu}. \tag{4.47}$$

In four dimensions, free Maxwell theory is conformal, and so we stick to four dimensions since that will simplify the calculation.

The general form of the EE for smooth regions of a $4d$ continuum field theory is

$$S_E = c_1 \frac{A}{\epsilon^2} + C \log \frac{A}{\epsilon^2} + O(\epsilon). \tag{4.48}$$

Here, $\epsilon$ is a UV cutoff required for calculating EE; universal quantities are those independent of $\epsilon$. Here, $c_1$ isn't universal, but $C$ is, and that is what we will focus on calculating in this section.

[1] Here, $\mathbf{k}$ means the set $r_1 \cdots r_L$.

We will also specialise to the case where the region is a ball of radius $R$. In four-dimensional Maxwell theory, using conformal field theory methods, one can show that [47, 48]

$$C = -\frac{31}{90}, \tag{4.49}$$

whereas by explicit calculation of thermal entropy on hyperbolic space, [49] found that

$$C = -\frac{16}{90}. \tag{4.50}$$

In section 4.3.1, we show that the replica trick agrees with (4.49) rather than (4.50) [3], and in section 4.3.2 we show that the difference is exactly the edge mode contribution [3, 24, 35, 50–53]. Finally, in section 4.3.3, we comment on whether the integrated Gauss' law gives a topological contribution to the EE similar to the toric code [3].

### 4.3.1 The Full Entropy

We would like to calculate the coefficient of the log term in the entanglement, $C$, for a spherical region of radius $R$. The boundary of this region is an $S^2$ with area

$$A = 4\pi R^2. \tag{4.51}$$

As is well known [26], the answer for $C$ can be readily extracted from well-known facts about the $U(1)$ theory since it is conformally invariant.

The argument is as follows. Consider infinitesimally rescaling the radius of the sphere

$$R \to R(1+\delta), \tag{4.52}$$

$\delta \ll 1$, while keeping the UV cut-off $\epsilon$ fixed. Then the change in $S_E$ is given by

$$\left.\frac{\partial S_E}{\partial \delta}\right|_{\delta=0} = 2C_1\frac{A}{\epsilon^2} + 2C. \tag{4.53}$$

The coefficient of interest, $C$, is the term on the RHS which is independent of $\epsilon$. Note that by scale-invariance the UV-finite terms must be independent of $A$ and therefore do not contribute to the RHS of eq.(4.53).

The rescaling of the radius $R$ can be carried out by rescaling the metric. To analyse the consequences, consider the path integral, but now in the presence of a background metric,

$$Z[g_{\mu\nu}] = \int [DA] e^{-S[g_{\mu\nu},A_\mu]} \delta(f(A_\mu)) \left|\det\left(\frac{\partial f(A^\omega_\mu)}{\partial \omega(\mathbf{x})}\right)\right|. \tag{4.54}$$

The metric appears in the action as shown explicitly above, but also in the measure and in general in the gauge-fixing delta function and associated determinant. The stress energy tensor is given by

$$\langle T_{\mu\nu}\rangle = \frac{\delta \ln Z[g_{\mu\nu}]}{\delta g_{\mu\nu}}. \tag{4.55}$$

For a conformal theory in $3+1$ dimensions, it is well known that

$$\int \sqrt{g}\langle T^{\mu}_{\mu}\rangle = aE_4 + cW^2 \tag{4.56}$$

where $E_4$ is the integral of the Euler density,

$$E_4 = \frac{1}{64\pi^2}\int \sqrt{g}\left(R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} - 4R_{\mu\nu}R^{\mu\nu} + R^2\right) \tag{4.57}$$

and $W^2$ is the integral of the square of the Weyl tensor given by

$$W^2 = -\frac{1}{64\pi^2}\int \sqrt{g}\left(R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} - 2R_{\mu\nu}R^{\mu\nu} + \frac{1}{3}R^2\right). \tag{4.58}$$

The coefficient of the $E_4$ term is called the "a-anomaly" coefficient.

Now if the metric is rescaled by $g_{\mu\nu} \to g_{\mu\nu}(1+2\delta)$ then this will accomplish the required scaling of $R$, eq.(4.52). The change $\partial_\delta \ln Z[g_{\mu\nu}]$ is given by,

$$\partial_\delta \ln Z[g_{\mu\nu}]|_{\delta\to 0} = 2\int_M \sqrt{g}T^{\mu}_{\mu} \tag{4.59}$$

$$= 2aE_4 + 2cW^2. \tag{4.60}$$

So, the derivative of the EE, using (2.30), is

$$\left.\frac{\partial S_E}{\partial \delta}\right|_{\delta\to 0} = (1-\partial_n)\partial_\delta \ln Z_n[\delta]\big|_{\delta\to 0, n\to 1} \tag{4.61}$$

Comparing with eq.(4.53) we see that $C$ can be obtained once the RHS in eq.(4.61) can be calculated.

An important point, already emphasised, is that the $n$-fold cover is a singular space. So the strategy we can use, as in [54], is to first slightly "de-singularise" the space and then take the singular limit of interest. For calculating $Z_n$ for the entanglement of the sphere of radius $R$ we work on a smoothed out space with metric,

$$ds^2 = r^2 d\tau^2 + \frac{r^2+b^2n^2}{r^2+b^2}dr^2 + (R + r^n c^{1-n}\cos\tau)^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{4.62}$$

Here $b, c$ are extra parameters introduced to smooth out the conical singularity at $r = 0$ along an $S^2$ of radius $R$. One then calculates the various integrals involved (for integer $n$) and takes the $n \to 1$ limit, in which the dependence on these extra parameters drops out.

In [54] it was shown that for the case, eq.(4.62) one gets that

$$\begin{aligned} \int \sqrt{g} R^2 &= O\left((n-1)^2\right) \\ \int \sqrt{g} R_{\mu\nu} R^{\mu\nu} &= 32\pi^2 (n-1) \\ \int \sqrt{g} R_{\mu\nu\alpha\beta} R^{\mu\nu\alpha\beta} &= 64\pi^2 (n-1). \end{aligned} \tag{4.63}$$

As a result $W^2 \sim O((n-1)^2)$, while

$$E_4 = -(n-1). \tag{4.64}$$

From eq.(4.61) it then follows that

$$C = a. \tag{4.65}$$

In the $U(1)$ it is well known that $a = -\frac{31}{90}$ leading to

$$C = -\frac{31}{90}, \tag{4.66}$$

which agrees with (4.49), as promised.

