Comments on replica wormholes

Douglas Stanford

Stanford University

July 3, 2020

Based on work with Geoff Penington, Steve Shenker and Zhenbin Yang. And on [DS, work in progress].

1. How can acting in the radiation create a particle behind the horizon?

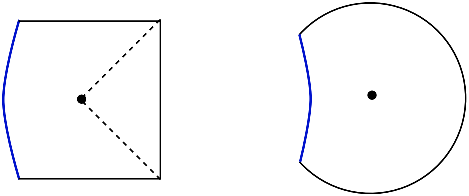
2. To what extent do the replica wormholes involved in the Page curve require an ensemble interpretation?

Part I: how to create a particle behind the horizon.

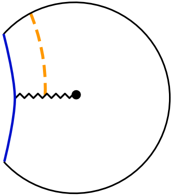
Main reference: [Penington/Shenker/DS/Yang]

See also [Cotler/Hayden/Penington/Salton/Swingle/Walter] where the Petz map was introduced to holographers. And see [Yiming Chen] for another method based on modular flow.

We will continue with the JT gravity + EOW brane model from Geoff's talk. Here is the basic BH solution in Lorentzian and Euclidean signature:



We will work in Euclidean signature. A creation operator for a particle behind the horizon should produce a spacetime like this:

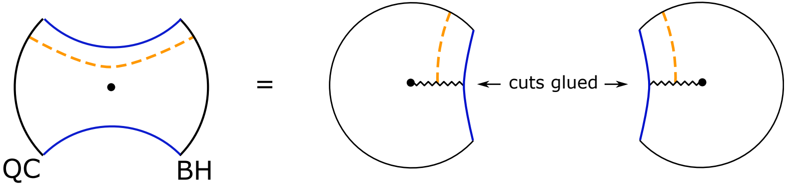


The branch cut separates versions of the interior with and without the particle.

Q: How does this happen?

Q: How does this happen?

Basic picture: one acts with a special operator in the radiation, and the dominant contribution to its matrix element comes from a wormhole geometry



Focusing on the BH part (rightmost), the spacetime looks like what we wanted.

Q2: What operator do you act with on the radiation?
Q3: How can a wormhole connect to a quantum computer?
Q4: That equation is true topologically but not geometrically.

Q2: What operator do you act with on the radiation?

Here's an example. Let

$$|0\rangle = |\text{starting entangled state of BH and radiation}\rangle$$
$$|1\rangle = |\text{same, but with particle behind horizon}\rangle$$

The operator

$$|1\rangle\langle 0| \qquad \text{acting on} \qquad B \otimes R$$

would turn $|0\rangle$ into $|1\rangle$. But we want an operator that acts only on R. The most naive possible guess would be

$$\mathcal{O}_R \overset{?}{\propto} \mathrm{Tr}_B\left(|1\rangle\langle 0|\right).$$

This sometimes works, but a choice that works better is a fancier version defined using something called the "Petz map:"

$$\mathcal{O}_R = \sigma^{-1/2}\mathrm{Tr}_B\left(|1\rangle\langle 0|\right)\sigma^{-1/2}$$

where

$$\sigma = \mathrm{Tr}_B\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right).$$

When you compute matrix elements of this operator after the Page time, a wormhole geometry dominates. Very similar to the discussion of the entropy from Geoff's talk.

Note, the operator is highly state-dependent, [Papadodimas/Raju, Verlinde/Verlinde]

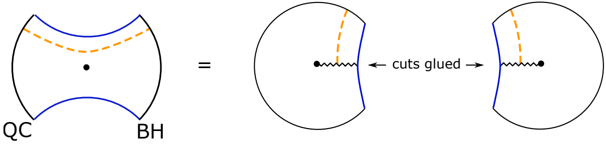
Q3: How can a wormhole connect to a quantum computer?

The operator

$$\mathcal{O}_R = \sigma^{-1/2} \mathrm{Tr}_B \left(|1\rangle\langle 0| \right) \sigma^{-1/2}$$

involves a partial trace over an operator that acts on the BH.

To implement this, the quantum computer might simulate a second copy of the black hole. This can connect via a wormhole to the physical black hole.

