The Explanatory Role of Entanglement Wedge Reconstruction in the Black Hole Information Loss Paradox

Jonathan Bain

Department of Technology, Culture and Society
NYU-Tandon, Brooklyn, New York

1. A Puzzle.
2. Two Paradoxes of Info Loss.
3. The RT Formula and the Page Curve.
4. EWR to the Rescue?

Not really...

1. A Puzzle

- *Proposal* (Penington 2020; Almheiri *et al*. 2019)

(1) Derivation of the *Page curve* using the *RT formula*.

Signature of unitarity of evaporation process (Page 1993)

Calculates entanglement entropy in AdS/CFT

(2) Explanation of info escape using *entanglement wedge reconstruction*.

Reconstruction of bulk d.o.f. using boundary d.o.f. in AdS/CFT

Puzzle: The RT formula is *equivalent* to EWR! (Harlow 2017)

If EWR explains information escape, shouldn't the RT formula, too; and vice-versa?

2. Two Paradoxes of Information Loss

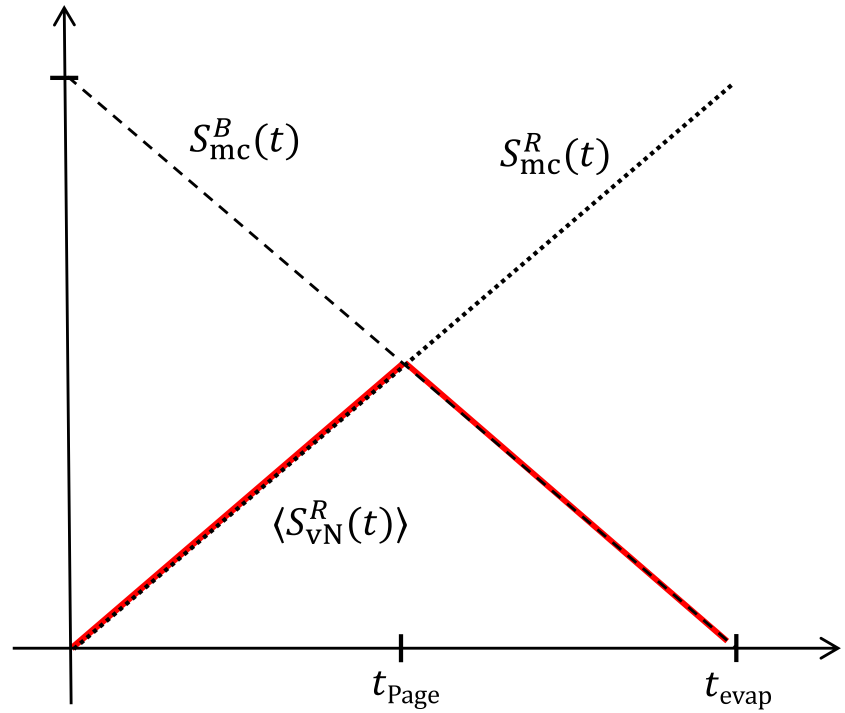
Wallace (2020)

The Page Time Paradox

Page's analysis (Page 1993)

Composite $B+R$ system characterized by:

- $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_R$, with $\dim(\mathcal{H}_R) \ll \dim(\mathcal{H}_B)$ in early stage, and *vice-versa* in late stage.
- Approx. pure initial state.



(P) Post-t_{Page} radiation is entangled with pre-t_{Page} radiation.

(H1) Radiation is in thermal states (at all times).

Hawking (1975)

No Thermal State Entanglement: Thermal states cannot be entangled with other thermal states.

