The black hole information paradox

and

the fate of the infalling observer

Kyriakos Papadodimas

CERN TH-seminar, 02 December 2015
Schwarzschild solution (1916)

\[ ds^2 = - \left( 1 - \frac{2GM}{r} \right) dt^2 + \left( 1 - \frac{2GM}{r} \right)^{-1} dr^2 + r^2 d\Omega^2 \]

Mysterious and fascinating objects in our Universe
Black Holes and puzzles of Quantum Gravity

- S-matrix including BH intermediate states
- Entropy-Area, $S = \frac{A}{4G}$, (UV/IR)
- BH Singularity
- Nature of horizon
- Information paradox
Do we need new physics?

- Modifications of effective field theory at large scales?
- Modifications of Quantum Mechanics in interior?
- Holography and emergence of spacetime
General Relativity: Equivalence Principle, black hole horizon is smooth

Quantum Mechanics: Unitarity, no information loss

Hawking: black holes evaporate

Conflict: Black hole information paradox, “firewall” paradox

Propose a possible way out? (based on work with S.Raju)
Basic info paradox

Hawking computation predicts thermal radiation

Photons thermal and independent (no correlations)

\[ |\psi\rangle_{\text{star}} \Rightarrow \rho_{\text{thermal}} \quad (\ast) \]

Information Loss?

In Quantum Mechanics time evolution is Unitary

\[ |\psi\rangle_{\text{final}} = e^{-iHt} |\psi\rangle_{\text{initial}} \]

Inconsistent with (\ast).
Normal “burning“

Radiation appears to be thermal

Small correlations between photons (of size $e^{-S}$)

Accurate measurement of correlations $\Rightarrow$ full information of initial state

No information loss problem
Resolution of basic version of info paradox

∃ quantum corrections to Hawking’s computation

$e^{-S_{BH}}$ deviations from Hawking’s predictions for simple observables (example: 2-point correlations between photons)

$\Rightarrow$ sufficient to restore unitarity

Reminder: for solar mass BH

$S_{BH} \approx 10^{77}$
Compare outgoing radiation

Hawking

$\Rightarrow \rho_{\text{thermal}}$

Hawking + "corrections"

$\Rightarrow |\psi\rangle_{\text{pure}}$

How different does radiation look?
Pure vs Mixed states

\[ |\psi\rangle = \sum_{i}^{N} c_i |E_i\rangle \quad \text{vs} \quad \rho_{\text{micro}} = \frac{1}{N} \mathbf{I} \]

\[ N = e^S = \text{number of eigenstates} \gg 1 \]

**Theorem:** In a large quantum system, for most pure states, and simple observables \( A \), we have

\[ \langle \psi | A | \psi \rangle = \text{Tr}(\rho_{\text{micro}} A) + O(e^{-S}) \]
Pure vs Mixed states

\[ |\Psi\rangle = \sum_{i} c_i |E_i\rangle \quad \text{vs} \quad \rho_{\text{micro}} = \frac{1}{N} I \]

\[ N = e^S = \text{number of eigenstates} \gg 1 \]

**Theorem:** In a large quantum system, for most pure states, and simple observables \( A \), we have

\[ \langle \Psi | A | \Psi \rangle = \text{Tr}(\rho_{\text{micro}} A) + O(e^{-S}) \]

(not true for complicated observables \( n \approx S \))

\[ \langle \Psi | A_1 ... A_n | \Psi \rangle = \text{Tr}(\rho_{\text{micro}} A_1 ... A_n) + O(e^{-(S-n)}) \]
Define $\langle A \rangle_{\text{micro}} = \text{Tr}(\rho_{\text{micro}}A)$

We also define the average over pure states in $\mathcal{H}_E$

$$\langle \psi | A | \psi \rangle \equiv \int [d\mu_\psi] \langle \psi | A | \psi \rangle$$

where $[d\mu_\psi]$ is the Haar measure. Then for any observable $A$ we have

$$\langle \psi | A | \psi \rangle = \langle A \rangle_{\text{micro}}$$

and

\[
\text{variance} \equiv \frac{1}{e^S+1} \left( \langle A^2 \rangle_{\text{micro}} - \left( \langle A \rangle_{\text{micro}} \right)^2 \right)
\]

"reasonable" observables have the same expectation value in most pure states, up to exponentially small corrections.
Compare outgoing radiation

Small number of photons ⇒ Predictions agree up to $O(e^{-S_{BH}})$

Need to measure correlator between $S_{BH}$ photons to get info

Hawking computation reliable for simple observables
Basic version of info paradox, where we only talk about radiation at infinity, can in principle be resolved: Hawking predicts thermal radiation. Exponentially small deviations $e^{-S_{BH}}$ to simple observables can restore unitarity.

