Entanglement entropies of equilibrated pure states and the origin of replica wormholes

Shreya Vardhan
Work with Hong Liu

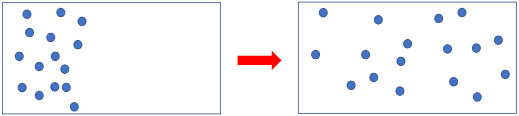
Massachusetts Institute of Technology

University of Ljubljana Holography Group Seminar, March 25th 2021
arXiv: 2008.01089
1912.08918, 2002.05734

Background and motivations: thermalization and unitarity

Equilibration

- Non-integrable many body systems are universally expected to approach equilibrium from an out-of-equilibrium initial state.



- Consider a quantum many-body system initially in a pure state $|\psi_0\rangle$.
$$|\psi(t)\rangle = U(t)\,|\psi_0\rangle$$

- At late times, $|\psi(t)\rangle$ should macroscopically resemble some equilibrium density matrix $\rho^{(\mathrm{eq})}$.

  For generic few-body observables O,
$$\langle\psi(t)|O|\psi(t)\rangle \approx \mathrm{Tr}[\rho^{(\mathrm{eq})} O]$$

Constraints from unitarity

- Under unitary evolution, pure states evolve to pure states. In particular, $\rho(t) = |\psi(t)\rangle \langle\psi(t)|$ cannot become equal to $\rho^{(\mathrm{eq})}$.

- Renyi entropies are one set of observables that can distinguish between pure and mixed states:
  (A is any subsystem, and $\bar{A}$ is its complement)

$$S_n^{(A)}(t) = -\frac{1}{n-1} \log \mathrm{Tr} \rho_A^n(t), \quad \rho_A(t) = \mathrm{Tr}_{\bar{A}} \rho(t), \quad n = 1, 2, \ldots$$

  In a pure state, for any n, we should have for any subsystem A

$$S_n^{(A)} = S_n^{(\bar{A})}.$$

  This is not true for a mixed state like $\rho^{(\mathrm{eq})}$.

- What form do the Renyi entropies of $|\psi(t)\rangle$ take at late times when the state has macroscopically equilibrated?

Black holes as thermalized pure states

Isolated star in pure state $|\psi_0\rangle$

$$\downarrow$$

Black hole

$$\downarrow$$

Black hole radiates (and often evaporates)

Black holes show macroscopic behaviour resembling thermal states:

- They emit thermal radiation at some temperature.
- Have an entropy satisfying standard thermodynamic relations.
- Correlation functions of few-body observables take thermal values.

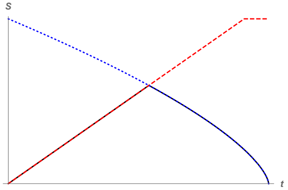
Black holes as thermalized pure states

- If black holes obey the unitarity of quantum mechanics, then the final state must be an equilibrated pure state.
- In particular, at all times, we should have

$$S_n^{(B)} = S_n^{(R)},$$

where B is the black hole and R is its emitted radiation.

- A calculation in semiclassical gravity by Hawking predicted that $S_n^{(R)} \gg S_n^{(B)}$ at late times. (S_R: red; S_B: blue)



- How can we obtain a unitary result for the entanglement entropies of an evaporating black hole?

Typical pure states

- Consider some system $L = A \otimes \bar{A}$. Average $S_n^{(A)}$ and $S_n^{(\bar{A})}$ over all pure states $|\psi\rangle \in L$ with the uniform measure. When either $d_A \ll d_{\bar{A}}$ or $d_{\bar{A}} \ll d_A$, Lubkin, Lloyd and Pagels, Page

$$\overline{S_n^{(A)}} = \overline{S_n^{(\bar{A})}} = \min\left(\log d_A, \log d_{\bar{A}}\right), \quad n = 1, 2, \cdots .$$

  If d_L is large, standard deviation about this average is small.

- Consistent with unitarity.

- When $d_A \ll d_{\bar{A}}$, the value of $S_n^{(A)}$ in typical pure states matches the value we would expect for a thermal state at infinite temperature, $\mathbf{1}/d_A$.