The Firewall Paradox

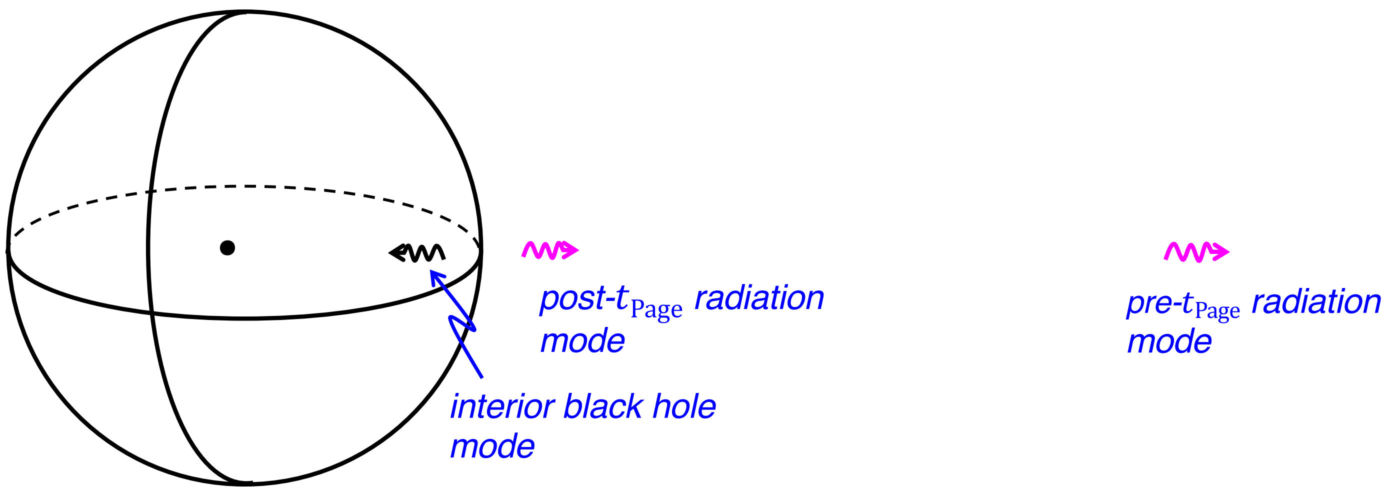
AMPS (2013)

(P) Post-t_{Page} radiation is entangled with pre-t_{Page} radiation.

(H2) Radiation outside horizon is max. entangled with interior black hole modes.

Hawking (1975)

Entanglement Monogamy: A physical system cannot be maximally entangled with two other systems.



The Firewall Paradox

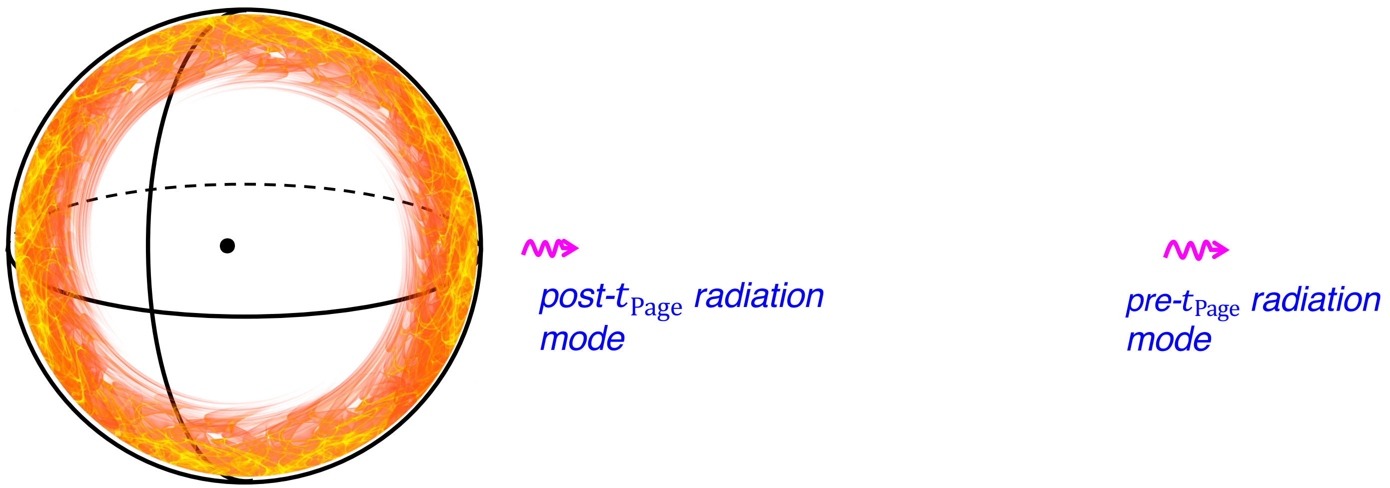
AMPS (2013)

(P) Post-t_{Page} radiation is entangled with pre-t_{Page} radiation.