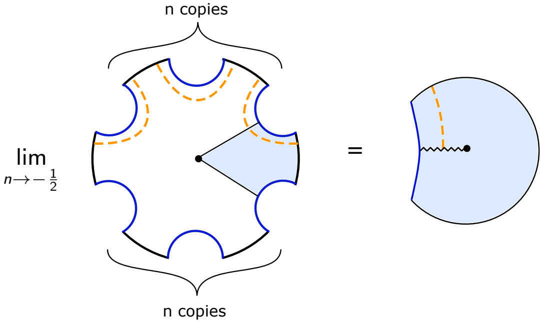


Q4: That equation is true topologically but not geometrically.

The factors of $\sigma^{-1/2}$ in the Petz map fix this. Use a replica trick

$$\mathcal{O}_R = \lim_{n \to -\frac{1}{2}} \sigma^n \mathrm{Tr}_B(|1\rangle\langle 0|)\sigma^n$$

Then the actual geometrical statement is



Part II: gravity and ensembles

See also [Bousso/Tomasevic/Wildenhain] for another perspective

It was a surprise that semiclassical gravity can compute the Page curve [Penington, Almheiri/Engelhardt/Marolf/Maxfield] and also average values of late-time correlation functions [Saad/Shenker/DS, Blommaert/Mertens/Verschelde, Saad].

Previously, it was thought that to compute these things, we would need a bulk description with manifest microstates. In retrospect, we may have been misunderstanding what semiclassical gravity describes. Possible answer:

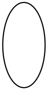
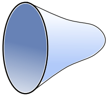
Gravity describes an ensemble average of quantum systems, including finite entropy effects (over an ensemble that may or may not be well-defined).

There are some simple "pure gravity" theories where souped-up semiclassical gravity is all there is. Such theories are fundamentally dual to ensembles.

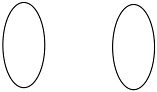
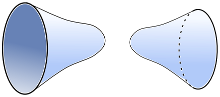
But in string theory, semiclassical gravity is **not** all there is, and we have more work to do. [Mathur/Warner/Bena/Turton/...] [Blommaert/Mertens/Verschelde] [Marolf/Maxfield] [Saad/Shenker/Yao]

Wormholes and ensembles

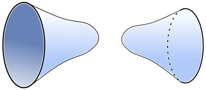
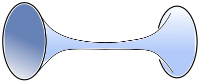
In AdS_2/CFT_1, time contours of the boundary theory become boundary conditions, to be filled in by the bulk theory, e.g.

$Z(\beta) \underset{\text{AdS/CFT dict.}}{=}$  $\underset{\text{fill in with gravity}}{\longrightarrow}$  (1)

Consider the quantity $Z(\beta)^2$. We might expect to get

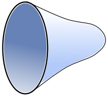
$Z(\beta)^2 \;=$  $\overset{?}{\longrightarrow}$ 

However, there can also be a contribution from a connected geometry

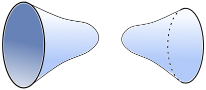
$Z(\beta)^2 \;\longrightarrow$  $+$  (2)

There is a flat contradiction between

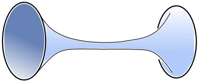
$$Z(\beta) =$$



$$Z(\beta)^2 =$$

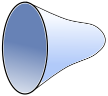


$$+$$

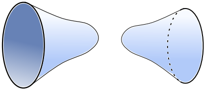


One resolution: interpret gravity answer as ensemble average

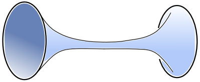
$$\langle Z(\beta)\rangle_{\text{ensemble}} =$$