- How should this expression be generalized to finite temperature?

- Generalization to infinite-dimensional Hilbert spaces (e.g. QFTs)?

- Can we find a systematically improvable approximation for these quantities?

Plan

- Derivation of approximation method for finding entanglement entropies in equilibrated pure states.

  (Applicable for finite temperatures and infinite-dimensional Hilbert spaces; can in principle be systematically improved.)

- Mathematical structure and physical implications.
- Applications to the black hole information loss paradox.

Developing the approximation method

- Starting point: assume we can identify some equilibrium density matrix $\rho^{(\mathrm{eq})}$ that $|\psi(t)\rangle = U|\psi_0\rangle$ resembles at late times ($t \gg t_s$).

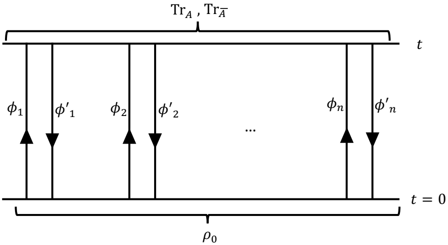
$$\rho^{(\mathrm{eq})} = \frac{1}{Z(\alpha)}\mathcal{I}_\alpha, \quad Z(\alpha) = \mathrm{Tr}\,\mathcal{I}_\alpha$$

- We will assume $Z(\alpha) \gg 1$.
- $\mathcal{I}_\alpha$ is invariant under the time-evolution, $U\mathcal{I}_\alpha U^\dagger = \mathcal{I}_\alpha$.
- Examples:
  - Infinite temperature:
  $$\mathcal{I} = \mathbf{1}, \quad Z = d$$
  - Microcanonical ensemble:
  $$\mathcal{I}_E = \sum_{E_n \in I} |n\rangle\langle n|, \quad Z(E) = \mathrm{Tr}\,\mathcal{I}_E = N_I, \quad I = (E - \Delta E, E + \Delta E)$$
  - Canonical ensemble:
  $$\mathcal{I}_\beta = e^{-\beta H}, \quad Z(\beta) = \mathrm{Tr} e^{-\beta H}$$

- We will now provide two kinds of expressions for the Renyi entropy as a transition amplitude on a replicated Hilbert space.
- Path integral representation:

$$\mathcal{Z}_n^{(A)} = e^{-(n-1)S_n^{(A)}} = \mathrm{Tr}_A \rho_A^n = \mathrm{Tr}_A \left(\mathrm{Tr}_{\bar{A}} U \rho_0 U^\dagger \right)^n$$

$$= (\dots) \prod_{i=1}^{n} \int_{\chi_i,\chi_i'}^{\psi_i,\psi_i'} D\phi_i(t) D\phi_i'(t) \exp\left(i \sum_{i=1}^{n} (S[\phi_i] - S[\phi_i']) \right)$$



- Exponent in path integral vanishes in configurations where Zhou and Nahum

$$\phi_i = \phi'_{\sigma(i)}, \quad \sigma \in S_n.$$

- Naively putting this into the path integral, we get a divergent answer.

Notation:

- σ, τ: permutations in S_n.
- e = identity permutation, $\eta = (n, n-1, ..., 1)$.
- For any operator $\mathcal{O}$ acting on $\mathcal{H}$, and any permutation $\sigma \in S_n$, define state $|\mathcal{O}, \sigma\rangle \in (\mathcal{H} \otimes \mathcal{H})^n$: $\quad$ ($|i\rangle$ is a basis for $\mathcal{H}$)

$$\langle i_1 \bar{i}'_1 i_2 \bar{i}'_2 \cdots i_n \bar{i}'_n | \mathcal{O}, \sigma\rangle = \mathcal{O}_{i_1 i'_{\sigma(1)}} \mathcal{O}_{i_2 i'_{\sigma(2)}} \cdots \mathcal{O}_{i_n i'_{\sigma(n)}}, \quad \mathcal{O}_{ij} = \langle i|\mathcal{O}|j\rangle$$