(H2) Radiation outside horizon is max. entangled with interior black hole modes.

Hawking (1975)

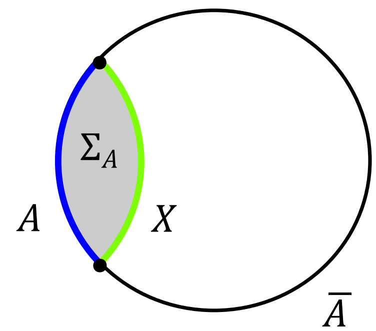
Entanglement Monogamy: A physical system cannot be maximally entangled with two other systems.

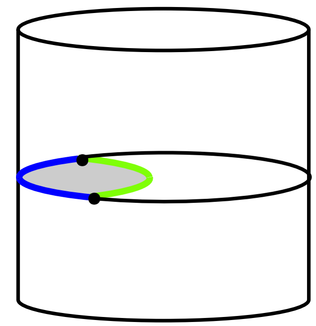


2. The RT Formula and the Page Curve

RT Formula

$$S_{\mathrm{vN}}^{A} = \min_X \left\{ \mathrm{ext}_X \left[\frac{\mathrm{Area}(X)}{4G} + S_{\mathrm{vN}}^{\Sigma_A} \right] \right\}$$





On the left: Entanglement entropy of a subsystem $\mathcal{H}_A$, located in a boundary subregion A, of a system $\mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$ located on the boundary $A \cup \bar{A}$ of a (spatial slice of a) bulk AdS spacetime.

On the right: Bulk surface X (*quantum extremal surface*) that satisfies:

- $\partial X = \partial A$
- $X \cup A = \partial \Sigma_A$
- Extremum of $[\mathrm{Area}(X)/4G + S_{\mathrm{vN}}^{\Sigma_A}]$.

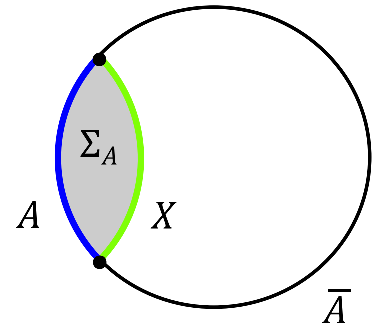
Σ_A = the homology surface determined by X

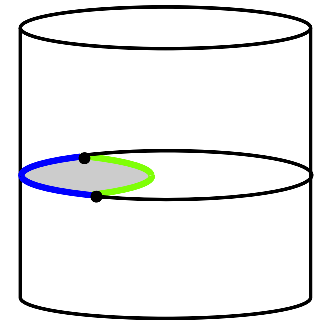
RT formula says: "To calculate S_{vN}^{A}, find a bulk QES X, and if there's more than one, take the minimal one."

2. The RT Formula and the Page Curve

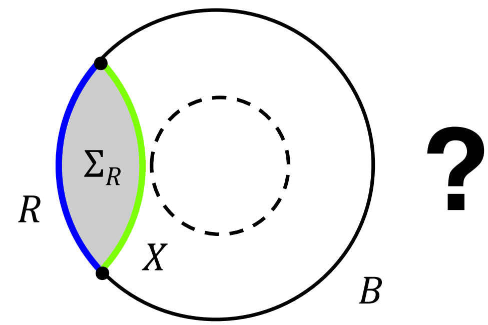
RT Formula

$$S_{\mathrm{vN}}^{A} = \min_X \left\{ \mathrm{ext}_X \left[\frac{\mathrm{Area}(X)}{4G} + S_{\mathrm{vN}}^{\Sigma_A} \right] \right\}$$

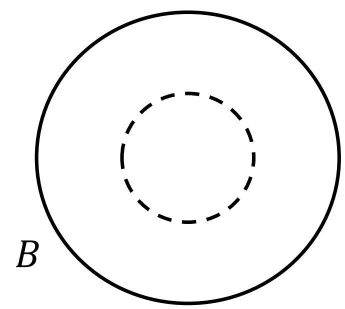


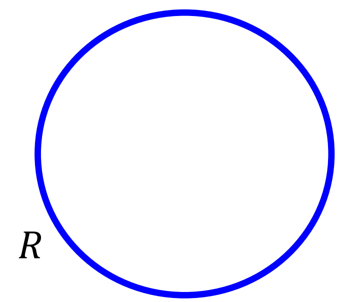


For an evaporating black hole $\mathcal{H}_B \otimes \mathcal{H}_R$...