### 4.3.2 The Non-Extractable Piece

Now that we have calculated the universal log term in the entanglement, the question naturally arises of how much of it is from the classical piece in the EE. In this section, we will independently calculate the classical piece and show that there is a log term with coefficient $-\frac{1}{6}$, exactly the discrepancy between the answers of [47, 48] and [49]. We will do so by finding the two-point function of the radial component of the electric field on the surface of the sphere. Since the theory is Gaussian, the probability distribution of the radial component is completely determined by the two-point function.

To see that it is the radial component whose probability distribution we need, we can imagine discretising the theory on a spherical lattice; then, the only outside link touching a boundary vertex is the radial one, and therefore it is the one whose value determines the superselection sectors.

Before calculating the two-point function, let us see how it can be used to calculate the entropy. Let us call the two-point function

$$G_{rr}(\mathbf{x}-\mathbf{y}) = \hat{x}_i \langle E_i(\mathbf{x}) E_j(\mathbf{y}) \rangle \hat{y}_j. \tag{4.67}$$

Since a Gaussian distribution is completely determined by its two-point functions, the correct probability distribution is one that reproduces all of the above. Such a probability distribution is

$$p(E_r(\mathbf{x})) = \frac{1}{\sqrt{\det(2\pi G)}} e^{-\frac{1}{2}\int d^2x d^2y E_r(\mathbf{x}) G^{-1}_{rr}(\mathbf{x}-\mathbf{y}) E_r(\mathbf{y})}, \tag{4.68}$$

where $G^{-1}_{rr}$ is the matrix inverse of $G_{rr}$. To find the matrix inverse, we need to diagonalise the two-point function. Since the problem has a spherical symmetry, the correct basis is that of spherical harmonics. The probability distribution then becomes

$$p(E_r(\mathbf{x})) = \prod_{l,m} \frac{1}{\sqrt{2\pi G^l_{rr}}} e^{-\frac{|E^{lm}_r|^2}{2G^l_{rr}}}. \tag{4.69}$$

The Green's function doesn't depend on $m$ because, as we will soon see, of the azimuthal symmetry of the problem.

Notice, now, that the above is a probability distribution and not a probability; it is somewhat unsuitable for calculating the entropy since it has a dimension and therefore taking its log doesn't make much sense. The way to tackle this is to pay careful attention to the measure. On the lattice, the electric field is discretised in units of 1; for the continuum limit, we define the effectively continuous variable

$$E_i = \frac{g}{\varepsilon^2} L_i. \tag{4.70}$$

So the microscopic measure is really just a sum over $L_r(\mathbf{n})$. The corresponding continuum measure, since the electric field can't change by less than $g/\varepsilon^2$, is[2]

$$\frac{dE_r(\mathbf{x})}{g/\varepsilon^2}. \tag{4.71}$$

So, the probability, with which we'll calculate the entropy, is

$$P(E_r(\mathbf{x})) = \frac{g}{\varepsilon^2} p(E_r(\mathbf{x})). \tag{4.72}$$

[2]This is essentially parallel to the modification of the measure of the phase space integration in classical statistical mechanics. There, the discreteness comes because it is impossible to localise a particle into a region of phase space with volume less than $\hbar$.

The entropy of this probability function is

$$\begin{aligned} S &= -\sum_{l,m} \int \frac{\varepsilon^2 dE_{l,m}}{g} P_{lm} \log P_{lm} \\ &= \sum_l (2l+1) \left\{ \log \frac{1}{g} + \frac{1}{2} \log(2\pi) + \frac{1}{2} \right\} + \frac{1}{2} \sum_l (2l+1) \log \left( \varepsilon^4 G^l_{rr} \right). \end{aligned} \tag{4.73}$$

While it was obvious that we'd get $\mathrm{tr} \log \sqrt{G}$, it was to get all the other factors that we have been so careful. The $\varepsilon^4$ accompanying $G$ comes from the normalisation of the measure, along with the coupling $g$, but we have put it in the second term to make manifest that this expression only has logarithms of dimensionless quantities.

With this expression in mind, we begin calculating the Green's function. The vector potential is quantised, in $A_0 = 0$ and $\vec{\boldsymbol{\nabla}} \cdot \vec{A} = 0$ gauge, as

$$\begin{aligned} A_i = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2k}} \sum_\alpha \left\{ a_{\mathbf{k},\alpha} \epsilon_i^\alpha(\mathbf{k}) e^{-i(kt - \mathbf{k}\cdot\mathbf{x})} + a^\dagger_{\mathbf{k},\alpha} \epsilon_i^{\alpha *}(\mathbf{k}) e^{i(kt-\mathbf{k}\cdot\mathbf{x})} \right\}, \\ \sum_\alpha \epsilon_i^\alpha(\mathbf{k}) \epsilon_i^{\alpha *}(\mathbf{k}) = \delta_{ij} - \frac{k_i k_j}{k^2}. \end{aligned} \tag{4.74}$$

Because of this, the electric field is

$$E_i = -i \int \frac{d^3k}{(2\pi)^3} \sqrt{\frac{k}{2}} \sum_\alpha \left\{ a_{\mathbf{k},\alpha} \epsilon_i^\alpha(\mathbf{k}) e^{-i(kt-\mathbf{k}\cdot\mathbf{x})} - a^\dagger_{\mathbf{k},\alpha} \epsilon_i^{\alpha *}(\mathbf{k}) e^{i(kt-\mathbf{k}\cdot\mathbf{x})} \right\} \tag{4.75}$$

and the Green's function is

$$\langle E_i(\mathbf{x}) E_j(\mathbf{y}) \rangle = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} k \left( \delta_{ij} - \frac{k_i k_j}{k^2} \right) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}. \tag{4.76}$$

The classical contribution is given by the two-point function with both points on the sphere. However, naively choosing both points on the sphere gives unphysical divergences, including the monopole term not vanishing as it should because of Gauss' law. So, we regularise by smoothing out the electric field a little in the radial direction; we choose the two points to be on spheres of radii $R_1 = R$ and $R_2 = R(1+\Delta)$ respectively. Since, we want the two spheres to be coincident, we take the spacing between the spheres to be much smaller than the lattice scale, $R\Delta \ll \varepsilon$.