- Special case of $\mathcal{O} = \mathbf{1}$:

$$\langle i_1 \bar{i}'_1 i_2 \bar{i}'_2 \cdots i_n \bar{i}'_n | \sigma\rangle = \delta_{i_1 i'_{\sigma(1)}} \delta_{i_2 i'_{\sigma(2)}} \cdots \delta_{i_n i'_{\sigma(n)}} \ .$$

- $\langle e|\mathcal{O}, e\rangle = \mathrm{Tr}[\mathcal{O}]^n, \qquad \langle \eta|\mathcal{O}, e\rangle = \mathrm{Tr}[\mathcal{O}^n]$
- Similarly, can define $|O, \sigma\rangle_A \in (\mathcal{H}_A \otimes \mathcal{H}_A)^n$.

Using this notation, we can rewrite

$$\mathcal{Z}_n^{(A)} = \mathrm{Tr}_A \left(\mathrm{Tr}_{\bar{A}} U \rho_0 U^\dagger\right)^n = \langle \eta_A \otimes e_{\bar{A}}|(U \otimes U^\dagger)^n|\rho_0, e\rangle$$

- Define projector onto states spanned by $\{|\mathcal{I}_\alpha, \sigma\rangle\}_{\sigma \in S_n}$:
$$P_\alpha = \frac{1}{Z_2^n} \sum_{\sigma,\tau} g^{\sigma\tau} \, |\mathcal{I}_\alpha, \sigma\rangle \langle \mathcal{I}_\alpha, \tau| \ .$$
- Decomposing the identity on $(\mathcal{H} \otimes \mathcal{H})^n$ as
$$\mathbf{1} = P_\alpha + Q, \qquad P_\alpha Q = Q P_\alpha = 0, \qquad Q^2 = Q$$
we can rewrite
$$\mathcal{Z}_n^{(A)} = \langle \eta_A \otimes e_{\bar{A}} | (P_\alpha + Q)(U \otimes U^\dagger)^n (P_\alpha + Q) | \rho_0, e \rangle$$
- But from $U \mathcal{I}_\alpha U^\dagger = \mathcal{I}_\alpha$,
$$(U \otimes U^\dagger)^n P_\alpha = P_\alpha (U \otimes U^\dagger)^n = P_\alpha \ .$$
- Hence,
$$\mathcal{Z}_n^{(A)} = \langle \eta_A \otimes e_{\bar{A}} | P_\alpha | \rho_0, e \rangle + \langle \eta_A \otimes e_{\bar{A}} | Q (U \otimes U^\dagger)^n Q | \rho_0, e \rangle$$
$$\equiv \mathcal{Z}_{n,P}^{(A)} + \mathcal{Z}_{n,Q}^{(A)} \ .$$
Note that the first term is time-independent.
- Proposal: At late times, and in systems with large Hilbert space dimension, $\mathcal{Z}_{n,Q}^{(A)}$ can be ignored.

$$\mathcal{Z}_n^{(A)} \approx \mathcal{Z}_{n,P}^{(A)} = \frac{1}{Z_2^n} \sum_{\sigma,\tau} g^{\tau\sigma} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \langle \mathcal{I}_\alpha, \sigma | \rho_0, e \rangle , \quad n = 1, 2, 3, \cdots \; .$$

Using the fact that the effective Hilbert space dimension Z_1 is large, we can simplify to:

$$\mathcal{Z}_n^{(A)} \approx \sum_{\tau \in \mathcal{S}_n} \mathcal{Z}_n^{(A)}(\tau), \qquad \mathcal{Z}_n^{(A)}(\tau) = \frac{1}{Z_1^n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \; .$$

In particular, note that the dependence on ρ_0 drops out, and the final expression is written only in terms of $\mathcal{I}_\alpha$.

