Black Holes, Entropy, and Holographic Encoding

Daniel Harlow

Massachusetts Institute of Technology

August 2, 2017

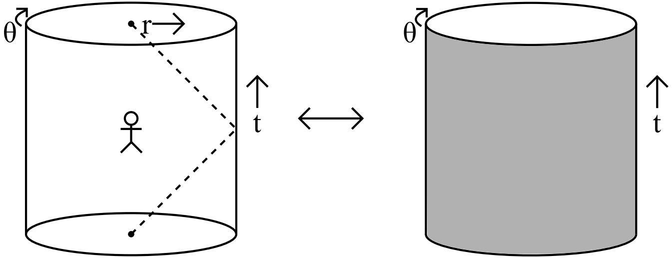
Our Favorite Theory of Quantum Gravity

- The Anti de Sitter/Conformal Field Theory (AdS/CFT) correspondence, our best-understood toy theory of quantum gravity, is now twenty years old!
- Until recently, most of the follow-up work has used classical gravity on the bulk side to learn about strongly-coupled QFT on the boundary side.
- This has been a reasonably successful approach (strongly-coupled plasmas, new understanding of transport in CMT, hydrodynamic anomalies, etc), but it is unlikely to tell us anything interesting about the deep puzzles of quantum gravity.
- To phrase that differently, we will never be able explain AdS/CFT to computer scientists this way!

- Now one might say that is *their* problem, after all for the most part computer scientists don't *know* anything about gravity, relativity, or quantum field theory: how can they understand 21st century physics without first understanding 20th century physics?
- An interesting realization for many of us physicists over the last few years (or earlier for a few wise sages like John) is that in fact there is something to be gained by learning how to translate "our" problems into "their" language.
- The problem I will study today, characterizing the holographic map that tells us which states and operators in the bulk AdS get mapped to which states and operators in the boundary CFT, turns out to be just such a problem.
- I'll describe three features of this map which are quite surprising from the traditional boundary QFT point of view, but which we'll see are quite natural once we learn to re-interpret the holographic map as encoding a quantum error-correcting code. Almheiri/Dong/Harlow 14, Dong/Harlow/Wall 16, Harlow 16, Harlow/Ooguri 17

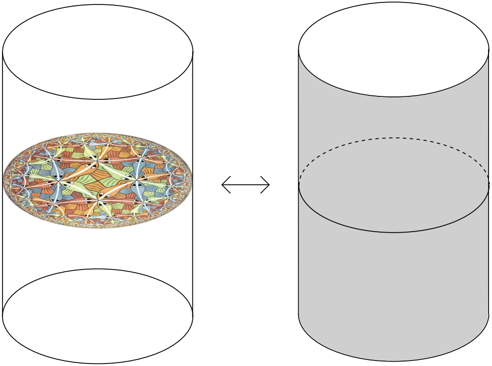
AdS/CFT Review

AdS/CFT says that quantum gravity in asymptotically AdS space is equivalent to conformal field theory on its boundary:



$G \neq 0$ $\longleftrightarrow$ $G = 0$

This correspondence is a quantum correspondence:



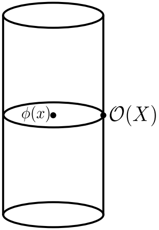
- $|\psi_{bulk}\rangle \longleftrightarrow |\psi_{boundary}\rangle$
- $H, J, \ldots \longleftrightarrow H, J, \ldots$
- $\lim_{r\to\infty} r^{\Delta}\phi(r, t, \Omega) \longleftrightarrow \mathcal{O}(t, \Omega)$.
- Vacuum perturbations $\longleftrightarrow$ low-energy states
- Black holes $\longleftrightarrow$ high-energy states

Radial Commutativity

In quantum field theory, causality is enforced by locality:

$$[\mathcal{O}(X), \mathcal{O}(Y)] = 0 \qquad (X - Y)^2 > 0.$$

We can consider this in the bulk as well:



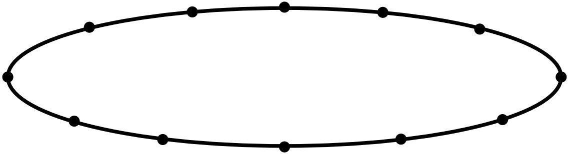
x and X are spacelike-separated in the bulk, so we might expect that

$$[\phi(x), \mathcal{O}(X)] = 0.$$

This however is impossible in the boundary quantum field theory!