$$\langle Z(\beta)^2\rangle_{\text{ensemble}} =$$

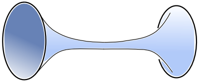


$$+$$



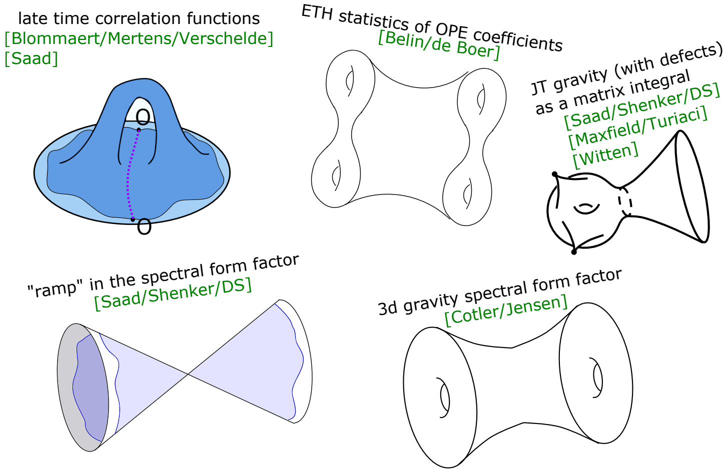
The variance is given by the wormhole

$$\langle Z(\beta)^2\rangle - \langle Z(\beta)\rangle\langle Z(\beta)\rangle =$$



[Coleman] [Giddings/Strominger] [Maldacena/Maoz] [Arkani-Hamed/Polchinski/Orgera]

In the last two years, we have seen something positive come out of this: wormholes give the right answer for the **ensemble average** of the “noise” in various quantities due to the underlying discreteness of quantum mechanics.



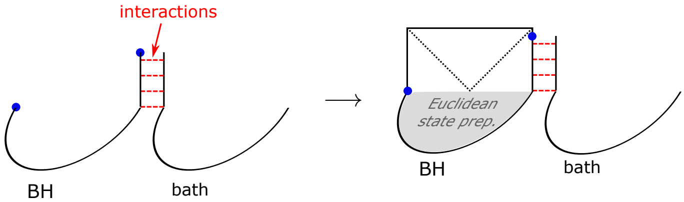
These things all give answers that require an ensemble interpretation.

Do the replica wormholes for the Page curve need an ensemble interpretation?

Replica wormholes for the Page curve

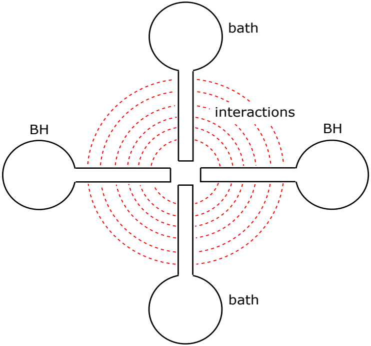
Consider the state of an evolving black hole, coupled to a bath:

$|\Psi\rangle =$

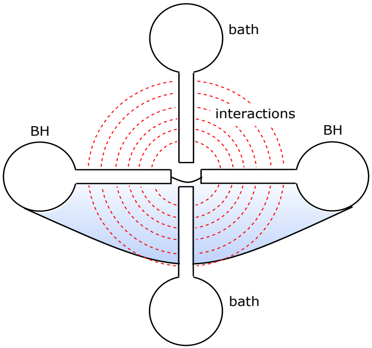


The time contour for the Renyi entropy of the blue dots is

$\mathrm{Tr}(\rho_{\bullet}^2) =$



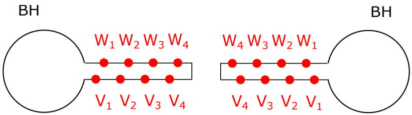
At late times, the dominant way of filling in this geometry is a wormhole between the two copies of the black hole,

$\mathrm{Tr}(\rho_{\bullet}^2) \longrightarrow$ 

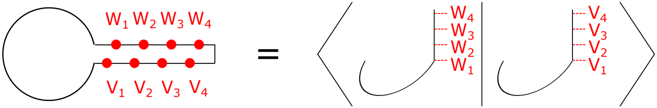
[Almheiri/Hartman/Maldacena/Shaghoulian/Tajdini][Penington/Shenker/DS/Yang]

The two replicas of the BH are connected. That's OK, though, because the boundaries themselves are connected. We don't yet require an ensemble.

But what if we replace the interactions with the bath by some fixed operator insertions? **(a)** what does this quantity mean? **(b)** is the wormhole still there?



(a) Each time contour computes an inner product of states of the BH left behind after two different specific histories of Hawking radiation:



$= \langle W|V\rangle$

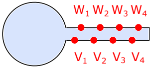
In Hawking's calculation, if the histories W, V are different, the leftover states will be close to orthogonal, $|\langle W|V\rangle|^2 \lesssim e^{-\#\text{different op.s}} \to 0$

At late time, this is too many orthogonal or nearly-orthogonal states. For a real quantum system, $\langle W|V\rangle$ will be erratic, with typical value $|\langle W|V\rangle|^2 \gtrsim e^{-S}$.