Justification in infinite temperature case

- Exact expression:
$$\mathcal{Z}_n^{(A)} = \mathrm{Tr}_A \left(\mathrm{Tr}_{\bar{A}} U \rho_0 U^\dagger\right)^n = \langle \eta_A \otimes e_{\bar{A}} | (U \otimes U^\dagger)^n | \rho_0, e \rangle$$
- Averaging over all unitary time-evolution operators U with the Haar measure, we get the equilibrium approximation for infinite temperature Zhou and Nahum
$$\overline{(U \otimes U^\dagger)^n} = P_{\mathcal{I}=\mathbf{1}} = \frac{1}{d^n} \sum_{\sigma,\tau} g^{\sigma\tau} \left|\sigma\right\rangle \left\langle\tau\right|$$
- Consider the variance under the Haar average:
$$\overline{\left(\mathcal{Z}_{n,Q}^{(A)}\right)^2} = \overline{(\mathcal{Z}_n^{(A)})^2} - (\mathcal{Z}_{n,P}^{(A)})^2.$$
- In the large d limit,
$$\overline{\left(\mathcal{Z}_{n,Q}^{(A)}\right)^2} \ll (\mathcal{Z}_{n,P}^{(A)})^2.$$
- For other choices of $\rho^{(\mathrm{eq})}$, we do not have a similar average over relevant time-evolution operators.

Self-consistency criterion

- For any time-evolved quantity that can be written in the form

$$F(t) = \langle A| \; (U \otimes U^{\dagger})^{q} \; |\rho_0, e\rangle$$

we can define the equilibrium approximation

$$[F]_{\text{eq app}} = \langle A| \; P_{\alpha} \; |\rho_0, e\rangle$$

- To see if the approximation is valid for a quantity F, we use the self-consistency criterion

$$[F^2]_{\text{eq app}} - ([F]_{\text{eq app}})^2 \ll ([F]_{\text{eq app}})^2.$$

- For the Renyi entropies, we can show

$$\Delta \equiv \left[\left(\mathcal{Z}_n^{(A)}\right)^2\right]_{\text{eq app}} - \left(\mathcal{Z}_{n,P}^{(A)}\right)^2 \ll \left(\mathcal{Z}_{n,P}^{(A)}\right)^2$$

when the effective Hilbert space dimension is large.

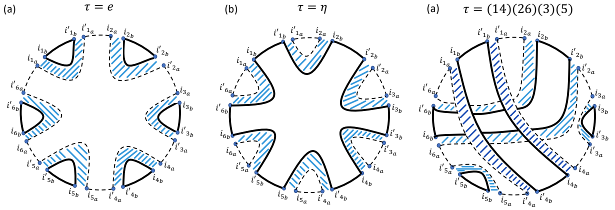
Equilibrium approximation for other observables

- Various other quantities can be expressed as transition amplitudes on replicated Hilbert spaces:
  n-point functions of observables, spectral form factors, Renyi negativities, and Renyi relative entropies.
- To see whether the equilibrium approximation is valid, need to check self-consistency criterion.
- **Expectation values of observables:**
  $$\langle O(t)\rangle_{\text{eq app}} = \text{Tr}[\rho_{\text{eq}} O]$$
  If O has support only on subsystem A, then the self-consistency criterion is met if $d_A \ll d_{\bar{A}}$, otherwise not valid.
- $n-$**th power of spectral form factor:**
  $$\mathcal{S}_n \equiv (\text{Tr}U\text{Tr}U^\dagger)^n = \hat{\text{Tr}}[(U \otimes U^\dagger)^n]$$
  For any choice of $\mathcal{I}_\alpha$, we get
  $$[\mathcal{S}_n]_{\text{eq app}} = n!$$
  So the self-consistency criterion is not met for any n. (Also, does not capture expected late-time features of ramp, plateau.)

Mathematical structure and physical consequences

Diagrammatic interpretation

$$\mathcal{Z}_n^{(A)} \approx \sum_{\tau \in \mathcal{S}_n} \mathcal{Z}_n^{(A)}(\tau), \qquad \mathcal{Z}_n^{(A)}(\tau) = \frac{1}{Z_1^n} \langle \eta_A \otimes e_{\bar{A}} | \mathcal{I}_\alpha, \tau \rangle \ .$$



$k(\tau)$: number of cycles in τ, number of solid loops in diagram.
$k(\eta\tau^{-1})$: number of dashed loops.