The problem is that in a quantum field theory, such as the boundary CFT, any operator that commutes with all local operators at a fixed time must be trivial. Streater/Wightman

For example consider a chain of Pauli spins:



The set of products of Pauli operators, eg

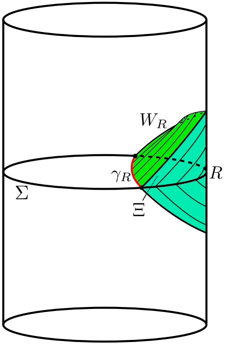
$$Z_1 X_4 Y_7 \ldots,$$

gives a basis for all operators, so an operator which commutes with all of the individual Pauli operators must be proportional to the identity.

This is a basic expression of the local structure of the Hilbert space in a QFT. But then how can we get an extra dimension to emerge?

Ryu-Takayanagi Formula

Given a boundary subregion R we can define the (H)RT surface γ_R and entanglement wedge W_R such that in "good" states ρ we have:



$$S(\rho_R) = \mathrm{Tr}\left(\rho \mathcal{L}_R\right) + S\left(\rho_{W_R}\right)$$

$$\mathcal{L}_R = \frac{\mathrm{Area}\left(\gamma_R\right)}{4G} + \dots$$

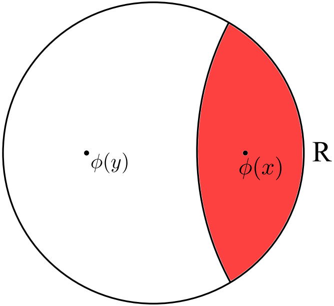
In other words, the entropy of a boundary region R is equal to the expectation value of the area of γ_R in Planck units plus the bulk entropy in W_R!

Although checked many times, this formula has some surprising features as well:

- Given that states are dual to states, why isn't entropy just dual to entropy?
- To the extent that the first term is dominant, how is this consistent with the linearity of QM? Papadodimas/Raju, Almheiri/Dong/Swingle
- What are the "good" states?

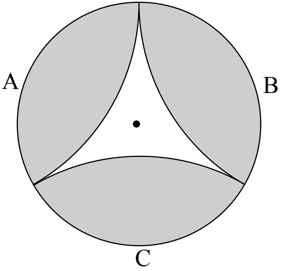
Subregion duality

Evidence has gradually accumulated for the following proposal: given a bulk operator ϕ with in W_R, there exists a CFT operator ϕ_R with support only in R which represents ϕ in the CFT. Let's think about it from above:



The operator $\phi(x)$ can be represented on R, but the operator $\phi(y)$ cannot.

This again leads to some surprising situations:



The operator in the center has no representation on A, B, or C, but it does have a representation either on AB, AC, or BC!
Where is the information?

- Each of these puzzles arises from trying to understand how the bulk physics can be consistent with spatial locality in the CFT: how can the same theory be local with two different spacetime dimensionalities?
- The key point is that the picture we have been using of the bulk, based in semiclassical effective field theory, is *not* valid throughout the Hilbert space of the CFT for any particular bulk observable.
- If we go to high enough energy states, any given observable will always end up far behind the horizon of a black hole, and become operationally inaccessible.
- We will now see that by asking for radial commutativity, the RT formula, and subregion duality to hold only in a *subspace* of the CFT Hilbert space, we can naturally realize all of them.
- Since this is a short talk, we will mostly do this in the context of a simple model of holography: the three-qutrit code.

Three Qutrits

The three qutrit code embeds a single "logical" qutrit into three "physical" qutrits as follows: Cleve/Gottesman/Lo

$$
\begin{aligned}
|\widetilde{0}\rangle &= \frac{1}{\sqrt{3}} \left(|000\rangle + |111\rangle + |222\rangle\right) \\
|\widetilde{1}\rangle &= \frac{1}{\sqrt{3}} \left(|012\rangle + |120\rangle + |201\rangle\right) \\
|\widetilde{2}\rangle &= \frac{1}{\sqrt{3}} \left(|021\rangle + |102\rangle + |210\rangle\right).
\end{aligned}
$$

This subspace is symmetric under cyclic permutations of the physical qutrits, and there is clearly a lot of entanglement in all three states.