We do not know how to calculate these corrections, but we do expect them on general grounds so there is no paradox.

Computing these corrections, and understanding the microscopic mechanism of information transfer is a bigger problem (S-matrix of Quantum Gravity) but is not really a "paradox".

So far we have not said anything about the BH interior...
Modern info paradox, infalling observer

Curvature at horizon

\[ R^2 \sim \frac{1}{(GM)^4} \]

General Relativity/Equivalence Principle, predicts:

free fall through horizon ⇒ will not notice anything

What if we include Quantum Mechanics?

Problem with Entanglement

Dramatic modification of horizon/interior?
Entanglement Reminder

Two sub-systems $A, B$ then

$$\mathcal{H}_{\text{full}} = \mathcal{H}_A \otimes \mathcal{H}_B$$

Typical state $|\Psi\rangle = \sum_{ij} c_{ij} |i\rangle_A \otimes |j\rangle_B$ does not factorize = ”is entangled“

Example: two spins

Non-entangled state

$$|\Psi\rangle = |\uparrow\rangle_A \otimes |\uparrow\rangle_B$$

Entangled state (EPR)

$$|\Psi\rangle = \frac{|\uparrow\rangle_A \otimes |\uparrow\rangle_B + |\downarrow\rangle_A \otimes |\downarrow\rangle_B}{\sqrt{2}}$$
Ground state of QFT is entangled

\[ \langle \phi(0, x) \phi(0, y) \rangle = \frac{1}{|x - y|^2} \]

\[ |0\rangle_M = \frac{1}{\sqrt{Z}} \prod_\omega \sum_{n=0}^\infty e^{-\pi \omega n} |n\rangle_L \otimes |n\rangle_R \]
Smooth spacetime needs entanglement

\[ \frac{1}{\sqrt{Z}} \prod \sum_{n=0}^{\infty} e^{-\pi \omega n} e^{i \theta n} |n\rangle_L \otimes |n\rangle_R \]

\[ \langle T_{\mu \nu} \rangle \neq 0 \]

Rindler Horizon excited
Monogamy of entanglement

\[ \text{Strong subadditivity of Entanglement Entropy} \]

\[ S_{AB} + S_{BC} \geq S_A + S_C \]
Monogamy of entanglement

$A, B, C$ independent systems

Strong subadditivity of Entanglement Entropy

\[ S_{AB} + S_{BC} \geq S_A + S_C \]
Hawking pair production

Particles of each pair highly entangled

Entanglement required for smoothness of horizon
Modern info Paradox
Mathur [2009], Almheiri, Marolf, Polchinski, Sully (AMPS) [2012]

**General Relativity:** smooth horizon, \( B \) entangled with \( C \)

**Quantum Mechanics:** information preserved, \( B \) entangled with \( A \)

\( B \) violates monogamy?

Mathur’s theorem: small corrections cannot fix the problem (?)
Which one survives, Unitarity or Smooth Horizon?

Giving up B-C entanglement?

Firewall, fuzball proposals $\Rightarrow \langle T_{\mu\nu} \rangle$ at horizon is very large, BH interior geometry is completely modified (maybe no interior at all)

Infalling observer ”burns“ upon impact on the horizon.

Dramatic modification of General Relativity/Effective Field Theory over macroscopic scales, due to quantum effects
Chaos vs entanglement

Black Holes are Chaotic Quantum Systems

How can typical states have specific entanglement between $B$, $C$ which is needed for smoothness?

Correct entanglement fragile under perturbations due to chaotic nature of system [Shenker, Stanford]
Summary

- The modern version of the info paradox, is intimately related to the smoothness of the horizon and to what happens to the infalling observer.