We can now proceed to calculate this regularised Green's function. Defining $\xi = \mathbf{x} - \mathbf{y}$, we rewrite the Green's function as

$$G_{rr}(\xi) = \frac{1}{2(2\pi)^2} \int dk d(\cos\theta) \left( k^3 \hat{x} \cdot \hat{y} - k \mathbf{k} \cdot \hat{x} \mathbf{k} \cdot \hat{y} \right) e^{ik\xi\cos\theta}. \tag{4.77}$$

The first term is

$$\begin{aligned}\frac{\hat{x}\cdot\hat{y}}{8\pi^2}\int_0^\infty dkk^3\int_{-1}^1 d(\cos\theta)e^{ik\xi\cos\theta} &= \frac{\hat{x}\cdot\hat{y}}{8\pi^2}\int dkk^3\frac{e^{ik\xi}-e^{-ik\xi}}{ik\xi}\\ &= \frac{\hat{x}\cdot\hat{y}}{4\pi^2\xi}\,\mathrm{Im}\int dkk^2e^{ik\xi}\\ &= \frac{\hat{x}\cdot\hat{y}}{4\pi^2\xi}\left(-\partial_\xi^2\right)\mathrm{Im}\int dke^{ik\xi}\\ &= \frac{\hat{x}\cdot\hat{y}}{4\pi^2\xi}\left(-\partial_\xi^2\right)\mathrm{Im}\,\frac{-1}{i\xi}\\ &= -\frac{\hat{x}\cdot\hat{y}}{2\pi^2\xi^4}.\end{aligned} \tag{4.78}$$

And the second term is

$$\begin{aligned}-\frac{1}{8\pi^2}\int dkd(\cos\theta)k\mathbf{k}\cdot\hat{x}\mathbf{k}\cdot\hat{y}e^{i\mathbf{k}\cdot\xi} &= -\frac{1}{8\pi^2}\left(-\hat{x}\cdot\vec{\boldsymbol{\nabla}}_\xi\hat{y}\cdot\vec{\boldsymbol{\nabla}}_\xi\right)\int dkk\int d(\cos\theta)e^{ik\xi\cos\theta}\\ &= -\frac{1}{8\pi^2}\left(-\hat{x}\cdot\vec{\boldsymbol{\nabla}}_\xi\hat{y}\cdot\vec{\boldsymbol{\nabla}}_\xi\right)\frac{2}{\xi}\,\mathrm{Im}\int dke^{ik\xi}\\ &= \frac{\hat{x}_i\hat{y}_j}{4\pi^2}\partial_i\partial_j\frac{1}{\xi^2}\\ &= \frac{\hat{x}_i\hat{y}_j}{4\pi^2}\frac{-2}{\xi^4}\left(\delta_{ij}-4\frac{\xi_i\xi_j}{\xi^2}\right)\\ &= -\frac{1}{2\pi^2\xi^4}\left(\hat{x}\cdot\hat{y}-4\frac{\xi\cdot\hat{x}\xi\cdot\hat{y}}{\xi^2}\right).\end{aligned} \tag{4.79}$$

Calling the angle between the points $\gamma$, the various inner products above are

$$\begin{aligned}\hat{x}\cdot\hat{y} &= \cos\gamma\\ \xi\cdot\hat{x} &= R_1-R_2\cos\gamma\\ \xi\cdot\hat{y} &= R_1\cos\gamma-R_2\\ \xi^2 &= R_1^2+R_2^2-2R_1R_2\cos\gamma.\end{aligned} \tag{4.80}$$

So, the Green's function is

$$G_{rr} = -\frac{1}{\pi^2(R_1^2+R_2^2)^2}\frac{\alpha-\cos\gamma}{(1-\alpha\cos\gamma)^3},\quad \alpha=\frac{2R_1R_2}{R_1^2+R_2^2}=1-\frac{\Delta^2}{2}. \tag{4.81}$$

To diagonalise it, we will expand this in a basis of Legendre functions and use the relation

$$P_l(\cos\gamma)=\frac{4\pi}{2l+1}\sum_m Y_l^m(\theta_1,\phi_1)Y_l^{m*}(\theta_2,\phi_2). \tag{4.82}$$

To expand it in terms of spherical harmonics, we first expand out the denominator to write it as a power series in $\alpha$,

$$\begin{aligned} G_{rr} &= -\frac{1}{\pi^2(R_1^2+R_2^2)^2}(\alpha-\cos\gamma)\sum_{n=0}^{\infty}\frac{(n+1)(n+2)}{2}\alpha^n\cos^n\theta \\ &= \frac{1}{\pi^2(R_1^2+R_2^2)^2}\sum_{n=0}^{\infty}\frac{n+1}{2}\left[\frac{1-\alpha^2}{\alpha}n-2\alpha\right]\alpha^n\cos^n\gamma, \end{aligned} \tag{4.83}$$

and then use the relations

$$\begin{aligned} t^{2n} &= \sum_{k=0}^{\infty}(4k+1)\frac{2n!!}{(2n-2k)!!}\frac{(2n-1)!!}{(2n+2k+1)!!}P_{2k}(t) \\ t^{2n+1} &= \sum_{k=0}^{\infty}(4k+3)\frac{2n!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+3)!!}P_{2k+1}(t). \end{aligned} \tag{4.84}$$

To plug (4.84) into (4.83), we separate the even and odd powers of $\cos\gamma$,

$$\begin{aligned} \pi^2(R_1^2+R_2^2)^2G_{rr} &= \sum_{n=0}^{\infty}\frac{2n+1}{2}\left[\frac{1-\alpha^2}{\alpha}2n-2\alpha\right]\alpha^{2n}\sum_{k=0}^{\infty}(4k+1)\frac{2n!!}{(2n-2k)!!}\frac{(2n-1)!!}{(2n+2k+1)!!}P_{2k}(\cos\gamma) \\ &\quad+\sum_{n=0}^{\infty}\frac{2n+2}{2}\left[\frac{1-\alpha^2}{\alpha}(2n+1)-2\alpha\right]\alpha^{2n+1}\times \\ &\qquad\sum_{k=0}^{\infty}(4k+3)\frac{2n!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+3)!!}P_{2k+1}(\cos\gamma) \\ &= \sum_{k=0}^{\infty}(4k+1)\left\{\sum_{n=0}^{\infty}\left(\frac{1-\alpha^2}{\alpha}n-\alpha\right)\frac{2n!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+1)!!}\right\}P_{2k} \\ &\quad+\sum_{k=0}^{\infty}(4k+3)\left\{\sum_{n=0}^{\infty}\left(\frac{1-\alpha^2}{\alpha}n-\frac{3\alpha^2-1}{2\alpha}\right)\frac{(2n+2)!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+3)!!}\alpha^{2n+1}\right\}P_{2k+1}. \end{aligned} \tag{4.85}$$

Note the change in the arguments of the double factorials from the first line to the second; a factor has been absorbed in each term. Also note that the factors of $4k+1$ and $4k+3$ will exactly cancel those that come from converting the Legendre polynomial into spherical harmonics.