$$\mathcal{Z}_n^{(A)}(\tau) \sim \frac{1}{Z_1^n} d_A^{k(\eta^{-1}\tau)} d_{\bar{A}}^{k(\tau)}$$

For any τ,

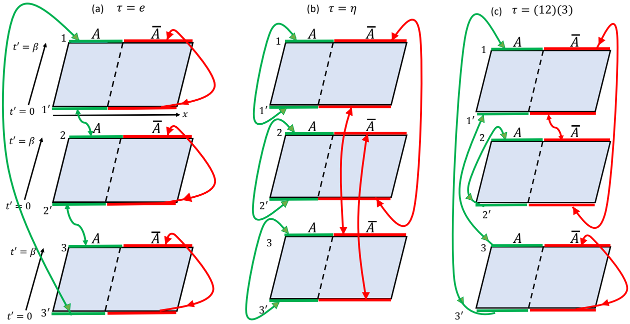
$$k(\eta^{-1}\tau) + k(\tau) \leq n + 1 \ .$$

Saturated for non-crossing permutations/ planar diagrams.

Path integral representation

$$\mathcal{Z}_n^{(A)} \approx \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \langle i_{\eta(1)_a} i_{1_b} | \mathcal{I}_\alpha | i_{\tau(1)_a} i_{\tau(1)_b} \rangle \cdots \langle i_{\eta(n)_a} i_{n_b} | \mathcal{I}_\alpha | i_{\tau(n)_a} i_{\tau(n)_b} \rangle$$

$$= \frac{1}{Z_1^n} \sum_{\tau \in \mathcal{S}_n} \int \prod_{i=1}^{n} D\psi_i D\psi_i' \, \delta(\psi_{iA} - \psi'_{\tau^{-1}\eta(i)A}) \delta(\psi_{i\bar{A}} - \psi'_{\tau^{-1}(i)\bar{A}}) \prod_{i=1}^{n} \int_{\psi_i'}^{\psi_i} D\phi_i \, e^{-S_E[\phi_i]}$$



A sum of $|S_n|$ Euclidean path path integrals on n copies of the system emerges as an approximation to the original Lorentzian path integral on $2n$ copies.

Universal implications

- Unitarity constraint is satisfied:

$$\mathcal{Z}_n^{(A)} = \sum_\tau \mathcal{Z}_n^{(A)}(\tau) = \frac{1}{2}\sum_\tau \left(\mathcal{Z}_n^{(A)}(\tau) + \mathcal{Z}_n^{(\bar{A})}(\tau)\right)$$

  so that $\mathcal{Z}_n^{(A)} = \mathcal{Z}_n^{(\bar{A})}$.

- Dominant contributions in various regimes:

$$\mathcal{Z}_n^{(A)}(\tau) \sim \frac{1}{Z_1^n}\, d_A^{k(\eta^{-1}\tau)} d_{\bar{A}}^{k(\tau)}$$

  - (i) When $d_A \ll d_{\bar{A}}$, the $\tau = e$ term dominates,

$$\mathcal{Z}_n^{(A)} \approx \mathrm{Tr}_A\left[\left(\mathrm{Tr}_{\bar{A}}\rho_{\mathrm{eq}}\right)^n\right]$$

    (ii) When $d_{\bar{A}} \ll d_A$, the $\tau = \eta$ term dominates,

$$\mathcal{Z}_n^{(A)} \approx \mathrm{Tr}_{\bar{A}}\left[\left(\mathrm{Tr}_A\rho_{\mathrm{eq}}\right)^n\right]$$

    Expected generalization of Page's result:

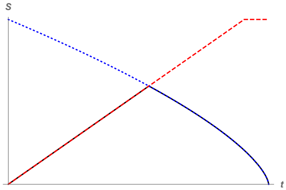
$$S_n^{(A)} = \min\left(S_n^{(A,\mathrm{eq})}, S_n^{(\bar{A},\mathrm{eq})}\right), \quad n \geq 2 \quad \Rightarrow S_1^{(A)} = \min\left(S_1^{(A,\mathrm{eq})}, S_1^{(\bar{A},\mathrm{eq})}\right)$$

  - When $d_{\bar{A}}$ and d_A are both large and comparable, the non-crossing permutations/planar diagrams dominate.