One way of understanding this code is to note that there is a unitary on the first two physical qutrits, U_{12}, such that

$$|\widetilde{i}\rangle = U_{12}^{\dagger}\left(|i\rangle_1 \otimes |\chi\rangle_{23}\right),$$

where

$$|\chi\rangle \equiv \frac{1}{\sqrt{3}}\left(|00\rangle + |11\rangle + |22\rangle\right).$$

Explicitly

$$\begin{array}{lll} |00\rangle \to |00\rangle & |11\rangle \to |01\rangle & |22\rangle \to |02\rangle \\ |01\rangle \to |12\rangle & |12\rangle \to |10\rangle & |20\rangle \to |11\rangle \\ |02\rangle \to |21\rangle & |10\rangle \to |22\rangle & |21\rangle \to |20\rangle \end{array}.$$

This means that we can recover any logical state $|\widetilde{\psi}\rangle$ from just the first two qutrits:

$$U_{12}|\widetilde{\psi}\rangle = |\psi\rangle_1 \otimes |\chi\rangle_{23}.$$

By symmetry there is also a U_{13} and U_{23}.

In additional to logical states, we also have logical operators

$$\widetilde{O}|\widetilde{i}\rangle = \sum_j (O)_{ji}\, |\widetilde{j}\rangle.$$

In general we expect these to act nontrivially on all three qutrits, but given our U_{12} we can do something clever: if we define

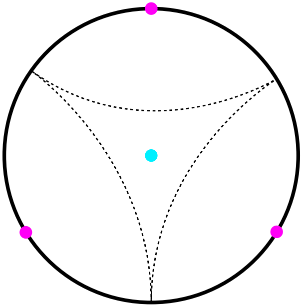
$$O_{12} \equiv U_{12}^\dagger O_1 U_{12},$$

then we have

$$\begin{aligned} O_{12}|\widetilde{\psi}\rangle &= \widetilde{O}|\widetilde{\psi}\rangle \\ O_{12}^\dagger|\widetilde{\psi}\rangle &= \widetilde{O}^\dagger|\widetilde{\psi}\rangle. \end{aligned}$$

Now using the symmetry, we see that any logical operator can be represented on any two of the physical qutrits. This is reminiscent of subregion duality!

Let's now make the analogy precise:



- Three "physical" qutrits are local CFT degrees of freedom on the boundary
- One "logical" qutrit is a field in the center of the bulk
- The correctability we just discussed ensures that subregion duality holds provided we say that our bulk point lies in the entanglement wedge of any two boundary qutrits.

We can now see radial commutativity:

- Consider

$$\langle\widetilde{\psi}|[\widetilde{O}, X_3]|\widetilde{\phi}\rangle,$$

  where X_3 is some operator on the third qutrit and $|\widetilde{\phi}\rangle$, $|\widetilde{\psi}\rangle$ are arbitrary states in the code subspace.
- Since $\widetilde{O}$ always acts either to the left on a state in the code subspace, we can replace it by O_{12}. But then the commutator is zero! This would have worked for X_1 or X_2 as well, so we see that on the code subspace $\widetilde{O}$ commutes with all "local" operators.
- It is because we are working only in the code subspace that we are able to circumvent the algebraic puzzle we discussed before.

Finally we can study the RT formula in this model.
We first note that any logical mixed state is of the form

$$\widetilde{\rho} = U_{12}^{\dagger}\Big(\rho_1 \otimes |\chi\rangle\langle\chi|_{23}\Big)U_{12}$$

We may then compute

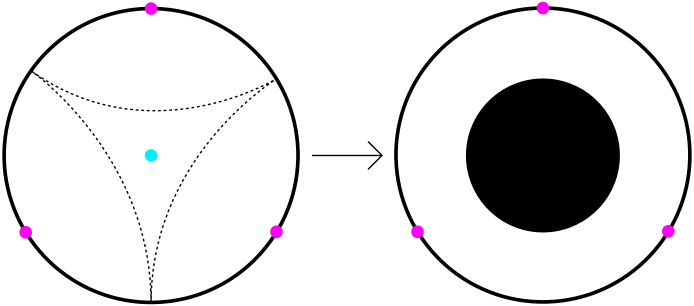
$$\begin{aligned} S\left(\widetilde{\rho}_3\right) &= \log 3 \\ S\left(\widetilde{\rho}_{12}\right) &= \log 3 + S\left(\widetilde{\rho}\right). \end{aligned}$$

If we define

$$\mathcal{L}_R \equiv \log 3,$$

then we see that the RT formula indeed holds!
Note the "area term" comes from the entanglement in $|\chi\rangle$, which we needed to be nonzero to have a good code.