The part of the sum contributing to its divergence is

$$n \sim \frac{1}{1-\alpha^2} = \frac{1}{\Delta^2}. \tag{4.86}$$

This is very good, since the maximum angular momentum allowed is

$$l_{max} \sim \frac{R}{\varepsilon} \ll \frac{1}{\Delta} \ll \frac{1}{\Delta^2}, \tag{4.87}$$

and therefore, as long as we're interested in only the divergent pieces, we can safely work in the regime

$$k \ll n \tag{4.88}$$

and therefore look at terms order by order in a $1/n$ expansion.

To do this expansion, first let us look at the double factorials for the even part,

$$\begin{aligned}\frac{2n!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+1)!!} &= \frac{2n}{2n+2k+1}\frac{2n-2}{2n+2k+1-2}\cdots\frac{2n-2(k-1)}{2n+2k+1-2(k-1)} \\ &= \prod_{r=0}^{k-1}\frac{2n-2r}{2n+2k+1-2r} \\ &= \prod_{r=0}^{k-1}\left(1-\frac{1}{2}\frac{2k+1}{n+k+\frac{1}{2}-r}\right).\end{aligned} \tag{4.89}$$

Similarly, the double factorial in the odd part is

$$\frac{(2n+2)!!}{(2n-2k)!!}\frac{(2n+1)!!}{(2n+2k+3)!!} = \prod_{r=-1}^{k-1}\left(1-\frac{1}{2}\frac{2k+1}{n+k+\frac{1}{2}-r}\right). \tag{4.90}$$

Clearly, in both terms this part goes as $1+1/n+1/n^2+\cdots$ at large $n$. Now, there are two terms multiplying the double factorial, one proportional to $n$ and one proportional to 1. Thus, the even term splits into two terms, one of which goes as $n+1+1/n+\cdots$ and the other of which goes as $1+1/n+1/n^2+\cdots$, and similarly for the odd term. Naively, then, the leading divergence is quadratic. However, the term proportional to $n$ comes with a factor of $1-\alpha^2$, which reduces the power of the divergence by one order because of (4.86), and so the leading divergence is linear.

In fact, the linear divergence actually cancels. The terms corresponding to the leading divergences all come from when the term which is only a product of 1s for every $r$ in (4.89) or (4.90). For the even term, this part is

$$\frac{1-\alpha^2}{\alpha}\sum n\alpha^{2n} - \alpha\sum\alpha^{2n} = \frac{1-\alpha^2}{\alpha}\frac{\alpha^2}{(1-\alpha^2)^2} - \alpha\frac{1}{1-\alpha^2} = 0. \tag{4.91}$$

Similarly, for the odd term, this part is

$$\frac{1-\alpha^2}{\alpha}\sum n\alpha^{2n+1} - \frac{3\alpha^2-1}{2\alpha}\sum\alpha^{2n+1} = \frac{1-\alpha^2}{\alpha}\frac{\alpha^3}{(1-\alpha^2)^2} - \frac{3\alpha^2-1}{2\alpha}\frac{\alpha}{1-\alpha^2} = \frac{1}{2}, \tag{4.92}$$

which, while not 0 per se, is regular.

So, the leading divergence is logarithmic. Only the part that didn't have an $n$ multiplying the double factorials can contribute to the log divergence, since the log divergent piece in

the part with the $n$ vanishes because of the multiplication with $1-\alpha^2$. In the part without the $n$, the first subleading term is the one where exactly one of the factors in the products (4.89) or (4.90) doesn't contribute a 1, resulting in a piece that is overall a $1/n$ piece. For the even term, it is

$$\begin{aligned}\alpha\sum_n\sum_{r=0}^{k-1}\frac{2k+1}{2n+2k+1-2r}\alpha^{2n} &= \frac{\alpha}{2}(2k+1)\sum_{r=0}^{k-1}\sum_n\left(\frac{\alpha^{2n}}{n}+O(1/n^2)\right)\\ &= \frac{\alpha}{4}2k(2k+1)\log\left(\frac{1}{1-\alpha^2}\right)\\ &= \frac{l(l+1)}{4}\log\left(\frac{1}{1-\alpha^2}\right).\end{aligned} \tag{4.93}$$

And for the odd term it is

$$\begin{aligned}\frac{3\alpha^2-1}{2\alpha}\sum_n\sum_{r=-1}^{k-1}\frac{2k+1}{2n+2k+1-2r}\alpha^{2n+1} &= \frac{3\alpha^2-1}{2\alpha}\sum_{r=-1}^{k-1}\sum_n\left(\frac{\alpha^{2n+1}}{n}+O(1/n^2)\right)\\ &= \frac{3\alpha^2-1}{8}(2k+1)(2k+2)\log\left(\frac{1}{1-\alpha^2}\right)\\ &= \frac{l(l+1)}{4}\log\left(\frac{1}{1-\alpha^2}\right).\end{aligned} \tag{4.94}$$

All the rest of the terms won't contribute to the log part and we ignore them. Thus, putting $R_1 = R$ and $R_2 = R(1+\Delta)$, the Green's function to leading order is

$$G^{lm}_{rr} = \frac{1}{\pi R^4}l(l+1)\log\frac{R^2}{\Delta^2}. \tag{4.95}$$

The relevant part of the entropy is $\frac{1}{2}\operatorname{tr}\log\left(-\boldsymbol{\nabla}^2\right)$, where $\boldsymbol{\nabla}^2$ is the Laplacian on the two-dimensional sphere. To evaluate this, we use a heat kernel expansion, see for example [7],

$$\frac{1}{2}\operatorname{tr}\log\left(-\boldsymbol{\nabla}^2\right) = -\frac{1}{2}\int_{\varepsilon^2}^{\infty}\frac{dt}{t}\operatorname{tr}e^{t\boldsymbol{\nabla}^2}. \tag{4.96}$$

For two-dimensional manifolds without boundary, the short-time asymptotic expansion of the heat kernel for the Laplacian is known to be [55]

$$\operatorname{tr}e^{t\boldsymbol{\nabla}^2} \approx \frac{1}{4\pi t}\left\{\operatorname{tr}\mathbb{1}+t\operatorname{tr}\left(\frac{R}{6}\mathbb{1}\right)\right\}, \tag{4.97}$$

where $R$ is not the radius but the Ricci scalar.

