The Black Hole Information Paradox in AdS-CFT

Suvrat Raju

International Centre for Theoretical Sciences

Regional Meeting in String Theory
20 June 2013

Based on `arXiv:1211.6767` (with Kyriakos Papadodimas)

Information Paradox: Two Aspects

- Can "small corrections" restore unitarity to the density matrix of the Hawking radiation outside the black hole.
- Strong Subadditivity paradox a.k.a. firewall/fuzzball paradox.

Efficacy of small corrections

Exponentially small corrections of the order of e^{-S} can restore unitarity to Hawking radiation.

Producing Entangled Pairs

- Simple understanding of Hawking radiation: imagine particles and anti-particles are produced at the horizon by vacuum fluctuations.
- One of the pair falls into the black-hole, and the other falls out.
- For photons, for example, vacuum fluctuations could create an entangled pair

$$|\Psi\rangle_{\text{hawk}} = \frac{1}{\sqrt{2}} \left(|+\rangle|-\rangle + |-\rangle|+\rangle \right)$$

- The density matrix of the bit falling outside is

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Let us naively assume that the *same process* of Hawking radiation repeats K times.

- The density matrix of the radiation outside will look like

$$\rho_0^K \equiv \rho \otimes \rho \otimes \rho \ldots \text{K times},$$

- The entropy of this density matrix is

$$S_{\text{hawk}} = -\mathrm{Tr}(\rho_K \ln \rho_K) = K \ln 2.$$

Can Small Corrections Unitarize Hawking Radiation?

- If we modify each individual density matrix by a small amount,

$$\rho_{\text{str}} = \rho_1 + \epsilon \rho_{\text{corr}},$$

- But assume that these corrections are uncorrelated then

$$S_{\text{hawk}} - \left[-\text{Tr}(\rho_{\text{str}}^K \ln(\rho_{\text{str}})^K) \right] \sim O(\epsilon).$$

- This is Mathur's argument for why small corrections cannot unitarize Hawking radiation.

Unitarizing Hawking Radiation with small corrections

- The point is that exponentially small correlations between different Hawking quanta can unitarize Hawking radiation.
- This correction cannot be detected at any order in perturbation theory: it is inherently non-perturbative.

Aside on Exponentially Suppressed corrections

- The entropy of a solar-mass black hole is approximately 10^{77}.
- So, exponentially suppressed corrections are of the order $e^{-10^{77}}$!

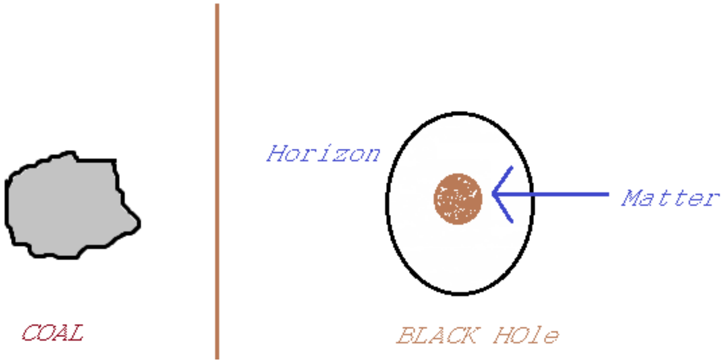
Path Integral Perspective

- Imagine formulating quantum gravity through the Feynman path integral

$$\mathcal{Z} = \int e^{-S} \mathcal{D} g_{\mu\nu}$$

- A semi-classical spacetime is a saddle point of this path-integral.
- Perturbative effective field theory (used to derive the Hawking answer) is an asymptotic series expansion of this path-integral.
- Non-perturbatively, the notion of spacetime breaks down.

Non-perturbative nonlocality!



Form of the Corrections

- We need:

$$\rho_{\text{exact}} = \rho_{\text{hawk}} + 2^{-N} \rho_{\text{corr}},$$

- The condition is that in a natural basis of observables, ρ_{corr} has elements that are $O(1)$.

Exponential small corrections contd ...

- Now after N steps, the Hawking density matrix looks like a big identity matrix

$$\rho_{\text{hawk}} = \frac{1}{2^N} I_{2^N \times 2^N},$$

- If ρ_{exact} is unitary, we must have (in some basis):

$$\rho_{\text{exact}} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}.$$

- So (in a possibly very unnatural basis) the correction must be:

$$\rho_{\text{corr}} = \begin{pmatrix} 2^N - 1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 \end{pmatrix}.$$

Exponential Small Corrections contd...

- This is perfectly consistent with the statement that the elements of ρ_{corr} are O(1) in some natural basis.
- Consistency check: we must have

$$\text{Tr}(\rho_{\text{corr}}^2) \approx 2^{2N}.$$

- But,

$$\text{Tr}(\rho_{\text{corr}}^2) = \sum_{ij} \rho_{\text{corr}}^{ij} \rho_{\text{corr}}^{ij} = \text{O}\left(2^{2N}\right),$$

which is just the number of elements of ρ_{corr}.