This e^{-S} value is necessary to get the Page curve. (Roughly, $\mathrm{Tr}(\rho_{\bullet}^2)$ is related to a sum over V, W of $|\langle W|V\rangle|^2$.)

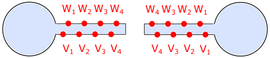
(b) The wormhole is still there. Gravity gives the following answers [in progress]:

$\langle W|V\rangle$ =

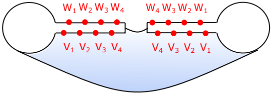


$= \delta_{W,V}$

$|\langle W|V\rangle|^2$ =



+



$$= \delta_{W,V} + e^{-S}.$$

The e^{-S} term from the wormhole is consistent with the Page curve. But the two answers together imply that we need an ensemble interpretation

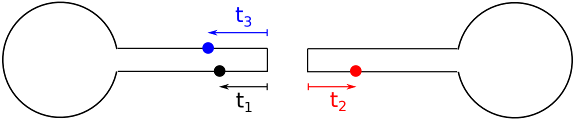
$$\Big\langle \langle W|V\rangle \Big\rangle_{\text{ensemble}} = \delta_{W,V}$$
$$\Big\langle |\langle W|V\rangle|^2 \Big\rangle_{\text{ensemble}} = \delta_{W,V} + e^{-S}.$$

So the inner products underlying the computation of $\mathrm{Tr}(\rho_{\bullet}^2)$ are given by the "same wormhole," but now we do need an ensemble interpretation.

Classical solution for the wormhole in JT gravity coupled to matter

In the microcanonical ensemble the wormhole is a classical solution.

Can get a sense for the geometry by studying the distances through the bulk between boundary points



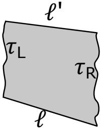
The points 1 and 2 are close together, like on opposite sides of a thermofield double

$$e^{D_{12}} \propto \cosh^2(\tfrac{\pi}{\beta_E}(t_1 - t_2)),$$

while points 1 and 3 are far apart, like on opposite sides of an "OTOC time fold"

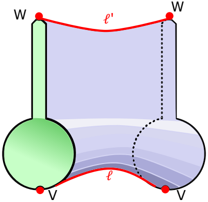
$$e^{D_{13}} \propto \left[\mathrm{i}\sinh(\tfrac{\pi}{\beta_E}(t_1 - t_3)) + G_N \sinh(\tfrac{\pi}{\beta_E} t_1)\sinh(\tfrac{\pi}{\beta_E} t_3)\right]^2 .$$

Computation using exact quantization of JT gravity [Kitaev/Suh][Yang][Saad]

$P(\tau_L, \tau_R; \ell, \ell') =$ 

$= \int_0^\infty \mathrm{d}s \sinh(2\pi s) e^{-s^2(\tau_L+\tau_R)} K_{2is}(e^{-\ell}) K_{2is}(e^{-\ell'})$

Following work by [Saad], we can glue two of these rectangles together to form the cylinder wormhole:



$$= \int \mathrm{d}\ell \mathrm{d}\ell' e^{-\Delta_W \ell' - \Delta_V \ell} P(\tfrac{\beta}{2} + it, \tfrac{\beta}{2} - it; \ell, \ell') \times P(\tfrac{\beta}{2} - it, \tfrac{\beta}{2} + it; \ell', \ell)$$

$$= \int \mathrm{d}E \mathrm{d}E' \rho(E) \rho(E') e^{-\beta(E+E')} |W_{E,E'}|^2 |V_{E,E'}|^2$$

This precisely matches expectations for the ensemble average in a collection of theories described by ETH statistics.

Do the replica wormholes for the Page curve need an ensemble interpretation?

If you treat the replica wormhole computation as a black box, then no.

But you can look inside the box, at quantities like $|\langle W|V\rangle|^2$. The simple gravity description isn't rich enough to describe the true erratic dependence on W, V.

A less ambitious theory would break down when asked about such quantities. But gravity does not. Instead, it gives the answer for an ensemble average.

Key question: in cases where there is no ensemble, are replica wormholes right? Renyi entropy is self-averaging, so an ensemble averaged answer will be close anyhow... but we can ask for more: are replica wormholes the first term in some computation that can be systematically improved? What is the next term?

Thanks for listening!