First thing to note is that the leading area law contribution is $\sim -\frac{R^2}{\varepsilon^2}$ and thus negative in this regularisation, even though the answer clearly proceeds from a Hamiltonian calculation of the EE. As was mentioned in section 2.4.2.1, this is exactly the negative contribution found by [21], as was first pointed out by [24].

The second term gives a log divergence. Substituting the Ricci scalar in terms of the radius as $2/R^2$ and $\mathrm{tr}\,\mathbb{1} = 4\pi R^2$, we get the log divergent piece to be

$$\frac{1}{6}\log\varepsilon^2 \sim -\frac{1}{6}\log\frac{R^2}{\varepsilon^2}. \tag{4.98}$$

And thus, the full log term in the EE for a free $U(1)$ gauge field is

$$-\frac{31}{45}\log\frac{R}{\varepsilon}, \tag{4.99}$$

and its extractable part is only

$$-\frac{16}{45}\log\frac{R}{\varepsilon}. \tag{4.100}$$

Finally, let us close this section by arguing that, for a general region, the non-extractable piece will be given by $\frac{1}{2}\,\mathrm{tr}\log\left(-\boldsymbol{\nabla}^2_{bd}\right)$. By expanding (4.77) in Bessel functions for the radial part and spherical harmonics for the angular part, we see that the Green's function can be written as

$$G_{rr} = \frac{2}{\pi}\left[\frac{1}{12}\frac{(R^2-R'^2)^2}{2RR'}\sum_{l,m}\int dk k^5 j_l(kR)j_i(kR')Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') \right. \tag{4.101}$$
$$\left. + \frac{1}{\alpha}\sum_{lm}\int dk k^3 j_l(kR)j_l(kR')Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi')\right].$$

This makes sense because $\phi_{klm} = j_l(kr)Y_{lm}(\theta,\phi)$ is an eigenvalue of the scalar Laplacian in 3 dimensions,

$$\nabla^2\phi_{klm} = -k^2\phi_{klm} \tag{4.102}$$
$$\frac{1}{r^2}\partial_r(r^2\partial_r\phi_{klm}) - \frac{l(l+1)}{r^2}\phi_{klm} = -k^2\phi_{klm} \tag{4.103}$$

with $l(l+1)$ being the eigenvalue of the two dimensional Laplacian on $S^2$.

In the limit (4.87), since $l_{max}\sim\frac{R}{\varepsilon}$, the contribution to the sum, (4.101), is dominated by modes with radial momentum bigger than the momentum along the $S^2$ boundary, $k \gg \frac{l}{R}$.

We can use the WKB approximation to understand the behaviour of the modes in this limit. In this approximation a solution goes like,

$$\phi \sim e^{i\pm\int dr\sqrt{E-V}}, \tag{4.104}$$

where $E = k^2r^2$ and the potential term arises from the two dim. Laplacian, $V = l(l+1)/r^2$. The leading term in eq.(4.104) comes from neglecting $V$ and goes like $e^{\pm i(kr+\theta)}$, where $\theta$ is

a phase. The next term comes from expanding the square root $\sqrt{E-V} \simeq \sqrt{E}[1-\frac{1}{2}\frac{V}{E}]$, and is proportion to $l(l+1)$, the eigenvalue of the two-dim Laplacian.

Done more carefully, this gives rise to the standard asymptotic expansion [56],

$$\begin{aligned} j_l(kr) &= \sqrt{\frac{\pi}{2kr}}[H^{(1)}_{l+1/2}(kr) + c.c.] \\ &\sim \{e^{ikr-\theta}[P_{l+\frac{1}{2}}(kr) + iQ_{l+\frac{1}{2}}(kr)] + c.c.\}. \end{aligned} \tag{4.105}$$

The leading behaviour which comes from setting $Q_{l+\frac{1}{2}} = 0, P_{l+\frac{1}{2}} = 1$ cancels out in the two integrals in (4.101). The first non-trivial contribution then comes from keeping the sub-leading term. From arguments given above it follows that its coefficient is therefore proportional to the eigenvalue of the two dim. Laplacian. In addition, it is logarithmically divergent going like,

$$\int \frac{dk}{k} e^{ik(R-R')} \sim \log\left(\frac{R}{\Delta}\right). \tag{4.106}$$

We see that this reproduces the result, (4.95) up to an overall constant, which we did not keep track of.

These arguments make it clear that in the more general case as well, since we are working in the limit where the component of the momentum normal to the boundary is much bigger than the component along the boundary, the final result for the greens function will be given by the two dimensional Laplacian on the boundary, $\boldsymbol{\nabla}^2_{bd}$ multiplied by a logarithmic divergence,

$$G_{rr} \propto \log\left(\frac{R^2}{\Delta^2}\right)(-\boldsymbol{\nabla}^2_{bd}). \tag{4.107}$$

And thus the non-extractable piece in four dimensions, very generally, is (up to less interesting terms and multiplicative constants)

$$S_c = \frac{1}{2}\,\mathrm{tr}\log\left(-\boldsymbol{\nabla}^2_{bd}\right) \tag{4.108}$$

### 4.3.3 The Not-Quite-Topological Piece

In this section, we make the interesting observation that there is a correction to the area law like the one in the toric code, also caused by integrated Gauss' law. However, we also argue that this piece is not topological.