A Toy Model

- It is easy to produce a toy-model where the density matrix has these properties.
- Consider a system of N spin-(1/2) spins. This has 2^N states. We can label these states by numbers and read off the individual spins using the binary expansion of the number.

$$
\begin{aligned}
&|+++\ldots\ldots\ldots++\rangle \equiv |0\rangle \\
&|+++\ldots\ldots\ldots+-\rangle \equiv |1\rangle \\
&|+++\ldots\ldots\ldots-+\rangle \equiv |2\rangle \\
&|+++\ldots\ldots\ldots--\rangle \equiv |3\rangle \\
&\ldots
\end{aligned}
$$

Pure States and Hawking Evaporation

- Consider a generic pure state in this spin-model

$$|\Psi\rangle = \frac{1}{2^{\frac{N}{2}}} \sum_{i=0}^{2^N-1} a_i |i\rangle$$

where the a_i are chosen to be either 1 or -1 with probability $\frac{1}{2}$.

- Consider breaking off the spins one by one.

Thermal Density Matrices with Small Corrections

- Even though the full density matrix is pure, if we consider K-spins for $K << \frac{N}{2}$, their density matrix will look thermal up to exponentially small corrections.

- For example,

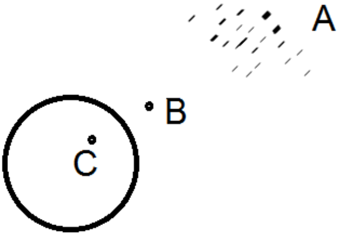
$$\rho_1 = \frac{1}{2^N} \left(\sum_{j=0}^{2^{N-1}-1} a_{2j}^2 |0\rangle\langle 0| + a_{2j+1}^2 |1\rangle\langle 1| + a_{2j} a_{2j+1} \Big(|0\rangle\langle 1| + |1\rangle\langle 0| \Big) \right)$$
$$= \frac{1}{2} \left(|0\rangle\langle 0| + |1\rangle\langle 1| + \mathrm{O}\left(2^{-\frac{N}{2}}\right) \Big(|0\rangle\langle 1| + |1\rangle\langle 0| \Big) \right).$$

- But if we start looking at $\frac{N}{2}$ spins or more, the exponentially small corrections become important.

Recent sharpening of the information paradox

- The info paradox was sharpened by Mathur in 2009.
- This argument has recently been expanded upon by Almheiri, Marolf, Polchinski, and Sully, and has attracted much attention.

Three Subsystems



They key point is to think of three subsystems

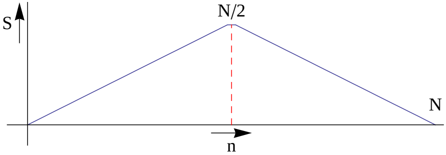
1. The radiation emitted long ago – A
2. The Hawking quanta just being emitted – B
3. Its partner falling into the BH – C

Entropy of A

- Say the Black Hole is formed by the collapse of a pure state.
- Consider the entropy of system A

$$S_A = -\mathrm{Tr}\rho_A \ln \rho_A$$

- Very general arguments tell us this must eventually start decreasing



Strong Subadditivity contradiction?

- Now, consider an old black hole, beyond its "Page time" where S_A is decreasing. We must have

$$S_{AB} < S_A$$

since B is purifying A.

- Second, the pair B, C is related to the Bogoliubov transform of the vacuum of the infalling observer, we have

$$S_{BC} = 0$$

- Finally, both B and C are thermal, so

$$S_B = S_C > 0$$

- However, a very general theorem tells us that for any three distinct systems A, B, C, we have

$$S_A + S_C < S_{AB} + S_{BC}$$

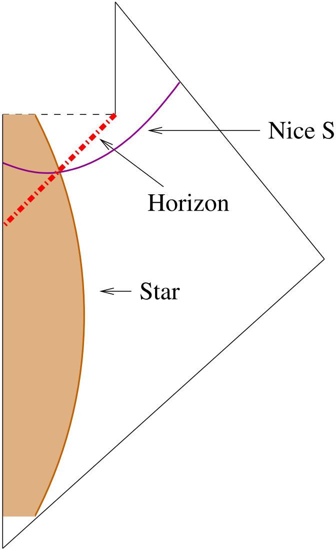
Resolution to the Strong Subadditivity Paradox

The resolution of the strong-subaddtivity paradox is through

Black Hole Complementarity: The interior and exterior of a black hole are not independent. The interior is a scrambled version of (part of the) exterior!

This resolves the strong subadditivity paradox because A and C are not independent.

Original Motivation for BH Complementarity



Objections to Black Hole Complementarity

- To explain the objection, let us go back to our spin-chain model.
- We can model the Hawking quanta outside the black-hole, as spins "breaking off" from the spin chain. [WARNING: May be misleading]
- What about the Hawking quanta that falls into the black hole? Where do we see that in this toy model?