...two disconnected bulks with separate boundaries:





Evaporating black hole AdS bulk corresponding to $\mathcal{H}_B$ on boundary B.

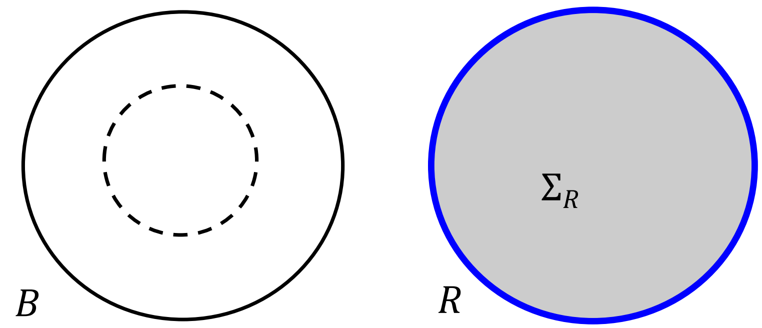
Radiation bulk corresponding to $\mathcal{H}_R$ on boundary R.

Main result: Two QES's

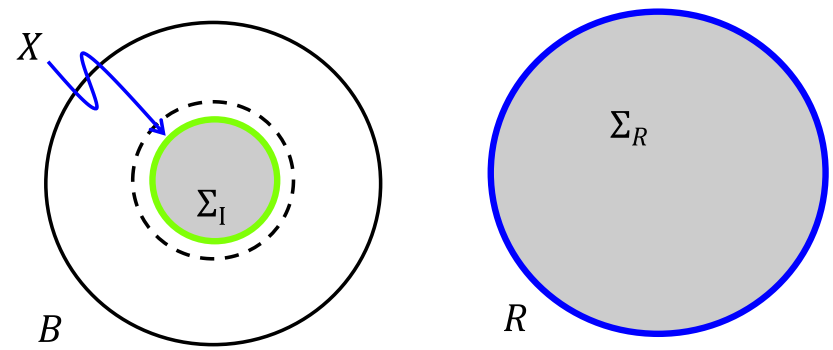
Claim 1. For $0 < t < t_{\text{Page}}$, the minimal QES is the empty surface: $X = \emptyset$.

Technical derivation in general relativity...

Claim 2. For $t_{\text{Page}} < t < t_{\text{evap}}$, the minimal QES X is a closed surface just inside the event horizon.



Homology surface of $X = \Sigma_R$



Homology surface of $X = \Sigma_R \cup \Sigma_{\text{I}}$

Consequence of Claim 1:

$S^R_{\text{vN}}(t)$ increases for $0 < t < t_{\text{Page}}$

Consequence of Claim 2:

$S^R_{\text{vN}}(t)$ decreases for $t_{\text{Page}} < t < t_{\text{evap}}$

Reproduces the Page Curve!

Not enough?

...[W]e emphasize that this [derivation of the Page curve] does not resolve the info paradox, since it supposes that the generalized entropy of the QES correctly captures the fine-grained entropy of the black hole, **without explaining the required correlations between early and late Hawking radiation** (Almheiri *et al*. 2019, pg. 30).

On its own, an explanation of the Page curve using the RT formula is not entirely satis-factory. **It does not explain *why* extracting Hawking radiation into the reservoir should decrease the entanglement entropy. In particular, it does not resolve the Firewall Problem**... Fortunately, we also know about entanglement wedge reconstruction (Penington 2020, pg. 6).

Explananda not explained by the RT formula?

(a) Post-t_{Page} correlations between early and late radiation.

Explained by Page curve?