To see where the topological piece of the entropy comes from, consider first a single-variable Gaussian distribution of variance $\sigma$,

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma}}. \tag{4.109}$$

To define its entropy, we discretise $x$ into steps of $\epsilon$. The corresponding probability is

$$P(x) = \frac{\epsilon}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma}} \tag{4.110}$$

and therefore the entropy is

$$S = \log 2\pi + 1 + \log \frac{\sigma}{\epsilon}. \tag{4.111}$$

As long as $\sigma >> \epsilon$, this expression makes perfect sense. But suppose the probability distribution was a delta function. This is the limit where $\sigma \to 0$.

In this limit, the entropy tends to $-\infty$! While this looks like nonsense, it's nonsense that we put in by hand, in the following sense. When we discretise $x$, we are saying that complete information of $x$ is knowing which region of size $\epsilon$ $x$ belongs to, and the entropy roughly measures how much information you gain by learning that information if you already knew that it was picked from the probability distribution $p(x)$. However, when we stipulate that $p(x)$ is a delta function, the probability distribution is localising the value of $x$ a lot more than the knowledge of which $\epsilon$-sized region $x$ belongs to, and therefore exchanging the probability distribution for that information actually means losing a lot of information! This is why the entropy is negative.

A similar situation holds with the monopole contribution to the classical term, since the probability distribution is set by Gauss' law to be $\delta(E_0)$. In fact, even on the lattice the sum of the $L$s leaving the region has to be 0. Thus, the distribution is in the $\sigma \ll \epsilon$ regime , since the distribution has no width even on the lattice. But, this fact also tells us how to fix the negative entropy: the correct probability function for the monopole is not the delta function $\delta(E_0)$ but the Kronecker delta $\delta_{L_0}$, which has 0 entropy – an eminently sensible answer, since the probability gives us complete knowledge of the value.

This has an interesting effect on the entropy. To see it, consider the first term in (4.73): it is an area law contribution proportional to the log of the coupling. The reason it's area law is that we choose the angular momentum cutoff $l_{max}$ by matching number of degrees of freedom with the lattice, that is by choosing it such that

$$\sum_{l=0}^{l_{max}} (2l+1) = \text{const} \frac{R^2}{\varepsilon^2}. \tag{4.112}$$

But, if we replace the monopole part by a Kronecker delta, $l = 0$ doesn't contribute to this term any more! Thus, that term is modified to be

$$\sum_{l=1}^{l_{max}} (2l+1) \log \frac{1}{g} = \left( \text{const} \frac{R^2}{\varepsilon^2} - 1 \right) \log \frac{1}{g}. \tag{4.113}$$

This correction, though having the same effect on the area law as the topological EE *and* having the same source, is not topological in the same sense. While it is robust to gauge-invariant deformations of the Hamiltonian and deformations of the surface, since its existence only depends on integrated Gauss' law being enforced, it is not the only constant contribution, and the other contributions depend sensitively on both the Hamiltonian and the shape of the surface. The reason for it is simply that the theory is massless, and so there is no length-scale/time-scale such the deformations of the shape/Hamiltonian on larger scales than that can't have any effect. This can be seen, for example, in the fact that the non-extractable piece is given by determinant of the massless scalar Laplacian. If it had been given, instead, by the determinant of $\nabla^2 + m^2$, the shape/Hamiltonian on scales longer than $1/m$ wouldn't have mattered since at that scale the eigenvalues are all dominated by the $m^2$ term anyway.

However, the fact that it has the same physical origin as the topological EE can still manifest; in particular, there is a limit in which it does become the topological EE. Consider a theory that, on the lattice, is $\mathbb{Z}_N$, with $N$ so large that it looks effectively like a $U(1)$ theory in the continuum — here, by continuum, we really mean the weak coupling limit.[3] Since, in terms of electric fields, the difference between $\mathbb{Z}_N$ and $U(1)$ is that in the former case the lattice electric field is cut off by $N$, this is the limit in which the size of the fluctuations of the lattice field is much smaller than $N$. The fact that the lattice and continuum electric fields are related by $E_{cont} \sim \left(g/\varepsilon^2\right) E_{latt}$ and the form of the two-point function (4.95) together mean the two-point function of the lattice electric field is

$$\langle L_{lm}^2 \rangle = \frac{f_l}{g^2}. \tag{4.114}$$

This means that the coupling regime where a $\mathbb{Z}_N$ theory can look like a $U(1)$ theory is

$$\frac{1}{g} \ll N. \tag{4.115}$$

Now, imagine making $g$ smaller and smaller. When $gN$ becomes $O(1)$, the lattice electric fields have an almost flat distribution over the entire allowed range $\{0, 1, \cdots N-1\}$, and the state looks like the ground state of the toric code. Also, in this limit,

$$\log \frac{1}{g} \sim \log N, \tag{4.116}$$

and we see that the term we identified above goes over exactly into the topological EE of the toric code. Finally, notice that even in the limit $gN \ll 1$ the piece above continues to be $\log N$, for the simple reason that reducing $g$ beyond $1/N$ just makes the distribution flatter on the $N$ states that are actually there. Thus, tuning $g$ to be smaller than $t/N$ results in a

[3] We thank Djordje Radicevic for pointing us to this argument.

theory whose ground state looks like the toric code, and the term (4.113) that we identified goes over into its topological EE.

# Chapter 5

# Conclusions

The basic problem in dfining entanglement in gauge theories between a region and its complement arises because of the Gauss' law constraint on the boundary., which equates incoming and outgoing electric flux, thus causing a lack of factorisation of the Hilbert space. One the lattice, in temporal gauge, we can define it anyway by embedding the gauge-invariant Hilbert space $\mathcal{H}_{ginv}$ in a larger extended Hilbert space $\mathcal{H}$ which contains states that do not satisfy Gauss' law and therefore does factorise into an inside and an outside Hilbert space. However, the state whose entanglement we are interested in calculating being gauge-invariant causes the reduced density matrix to have a block-diagonal structure,

$$\rho_{in} = \bigoplus_{\mathbf{k}} p_{\mathbf{k}} \rho_{in}^{\mathbf{k}} \otimes \frac{\mathbb{1}_{d_{\mathbf{k}}}}{d_{\mathbf{k}}}, \tag{5.1}$$

where $\mathbf{k}$ denotes the cofiguration of electric flux leaving the boundary, edge mode for short, and the identity piece comes from the fact that the inside and outside representations have to fuse to the identity.