But what about the rest of the states? There is a a 24-dimensional subspace orthogonal to the code subspace, what about bulk locality in those states?
This is where gravity comes to the rescue: these states are the microstates of a black hole that has swallowed our bulk point!

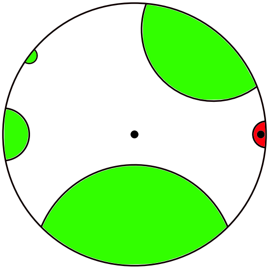


Conclusion

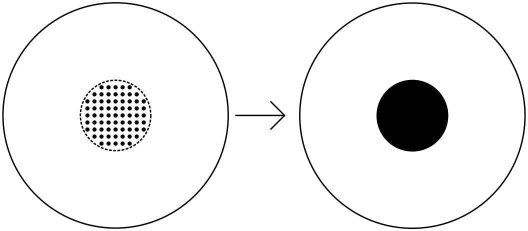
One can pursue this further, using the general theory of error correcting codes and the physics of the bulk to learn more about what kind of error correcting code AdS/CFT realizes.

I do not have time to discuss this in detail today, but I'll mention some of the highlights.

- The code in general has the property that quantum information located further from the boundary is better protected in the CFT:

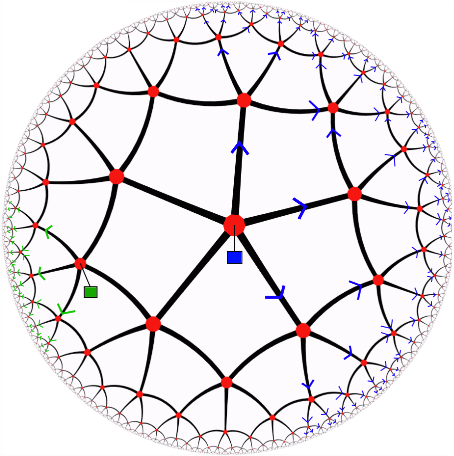


- We can also consider the question of how large the code subspace can be; on how many states can we realize the bulk algebra of operators? It turns out that the answer, which is given by a general theorem in quantum error correction, can be matched directly onto the bulk regime where we make a black hole:



- There is a nice generalization of the three-qutrit code to a larger code with a volume's worth of degrees of freedom in the bulk:

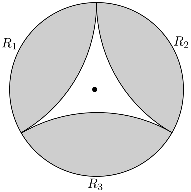
Pastawski/Yoshida/Harlow/Preskill, Hayden/Nezami/Qi/Thomas/Walter/Yang



New codes!

- One can prove a general algebraic theorem that subregion duality and the Ryu-Takayanagi formula are mathematically equivalent: any quantum error correcting code has a version of the RT formula! Harlow
- This theorem clarifies the regime of validity of the RT formula: it explains how the formula can hold in superpositions of geometries, and gives some guidance on to what extent it holds in the presence of black holes.
- This has led to a new focus on "operator algebra quantum error correction", see in particular Cotler/Hayden/Salton/Swingle/Walter,Pastawski/Preskill.
- Finally I'll close by sketching a new proof using subregion duality of the old conjecture that there are no global symmetries in quantum gravity, for simplicity specializing to continuous symmetries although our proof works in general. Harlow/Ooguri

Say there were a bulk global symmetry. By definition, there is some object in the bulk which is charged under it. Consider the algebra of an operator that creates this object in the center of the bulk with the symmetry operator $U(g)$:



By Noether's theorem we have

$$U(g) = U(g, R_1)U(g, R_2)U(g, R_3).$$

Since each $U(g, R_i)$ is localized in the boundary, it can only affect the bulk within the entanglement wedge of R_i. Since our charged operator is not in the entanglement wedge for any R_i, it must commute with all the $U(g, R_i)$. But then it must also commute with $U(g)$, which contradicts the assumption that the object is charged!

Clearly there is much that a coding perspective on holography still has to teach us! More examples?

- New QFT theorems? (QNEC)
- New GR theorems? Myers/Ooguri/et al
- More new codes? Brehm/Richter

Thanks!