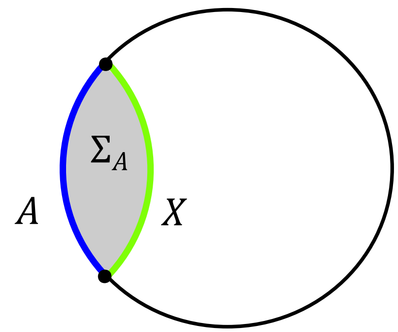
(b) The post-t_{Page} decrease in S^R_{vN}.

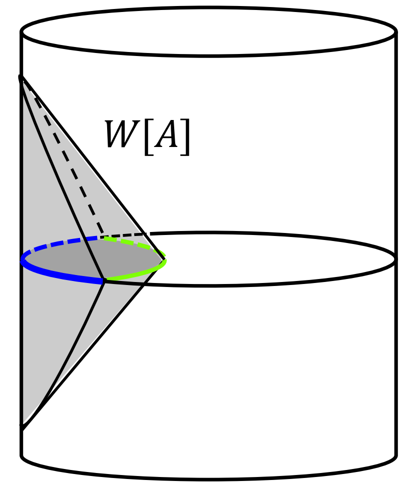
(c) The entanglement of late radiation with both early radiation and interior d.o.f., without a violation of entanglement monogamy.

3. EWR to the Rescue?

RT Formula

$$S_{\mathrm{vN}}^{A} = \min_X \left\{ \mathrm{ext}_X \left[\frac{\mathrm{Area}(X)}{4G} + S_{\mathrm{vN}}^{\Sigma_A} \right] \right\}$$





Entanglement wedge of A

$W[A] \equiv D_{\mathrm{bulk}}[\Sigma_A]$

Determined by X.

Entanglement Wedge Reconstruction (EWR)

Any bulk observable located in $W[A]$ can be represented by a boundary observable located in A.

One way to make EWR precise...

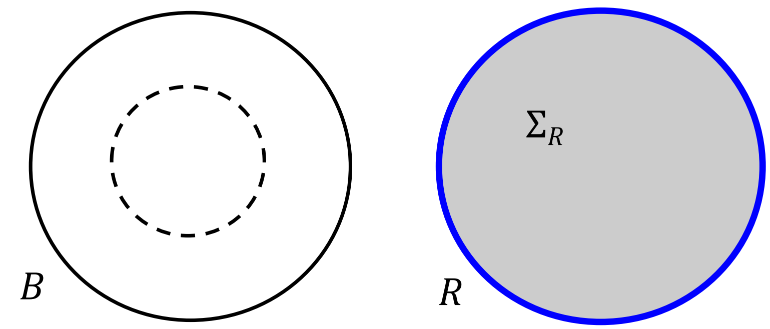
EWR$_{precise}$

Suppose $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$ has a subspace $\mathcal{H}_{\mathrm{C}} = \mathcal{H}_{W[A]} \otimes \mathcal{H}_{W[\bar{A}]}$ that encodes bulk d.o.f. located in $W[A]$ and $W[\bar{A}]$. Then for any operator $\mathcal{O}$ on $\mathcal{H}_{W[A]}$, there is an operator $\mathcal{O}_A$ on $\mathcal{H}_A$ such that, for any $|\psi\rangle \in \mathcal{H}_{\mathrm{C}}$,

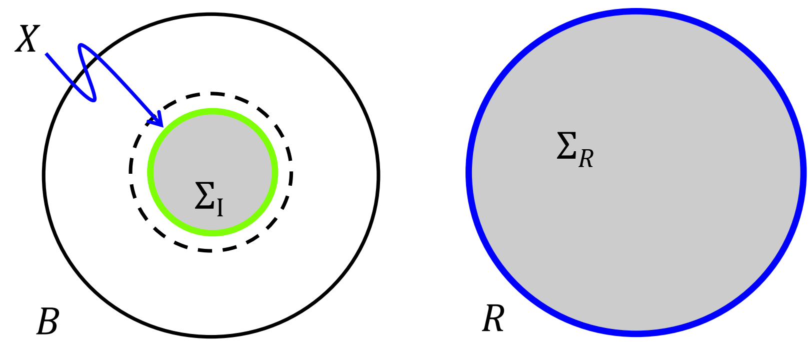
$$\mathcal{O}_A|\psi\rangle = \mathcal{O}|\psi\rangle \quad \text{and} \quad \mathcal{O}_A^{\dagger}|\psi\rangle = \mathcal{O}^{\dagger}|\psi\rangle$$