Infalling Quanta in the Toy Model

- Let us make a simple coarse-graining of the system:

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_{N-1}$$

  The coarse-grained d.o.f. is the first spin, and the fine-grained d.o.fs are all the other spins.

- Let us write out pure state as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|+\rangle |\phi_+\rangle + |-\rangle |\phi_-\rangle \right)$$

- We measure

$$S_1 = |-\rangle\langle -| - |+\rangle\langle +|$$

  We can define

$$\widetilde{S}_1 = |\phi_+\rangle\langle \phi_+| - |\phi_-\rangle\langle \phi_-|$$

- Measurement of $\widetilde{S}_1$ are precisely anti-correlated with measurements of S_1.

Large commutators in the Toy Model

- We could, for example, choose the any p-bits to correspond to the coarse d.o.fs and the other $N - p$ to correspond to the fine d.o.fs
- However, if we take $p > \frac{N}{2}$, then there is no $\widetilde{S}_1$ that has small commutators with $S_1, S_2, \ldots S_p$.
- The naive translation of this fact is: once more than half the black hole has evaporated, we are forced to have large commutators between operators outside and inside the black hole.

The Abstract Problem of BH Complementarity

Perhaps the spins-breaking-off model is not a good model. Abstractly, we need a setup with the following property:

1. A large Hilbert space $\mathcal{H}_{\text{full}}$ and a subspace $\mathcal{H}_{\text{in}}$.
2. A "natural basis" of operators O_n of $\mathcal{H}_{\text{full}}$ and a basis of operators $\widetilde{O}_n$ for $\mathcal{H}_{\text{in}}$.
3. These have the property that $[\widetilde{O}_n, O_m] \sim \frac{c_{nm}}{\dim(\mathcal{H}_{\text{full}})}$
4. Also, $\widetilde{O}_n$ and O_n are perfectly correlated in some given state.

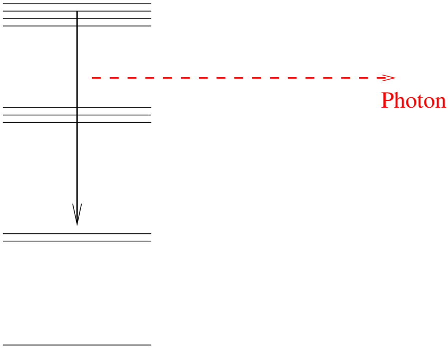
Coarse Graining Again

- We can produce this setup if we assume that the “natural observables” that a coarse-grained observer can access do not span the full space.
- Consider the following analogy: If I measure 10^{26} pieces of information about the gas in this room, there is no continuous density profile that is consistent with my data. At this level of accuracy, the gas is a bunch of distinct atoms.
- Nevertheless, for most purposes, such as waving my hand, a continuous density profile is perfectly good.

Coarse Graining Again

- We can produce this setup if we assume that the “natural observables” that a coarse-grained observer can access do not span the full space.
- Consider the following analogy: If I measure 10^{26} pieces of information about the gas in this room, there is no continuous density profile that is consistent with my data. At this level of accuracy, the gas is a bunch of distinct atoms.
- Nevertheless, for most purposes, such as waving my hand, a continuous density profile is perfectly good.
- If I keep track of all the information in the bits that are emitted by an old black hole, there is no semi-classical metric that can accurately reproduce my measurements. This measurement is inherently non-perturbative.

A Toy Model with Natural Coarse Graining



- Imagine a system, with fine-spacing in its energy levels.
- Transitions between these energy levels lead to the emission of a photon.

Coarse Graining the Photon Field

- The photon field outside can be quantized in terms of

$$A_\mu(x,t) = \sum_n a_{n,\mu} e^{in\omega_0(t-x)} + a'_{n,\mu} e^{in(\omega_0+\epsilon)(t-x)} + h.c$$

- However, a coarse grained observer will see an effective coarse grained field

$$A_\mu^{\text{coarse}}(x,t) = \sum_n (a_{n,\mu} + a'_{n,\mu}) e^{i\omega_n(t-x)} + h.c$$

- If we consider a configuration of photons, with total energy $E = N\omega_1$, then half the degrees of freedom are in excitations of the oscillators $\frac{1}{\sqrt{2}}(a_{n,\mu} - a'_{n,\mu})$.

Exponential degeneracies in the black hole

- The black-hole does have these closely spaced frequencies.
- The energy-levels of a black hole even in global AdS are separated by a spacing of order e^{-S}.
- Semi-classically, it appears that the black-hole emits into all frequencies: the spectrum of the wave-operator on the black hole background is continuous.

Summary: proposed resolution of the strong subadditivity paradox

- The information may be outside the black-hole, but is not accessible to a coarse-grained observer:

$$H_{\text{coarse}} \neq H_{\text{out}}$$

- We can use the fine structure of the emitted radiation to reconstruct the interior of the black hole. (Kyriakos talk.)

- This leads to commutators

$$[\phi_{\text{out}}, \phi_{\text{in}}] \sim e^{-S},$$

which is acceptable.