Application to the black hole information loss paradox

The black hole information problem

- Key to showing that black holes do not destroy information despite looking thermal: show that $S_n^{(B)} = S_n^{(R)}$ at all times.
- A calculation in semiclassical gravity by Hawking predicted that $S_n^{(R)} \gg S_n^{(B)}$ at late times. (S_R: red; S_B: blue)

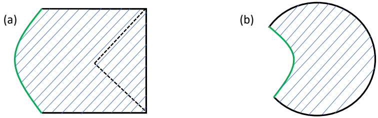


- Previously, it was expected that Hawking's information loss paradox could only be resolved using non-perturbative quantum gravity.
- Recent works Penington; Almheiri, Engelhardt, Marolf, Maxfield; Almheiri, Mahajan, Maldacena, Zhao showed that a semiclassical prescription with some new ingredients is sufficient to get a unitary result.

A simple model for black hole evaporation

Penington, Shenker, Stanford, Yang

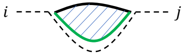
- After the evaporation process, we have a state $|\Psi\rangle \in B \otimes R$.
  $B =$ black hole, $R =$ radiation.
- Black hole lives in a (1+1)-D spacetime with JT gravity, with end-of-the-world (EOW) branes behind the horizon.



- R has N orthonormal states $|i\rangle$.
- Evaluate $(\rho_R)_{ij} = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)_{ij}$ using a Euclidean gravitational path integral:
  - Boundary condition is a segment of length β.

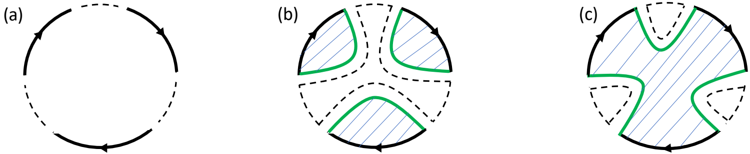
    i and j give b.c. for the state on the EOW brane.
  - This boundary condition is "filled in" with a bulk geometry. (Joining dotted lines gives δ_{ij}.)

A simple model for black hole evaporation

Penington, Shenker, Stanford, Yang

- Now to evaluate, for instance, $\mathcal{Z}_3^{(R)} = \mathrm{Tr}[\rho_R^3] = (\rho_R)_{ij}(\rho_R)_{jk}(\rho_R)_{ki}$



- Geometries like (c), where the Euclidean path integral connects different replicas, are important for ensuring $S_B = S_R$. "Replica wormholes."
- Why should the entanglement entropies of the state formed from the black hole collapse and evaporation be evaluated in this way?
- How do such Euclidean path integrals emerge from the Lorentzian dynamics?
- Why does including such geometries ensure unitarity?

Equilibrium approximation for the evaporation model

- Initial pure state $|\Psi_0\rangle$.
- Final state $|\Psi\rangle = U(t)\,|\Psi_0\rangle$ macroscopically resembles $\rho^{(\text{eq})}$ with

$$\mathcal{I}_\alpha = \mathbf{1}_R \otimes \mathcal{I}^{(B)}_\beta, \quad \mathcal{I}^{(B)}_\beta = e^{-\beta H_B}$$

  (Radiation has no energy constraint, dimension N; black hole associated with temperature β.)

- Applying the equilibrium approximation, we get

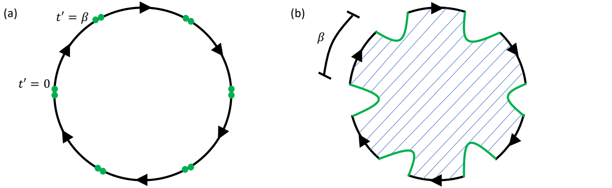
$$\mathcal{Z}^{(R)}_n \approx \frac{1}{\left(N Z^{(B)}_1\right)^n} \sum_\tau N^{k(\tau\eta^{-1})} Z^{(B)}_{n_1} \cdots Z^{(B)}_{n_{k(\tau)}}$$

  where $k(\tau)$: number of cycles in τ;
  $n_1, \dots, n_{k(\tau)}$: lengths of cycles of τ, and

$$Z^{(B)}_m = \mathrm{Tr}[e^{-m\beta H_B}] = e^{S_0} z_m(\beta)$$

- $Z^{(B)}_n$ for each n corresponds to a boundary Euclidean path integral on n copies. Using holography, equal to a bulk path integral.
- e^{S_0} and N both large $\to$ non-crossing permutations dominate.