Entanglement wedges for the radiation

For $0 < t < t_{\mathrm{Page}}$:
$W[R] = D_{\mathrm{bulk}}[\,\Sigma_R]$



For $t_{\mathrm{Page}} < t < t_{\mathrm{evap}}$:
$W[R] = D_{\mathrm{bulk}}[\,\Sigma_R \cup \Sigma_{\mathrm{I}}]$



- <u>*EWR entails*</u>: After t_{Page}, radiation located in R represents bulk d.o.f. located in the grey regions.

D.o.f. in the bulk radiation spacetime Σ_R ("far" from the black hole).

D.o.f. in the bulk region Σ_{I} inside the event horizon.

<u>*Intended interpretation*</u>: After t_{Page}, (most) interior black hole d.o.f. can be identified with early radiation d.o.f.

Explananda not explained by the RT formula?

(a) Post-t_{Page} correlations between early and late radiation.

Explained by Page curve?

(b) The post-t_{Page} decrease in S^{R}_{vN}.

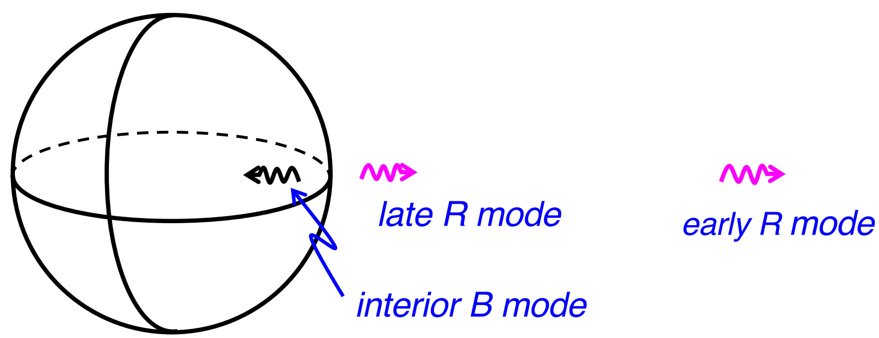
(c) The entanglement of late radiation with both early radiation and interior d.o.f., without a violation of entanglement monogamy.

EWR Explanations...

(a) Late radiation that is entangled with interior black hole d.o.f. is really entanged with early radiation.

(b) After t_{Page}, entanglement between (most) interior d.o.f. and radiation d.o.f. becomes entanglement between early and late radiation d.o.f.; hence entanglement between the B and R subsystems decreases.

(c) After t_{Page}, (most) interior d.o.f. are early radiation d.o.f.; hence a late radiation mode is entangled with just one other system.



EWR as a microphysical explanation of the Page curve?

Claim: The RT formula counts the number of radiation microstates.

- *Can be derived via Euclidean path integral.*
- *Similar to Gibbons & Hawking's (1977) derivation of black hole entropy.*

- *Implication*: Page curve charts the behavior of radiation microstates.

Reasonable: Page curve charts $\langle S_{\mathrm{vN}}^{R} \rangle$ as approximated by max value (i.e., ln[#microstates]).

EWR microphysical explanation of Page curve:
After t_{Page}, the number of radiation microstates decreases, since those formerly associated with early radiation are now black hole microstates.

Concerns:

- Isn't the Page time paradox a conflict between contrasting accounts of entanglement relations between black hole and radiation d.o.f.?
- Isn't the RT formula a relation between entanglement and geometry?
- Shouldn't S_{mc} be kept distinct from S_{vN}?

Swingle (2018)
Ney (2021)
Jaksland (2025)

Moot!