This results in the EE having the form

$$S_E = -\sum_{\mathbf{k}} p_{\mathbf{k}} \log p_{\mathbf{k}} - \sum_{\mathbf{k}} p_{\mathbf{k}} \rho_{in}^{\mathbf{k}} \log \rho_{in}^{\mathbf{k}} + \sum_{\mathbf{k}} p_{\mathbf{k}} \log d_{\mathbf{k}} \equiv S_c + S_q + S_f, \tag{5.2}$$

These three terms are interpreted as follows:

1. The classical term $S_c$: this is the entropy caused by the classical correlation between edge mode measurements on the inside and outside.

2. The quantum or extractable term $S_q$: this is a measure of the actual quantum entanglement between the two sides.

3. The fusion term $S_f$: this comes from the correlation between measurements of non-gauge-invariant combinations of the edge mode operators and so isn't even classical correlation.

We showed that the quantum term is actual entanglement by showing that only that the state can only be reversibly converted using LOCC into $S_q$ Bell pairs. We showed that the fusion term is information-theoretically empty by showing that it doesn't turn up in an algebraic electric centre definition of the EE, which only knows about statistics of gauge-invariant operators. From here, the interpretation of the classical term follows.

We also showed that despite these differing interpretations, it is still the full EHS EE that is intereting in some physical situations. First, the continuum replica trick, and therefore the Ryu-Takayanagi formula in AdS/CFT, calculates the EHS EE. Secondly, in toric code models, all three terms contribute to the topological EE, which has physical information. Finally, in four-dimensional free Maxwell theory, the classical term is important to obtain agreement with the CFT result for EE.

There are stil some open problems, however. One is a direct formulation of one of the definitions in the continuum after gauge-fixing, steps towards which have been made in [31, 34], which would involve thinking about the Hilbert space as formed from entangled large gauge transformation eigenvalues. Secondly, there is the related queston of lifting temporal gauge and considering nicer gauges like the Lorenz gauge, which would require a proper treatment of BRST ghosts; this is work in progress. Third, there is the question about how entanglement behaves under duality transformations; some work has been done [57, 58], and there is also work in progress. Finally, of course, one can try to calculate entanglement in various theories, especially non-Abelian or strongly coupled ones, and try to find physical signatures in the answers.

Further, we can leave the cosy lap of Yang-Mills gauge theory. Similar EHS-like ideas have been used in the study of entanglement in other constrained systems, like Chern-Simons theory [32], string theory [59, 60] and loop quantum gravity [61]. Finally, there is a suggestion in the literature that the area term in the Ryu-Takayanagi formula is actually a fusion term in the quantum gravity entanglement calculation [62, 63].

# Bibliography


[1] S. Ghosh, R. M. Soni and S. P. Trivedi, *On The Entanglement Entropy For Gauge Theories*, *JHEP* **09** (2015) 069, [1501.02593].

[2] R. M. Soni and S. P. Trivedi, *Aspects of Entanglement Entropy for Gauge Theories*, *JHEP* **01** (2016) 136, [1510.07455].

[3] R. M. Soni and S. P. Trivedi, *Entanglement entropy in* $(3+1)$*-d free* $U(1)$ *gauge theory*, *JHEP* **02** (2017) 101, [1608.00353].

[4] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, *Physics* **1** (1964) 195–200.

[5] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, *Proposed experiment to test local hidden variable theories*, *Phys. Rev. Lett.* **23** (1969) 880–884.

[6] E. Witten, *Notes on Some Entanglement Properties of Quantum Field Theory*, 1803.04993.

[7] T. Nishioka, *Entanglement entropy: holography and renormalization group*, 1801.10352.

[8] J. B. Kogut, *An Introduction to Lattice Gauge Theory and Spin Systems*, *Rev. Mod. Phys.* **51** (1979) 659.

[9] J. B. Kogut and L. Susskind, *Hamiltonian Formulation of Wilson's Lattice Gauge Theories*, *Phys. Rev.* **D11** (1975) 395–408.

[10] K. Bardakci, *Classical Equations for Nonabelian Gauge Theories in the Large N Limit*, *Nucl. Phys.* **B193** (1981) 245–256.

[11] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*. Princeton university press, 1994.

[12] T. Tao, "The peter-weyl theorem, and non-abelian fourier analysis on compact groups." https://terrytao.wordpress.com/2011/01/23/the-peter-weyl-theorem-and-non-abelian-fourier-analysis-on-compact-groups/, 2011.

[13] P. Woit, "Lie groups and representations." `http://www.math.columbia.edu/~woit/RepThy/repthynotes3.pdf`, 2007.

[14] P. V. Buividovich and M. I. Polikarpov, *Entanglement entropy in gauge theories and the holographic principle for electric strings*, *Phys. Lett.* **B670** (2008) 141–145, [0806.3376].

[15] W. Donnelly, *Decomposition of entanglement entropy in lattice gauge theory*, *Phys. Rev.* **D85** (2012) 085004, [1109.0036].

[16] S. Aoki, T. Iritani, M. Nozaki, T. Numasawa, N. Shiba and H. Tasaki, *On the definition of entanglement entropy in lattice gauge theories*, *JHEP* **06** (2015) 187, [1502.04267].

[17] H. Casini, M. Huerta and J. A. Rosabal, *Remarks on entanglement entropy for gauge fields*, *Phys. Rev.* **D89** (2014) 085012, [1312.1183].

[18] D. Radicevic, *Notes on Entanglement in Abelian Gauge Theories*, 1404.1391.

[19] H. Casini and M. Huerta, *Entanglement entropy for a Maxwell field: Numerical calculation on a two dimensional lattice*, *Phys. Rev.* **D90** (2014) 105013, [1406.2991].

[20] J. L. Cardy, O. A. Castro-Alvaredo and B. Doyon, *Form factors of branch-point twist fields in quantum integrable models and entanglement entropy*, *J. Statist. Phys.* **130** (2008) 129–168, [0706.3384].

[21] D. N. Kabat, *Black hole entropy and entropy of entanglement*, *Nucl. Phys.* **B453** (1995) 281–299, [hep-th/9503016].

[22] D. Iellici and V. Moretti, *Kabat's surface terms in the zeta function approach*, in *General relativity and gravitational physics. Proceedings, 12th Italian Conference, Rome, Italy, September 23-27, 1996*, pp. 317–321, 1996. hep-th/9703088.