- We have a conflict between QM and General Relativity because it seems impossible to have the entanglement of quantum fields, needed for smoothness, near the horizon.

- Is there a way out?
AdS/CFT

- AdS/CFT: non-perturbative definition of Quantum Gravity by dual gauge theory
- Black Holes in AdS $\Leftrightarrow$ Quark-Gluon-Plasma states in QFT
- BH formation + evaporation $\Leftrightarrow$ deconfinement + hadronization
- Very strong argument in favor of Unitarity
Non-perturbative Black Hole S-matrix encoded in CFT correlators

Manifestly Unitary
Suppose we completely solve the CFT (know all correlators exactly)

How do we reconstruct the black hole interior?

Well-defined question, conceptual/mathematical framework missing?

What computation do we have to do?
- AdS/CFT successful for certain black hole questions
- Until recently, understanding of BH interior was limited
- We identified CFT operators relevant for BH interior
- Seems to resolve the tension of entanglement in modern version of the info paradox
- It is important to make further checks and to expand into a complete mathematical framework
Local observables in AdS

\[ \phi(x) = \int dY \, K(x, Y) \, \mathcal{O}(Y) \]

\( \mathcal{O} \) = local CFT operator
\( K \) = known kernel

Locality in bulk is approximate:
1. True in \( 1/N \) perturbation theory
2. \( [\phi(P_1), \phi(P_2)] = 0 \) only up to \( e^{-N} \) accuracy
3. Locality may break down for high-point functions
Local observables in AdS

\[ \phi(x) = \int dY \, K(x, Y) \, \mathcal{O}(Y) \]

\( \mathcal{O} \) = local CFT operator
\( K \) = known kernel

Locality in bulk is approximate:

1. True in \( 1/N \) perturbation theory
2. \( [\phi(P_1), \phi(P_2)] = 0 \) only up to \( e^{-N^2} \) accuracy
3. Locality may break down for high-point functions
For smooth horizon effective field theory requires:

I) $\tilde{b}$ commute with $b$ \hspace{1cm} AND \hspace{1cm} II) $\tilde{b}$ entangled with $b$

$b \Leftrightarrow O$

$\tilde{b} \Leftrightarrow ?$

Which CFT operators $\tilde{O}$ correspond to $\tilde{b}$?
Smoothness of BH horizon and existence of interior, translated into concrete mathematical problem: can we find CFT operators $\tilde{\mathcal{O}}$ with desired properties.

i) for every single trace operator $\mathcal{O}$ there is a $\tilde{\mathcal{O}}$

ii) $\mathcal{O}$'s and $\tilde{\mathcal{O}}$'s must commute

ii) $\mathcal{O}$'s and $\tilde{\mathcal{O}}$'s must be entangled

We verified the existence of such operators

Interesting property: state-dependence.
Small algebra of observables

EFT operators in bulk correspond to a small sector of boundary
CFT operators (low $\Delta$). They form small algebra

$$\mathcal{A} \equiv \text{span}[\mathcal{O}(x_1), \mathcal{O}(x_1)\mathcal{O}(x_2), ...]$$

The algebra $\mathcal{A}$ acts on the state $|\psi\rangle$ of the system.
If $|\psi\rangle$ is a BH microstate, we have nontrivial property

$$A|\psi\rangle \neq 0 \quad \forall A \in \mathcal{A}, \; A \neq 0$$

Physically this means that the state seems to be entangled when
probed by the algebra $\mathcal{A}$.

Whatever it is entangled with, corresponds to the operators $\tilde{\mathcal{O}}$
Tomita-Takesaki modular theory

Algebra, cannot annihilate state.

⇒ the representation of the algebra is reducible, and the algebra has a nontrivial commutant acting on the same space.

Define antilinear map

\[ SA|\psi\rangle = A^\dagger |\psi\rangle \]

and

\[ \Delta = S^\dagger S \quad J = S\Delta^{-1/2} \]

Then the operators

\[ \tilde{O} = JOJ \]

i) commute with \( O \)

ii) are correctly entangled with \( O \)

These are the operators that we need for the Black Hole interior.
The operator $\Delta$ is a positive, hermitian operator and can be written as

$$\Delta = e^{-K}$$

where

$$K = \text{“modular Hamiltonian”}$$

For entangled bipartite system $A \times B$ this construction would give $K_A \sim \log(\rho_A)$ i.e. the usual modular Hamiltonian for $A$.