Harlow (2017)

<u>*EWR Theorem*</u>: Let $\mathcal{H}_{\mathrm{C}} = \mathcal{H}_a \otimes \mathcal{H}_{\bar{a}}$ be a subspace of $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. Then the following are equivalent:

(i) There exists an operator $\mathcal{L}$ on $\mathcal{H}_{\mathrm{C}}$ such that, for any state ρ on $\mathcal{H}_{\mathrm{C}}$,

$$S_{\mathrm{vN}}(\rho_A) = \mathrm{Tr}(\rho_a \mathcal{L}) + S_{\mathrm{vN}}(\rho_a)$$

RT formula!

(ii) For any operator $\mathcal{O}$ on $\mathcal{H}_a$, there is an operator $\mathcal{O}_A$ on $\mathcal{H}_A$ such that, for any $|\psi\rangle \in \mathcal{H}_{\mathrm{C}}$,

$$\mathcal{O}_A|\psi\rangle = \mathcal{O}|\psi\rangle \quad \text{and} \quad \mathcal{O}_A^{\dagger}|\psi\rangle = \mathcal{O}^{\dagger}|\psi\rangle$$

$\mathrm{EWR}_{precise}$!

- <u>*Proviso*</u>: Assumes the Quantum Erasure Code (QEC) interpretation of AdS/CFT, underwhich a subspace $\mathcal{H}_{\mathrm{C}}$ of the CFT Hilbert space $\mathcal{H}$ encodes bulk d.o.f. by means of a <u>*QEC map*</u>.

One way to understand $\mathrm{EWR}_{precise}$!

Proof Sketch

<u>*Claim*</u>: $\text{EWR}_{\text{precise}} \Leftrightarrow$ (RT formula)

<u>*Proof*</u>: "$\Rightarrow$":

- Appeal to QEC theorem to establish $\text{EWR}_{\text{precise}} \Leftrightarrow$ (QEC map)

<u>*QEC Theorem*</u>: Let $\mathcal{H}_C = \mathcal{H}_a \otimes \mathcal{H}_{\bar{a}}$ be a subspace of $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. Then the following are equivalent:

(i) Let $\mathcal{H}_A = (\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}) \oplus \mathcal{H}_{A_3}$, with $|A_1| = |a|$ and $|A_3| < |a|$. Then there exist unitaries U_A, $U_{\bar{A}}$ on $\mathcal{H}_A$, $\mathcal{H}_{\bar{A}}$, and a state $|\chi\rangle_{A_2\bar{A}_2} \in \mathcal{H}_{A_2} \otimes \mathcal{H}_{\bar{A}_2}$, such that

$$|\tilde{i}\tilde{j}\rangle = (U_A \otimes I_{\bar{A}})(U_{\bar{A}} \otimes I_A)\big(|i\rangle_{A_1}|j\rangle_{\bar{A}_1}|\chi\rangle_{A_2\bar{A}_2}\big)$$

QEC map!

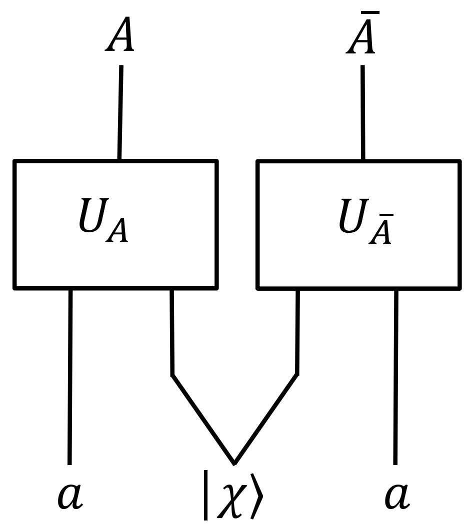
where $|\tilde{i}\rangle$, $|\tilde{j}\rangle$ are bases for $\mathcal{H}_a$, $\mathcal{H}_{\bar{a}}$, and $|i\rangle_{A_1}$, $|j\rangle_{\bar{A}_1}$ are bases for $\mathcal{H}_{A_1}$, $\mathcal{H}_{\bar{A}_1}$.