[23] W. Donnelly and A. C. Wall, *Do gauge fields really contribute negatively to black hole entropy?*, *Phys. Rev.* **D86** (2012) 064042, [1206.5831].

[24] W. Donnelly and A. C. Wall, *Entanglement entropy of electromagnetic edge modes*, *Phys. Rev. Lett.* **114** (2015) 111603, [1412.1895].

[25] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602, [hep-th/0603001].

[26] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045, [hep-th/0605073].

[27] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090, [1304.4926].

[28] A. P. Balachandran, L. Chandar and A. Momen, *Edge states and entanglement entropy*, *Int. J. Mod. Phys.* **A12** (1997) 625–642, [hep-th/9512047].

[29] M. Pretko and T. Senthil, *Entanglement entropy of $U(1)$ quantum spin liquids*, *Phys. Rev.* **B94** (2016) 125112, [1510.03863].

[30] M. Pretko, *On the Entanglement Entropy of Maxwell Theory: A Condensed Matter Perspective*, 1801.01158.

[31] W. Donnelly and L. Freidel, *Local subsystems in gauge theory and gravity*, *JHEP* **09** (2016) 102, [1601.04744].

[32] G. Wong, *A note on entanglement edge modes in Chern Simons theory*, 1706.04666.

[33] S. Elitzur, G. W. Moore, A. Schwimmer and N. Seiberg, *Remarks on the Canonical Quantization of the Chern-Simons-Witten Theory*, *Nucl. Phys.* **B326** (1989) 108–134.

[34] A. Blommaert, T. G. Mertens, H. Verschelde and V. I. Zakharov, *Edge State Quantization: Vector Fields in Rindler*, 1801.09910.

[35] W. Donnelly and A. C. Wall, *Geometric entropy and edge modes of the electromagnetic field*, *Phys. Rev.* **D94** (2016) 104053, [1506.05792].

[36] K. Van Acoleyen, N. Bultinck, J. Haegeman, M. Marien, V. B. Scholz and F. Verstraete, *The entanglement of distillation for gauge theories*, *Phys. Rev. Lett.* **117** (2016) 131602, [1511.04369].

[37] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, *Concentrating partial entanglement by local operations*, *Phys. Rev.* **A53** (1996) 2046–2052, [quant-ph/9511030].

[38] J. Preskill, *Lecture notes for physics 229: Quantum information and computation*, *California Institute of Technology* **16** (1998) .

[39] M. A. Nielsen and I. Chuang, *Quantum computation*, *Quantum Information. Cambridge University Press, Cambridge* (2000) .

[40] M. M. Wilde, *Quantum information theory*. Cambridge University Press, 2013.

[41] N. Schuch, F. Verstraete and J. I. Cirac, *Nonlocal resources in the presence of superselection rules*, *Physical review letters* **92** (2004) 087904.

[42] N. Schuch, F. Verstraete and J. I. Cirac, *Quantum entanglement theory in the presence of superselection rules*, *Physical Review A* **70** (2004) 042310.

[43] A. Yu. Kitaev, *Fault tolerant quantum computation by anyons*, *Annals Phys.* **303** (2003) 2–30, [quant-ph/9707021].

[44] A. Kitaev and J. Preskill, *Topological entanglement entropy*, *Phys. Rev. Lett.* **96** (2006) 110404, [hep-th/0510092].

[45] M. Levin and X.-G. Wen, *Detecting Topological Order in a Ground State Wave Function*, *Phys. Rev. Lett.* **96** (2006) 110405.

[46] T. Grover, A. M. Turner and A. Vishwanath, *Entanglement Entropy of Gapped Phases and Topological Order in Three dimensions*, *Phys. Rev.* **B84** (2011) 195120, [1108.4038].

[47] S. N. Solodukhin, *Entanglement entropy, conformal invariance and extrinsic geometry*, *Phys. Lett.* **B665** (2008) 305–309, [0802.3117].

[48] H. Casini, M. Huerta and R. C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **05** (2011) 036, [1102.0440].

[49] J. S. Dowker, *Entanglement entropy for even spheres*, 1009.3854.

[50] C. Eling, Y. Oz and S. Theisen, *Entanglement and Thermal Entropy of Gauge Fields*, *JHEP* **11** (2013) 019, [1308.4964].

[51] K.-W. Huang, *Central Charge and Entangled Gauge Fields*, *Phys. Rev.* **D92** (2015) 025010, [1412.2730].

[52] H. Casini and M. Huerta, *Entanglement entropy of a Maxwell field on the sphere*, *Phys. Rev.* **D93** (2016) 105031, [1512.06182].

[53] F. Zuo, *A note on electromagnetic edge modes*, 1601.06910.

[54] D. V. Fursaev, A. Patrushev and S. N. Solodukhin, *Distributional Geometry of Squashed Cones*, *Phys. Rev.* **D88** (2013) 044054, [1306.4000].

[55] L. Parker and D. Toms, *Quantum field theory in curved spacetime: quantized fields and gravity*. Cambridge university press, 2009.

[56] G. B. Arfken, H. J. Weber and F. E. Harris, *Mathematical methods for physicist a comprehensive guide*, *Academic Press* **20** (2013) 13.

[57] . Radievi, *Entanglement Entropy and Duality*, *JHEP* **11** (2016) 130, [1605.09396].

[58] W. Donnelly, B. Michel and A. Wall, *Electromagnetic Duality and Entanglement Anomalies*, *Phys. Rev.* **D96** (2017) 045008, [1611.05920].

[59] W. Donnelly and G. Wong, *Entanglement branes in a two-dimensional string theory*, *JHEP* **09** (2017) 097, [1610.01719].

[60] V. Balasubramanian and O. Parrikar, *Remarks on entanglement entropy in string theory*, *Phys. Rev.* **D97** (2018) 066025, [1801.03517].

[61] W. Donnelly, *Entanglement entropy in loop quantum gravity*, *Phys. Rev.* **D77** (2008) 104006, [0802.0880].

[62] D. Harlow, *The RyuTakayanagi Formula from Quantum Error Correction*, *Commun. Math. Phys.* **354** (2017) 865–912, [1607.03901].

[63] J. Lin, *Ryu-Takayanagi Area as an Entanglement Edge Term*, 1704.07763.