In the large $N$ gauge theory and using the KMS condition for correlators of single-trace operators we find that for equilibrium states

$$K = \beta(H_{CFT} - E_0)$$
\[ \tilde{\mathcal{O}}_\omega |\psi\rangle = e^{-\frac{\beta \omega}{2}} \mathcal{O}^\dagger_\omega |\psi\rangle \]

\[ \tilde{\mathcal{O}}_\omega \mathcal{O} \ldots \mathcal{O} |\psi\rangle = \mathcal{O} \ldots \mathcal{O} \tilde{\mathcal{O}}_\omega |\psi\rangle \]

\[ [H, \tilde{\mathcal{O}}_\omega] \mathcal{O} \ldots \mathcal{O} |\psi\rangle = \omega \tilde{\mathcal{O}}_\omega \mathcal{O} \ldots \mathcal{O} |\psi\rangle \]
Bulk field inside BH

$$\phi(t, r, \Omega) = \int_0^\infty d\omega \left[ O_\omega f_\omega(t, \Omega, r) + \tilde{O}_\omega g_\omega(t, \Omega, r) + \text{h.c.} \right]$$

Correlation functions of these operators

$$\langle \Psi | \phi(t_1, r_1, \Omega_1) \cdots \phi(t_n, r_n, \Omega_n) | \Psi \rangle$$

reproduce those of effective field theory in the exterior/interior of the black hole

AdS/CFT: Smooth spacetime at the horizon, no firewall

At the same time, Unitarity OK

We saved Unitarity + Equivalence Principle!
What about previous paradoxes?
Non-locality

\[ [\mathcal{O}, \tilde{\mathcal{O}}] \approx 0 \text{ in simple correlators} \]

Operators \( \tilde{\mathcal{O}} = \text{complicated combinations of } \mathcal{O} \)

Hilbert space of Quantum Gravity: \( \mathcal{H}_{\text{inside}} \otimes \mathcal{H}_{\text{outside}} \)

Solves problem of Monogamy of Entanglement

Concrete realization of “Black Hole Complementarity”, consistent with EFT
State-dependence

- Interior operators defined by

\[ \widetilde{O}_\omega |\Psi\rangle = e^{-\frac{\beta \omega}{2}} O_\omega^\dagger |\Psi\rangle \]

\[ \widetilde{O}_\omega O \ldots O |\Psi\rangle = O \ldots O \widetilde{O}_\omega |\Psi\rangle \]

- Solution depends on reference state |\Psi\rangle

- Operators cannot be upgraded to “globally defined” operators

- Solves Chaos vs Entanglement problem

- Unusual in Quantum Mechanics, needs further study
“Derivation” of $\text{ER} = \text{EPR}$
[K.P and S.R. (1503.08825)]

Entanglement & Wormholes (Maldacena, Susskind, Raamsdonk)

\[ H = H_L + H_R \]

\[ |\text{TFD}\rangle = \sum_E \frac{e^{-\beta E/2}}{\sqrt{Z}} |E\rangle_L \otimes |E\rangle_R \]
Time-shifted wormholes
[K.P and S.R. (1502.06692)]

$$|\Psi_T\rangle \equiv e^{iH_L T} |\text{TFD}\rangle$$

Strong evidence in favor of state-dependence
Thermalization in gauge theories

A class of “quasi-equilibrium“ states

$$|\psi'\rangle = U(\tilde{O}) |\psi\rangle = e^{-\frac{\beta H}{2}} U(O) e^{\frac{\beta H}{2}} |\psi\rangle$$
Outlook

Things to understand:

▶ Resolve certain subtleties
▶ $1/N$ corrections
▶ Thermalization, real time
▶ Time evolution + measurement behind horizon
▶ Singularity
Summary

- The modern version of the info paradox has to do with entanglement at the horizon

- State-dependence may be able to resolve the problem

- Proposal for holographic reconstruction of BH interior, important to develop further.
THANK YOU