(ii) For any operator $\mathcal{O}$ on $\mathcal{H}_a$, there is an operator $\mathcal{O}_A \equiv U_A \mathcal{O} U_A^\dagger$ on $\mathcal{H}_A$ such that, for any $|\psi\rangle \in \mathcal{H}_C$,

$$\mathcal{O}_A|\psi\rangle = \mathcal{O}|\psi\rangle \quad \text{and} \quad \mathcal{O}_A^\dagger|\psi\rangle = \mathcal{O}^\dagger|\psi\rangle$$

$\text{EWR}_{\text{precise}}$!

Harlow (2017)



QEC map encodes dof in a (resp. $\bar{a}$) in dof in A (resp. $\bar{A}$) via the ancillary entangled state $|\chi\rangle$, the entangled nature of which protects dof in a (resp. $\bar{a}$) against erasure of dof in $\bar{A}$ (resp. A).

Proof Sketch

Claim: $\mathrm{EWR}_{\mathrm{precise}} \Leftrightarrow$ (RT formula)

Proof: "$\Rightarrow$":

- Appeal to QEC theorem to establish $\mathrm{EWR}_{\mathrm{precise}} \Leftrightarrow$ (QEC map)
- Use QEC map to calculate $S_{\mathrm{vN}}(\rho_A)$, where $\rho_A = \mathrm{Tr}_{\bar{A}}\rho$, with $\rho = \sum_i c_i |\tilde{\imath}\rangle\langle\tilde{\imath}|$ an arbitrary mixed state on $\mathcal{H}_{\mathrm{C}}$. Result:

$$S_{\mathrm{vN}}(\rho_A) = \mathrm{Tr}(\rho_a \mathcal{L}_A) + S_{\mathrm{vN}}(\rho_a)$$

$$\mathcal{L}_A \equiv S_{\mathrm{vN}}\big(\mathrm{Tr}_{\bar{A}_2}|\chi\rangle\langle\chi|\big) I_a$$

Area term encodes entanglement of $|\chi\rangle$!

"$\Leftarrow$": Appeal to Dong, *et al.* (2016) who establish:

(RT formula) $\Rightarrow$ (relative entropy equivalence of bulk and boundary) $\Rightarrow$ $\mathrm{EWR}_{\mathrm{precise}}$

Claim: The QEC assumption is reasonable

- If EWR is to play the explanatory role that its advocates desire, it should be more than just a conjecture in AdS/CFT.
- And (so far) the only way to turn it into a theorem is *via* the QEC interpretation.

4. Conclusion

- EWR explains how "information escapes" (in any number of senses) from an evaporating black hole, provided we model the latter as an AdS/CFT system.
- But so does, and to the same extent, the RT formula.

References

- Almheiri, A., *et al*. (2019) 'The Entropy of Bulk Quantum Fields and the Entanglement Wedge of an Evaporating Black Hole', *JHEP 12*, 63.
- Almheiri, A., *et al*. (2013) 'Black Holes: Complementarity or Firewalls?', *JHEP 02*, 62.
- Cinti, E. & M. Sanchioni (2024) 'Beyond the Quantum Membrane Paradigm: A Philosophical Analysis of the Structure of Black Holes in Full QG', *Found Phys 54: 27*.
- Dong, X., et al. (2016) 'Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality', *Phys Rev Lett 117,* 021601.
- Harlow, D. (2017) 'The RT Formula from Quantum Error Correction', *Comm Math Phys 354*, 865-912.
- Jaksland, R. (2025) 'Spacetime From Entanglement: The Emergence of Metric, Gravity, or Topology', *Phil Phys 3*, 1-25.
- Ney, A. (2021) 'From Quantum Entanglement to Spatiotemporal Distance', in Wuthrich, C. *et al*. (eds.) *Philosophy Beyond Spacetime*, OUP, 78-102.
- Page, D. (1993) 'Information in Black Hole Radiation', *Phys Rev Lett 71*, 3743-6.
- Penington, G. (2020) 'EWR and the Information Paradox', *JHEP 09*, 002.
- Swingle, B. (2018) 'Spacetime from Entanglement', *Annu. Rev. Condens. Matter Phys. 2018*, 9: 345-58.
- Wallace, D. (2020) 'Why Black Hole Information Loss is Paradoxical', in Huggett, N., *et al*. (eds.) *Beyond Spacetime: The Foundations of Quantum Gravity*, CUP, 209-36