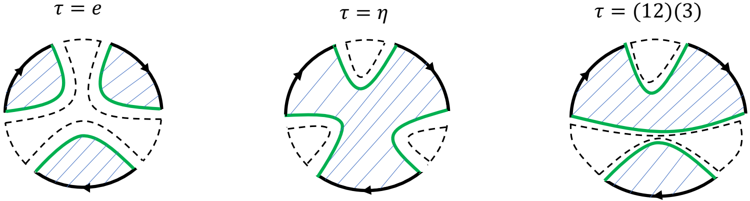
- To match the EOW brane model, $\mathcal{I}_\alpha = \mathbf{1}_R \otimes f(H_B)\mathcal{I}_\beta^{(B)}$.
- Then the bulk and boundary path integrals for calculating $Z_m^{(B)}$



- One-to-one correspondence between contributions from non-crossing permutations τ in

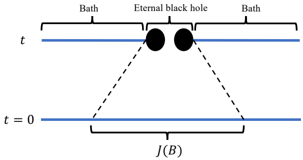
$$\mathcal{Z}_n^{(R)} \approx \frac{1}{\left(NZ_1^{(B)}\right)^n} \sum_\tau N^{k(\tau\eta^{-1})} Z_{n_1}^{(B)} \cdots Z_{n_{k(\tau)}}^{(B)}$$

and planar diagrams contributing in the same limit according to the prescription of Penington, Shenker, Stanford, Yang :



General conclusions

- The equilibrium approximation provides a derivation for the replica wormhole prescription for calculating $\mathcal{Z}_n^{(R)}$.
- Unconventional geometries for $\mathcal{Z}_n^{(R)}$ come from $\tau \neq e$ terms in the Euclidean path integral expression for the equilibrium approximation.
- Unitary result on very general grounds from our earlier argument.
- Can be applied to other black hole models, including:
  - Eternal black hole coupled to an infinite bath. Almheiri, Hartman, Maldacena, Shaghoulian, Tajdini



  - Non-evaporating black holes, such as big black holes in AdS.

Precise prescriptions for when and how replica wormholes should be included.

Non-factorization of matrix elements

- Penington, Shenker, Stanford, Yang noted that including replica wormholes leads to results that suggest that $\langle|(\rho_R)_{ij}|^2\rangle$ does not factorize.
- We find the same issue using the equilibrium approximation: Approximation for matrix elements:
$$(\rho_R)_{ij} = (\rho_R^P)_{ij} + \Delta_{ij}^Q, \qquad (\rho_R)_{ij} \approx (\rho_R^P)_{ij} = \frac{1}{N}\delta_{ij}$$
Approximation to the second Renyi entropy:
$$\mathcal{Z}_2^{(R)} = \sum_{ij} |(\rho_R)_{ij}|^2 \approx \frac{1}{N} + \frac{1}{e^{S_0}} \times O(1)$$
So the two expressions are incompatible when $N \gg e^{S_0}$.
- Using the self-consistency condition for $(\rho_R)_{ij}$, we can estimate
$$|\Delta_{ij}^Q|_{\text{eq app}} \sim \frac{1}{N} e^{-S_0/2}$$
    - Individual off-diagonal matrix elements, coming from fluctuations around the equilibrium value of $(\rho_R)_{ij}$, are suppressed.
    - They make a significant contribution to $\mathcal{Z}_2^{(R)}$ due to their large number.

Semiclassical gravity and averaging

- It has been suggested that the lack of factorization implies that semiclassical gravity computes an average over theories,

$$\overline{|\rho_{ij}|^2} \neq \overline{(\rho_{ij})}\ \overline{(\rho_{ij})^*}$$

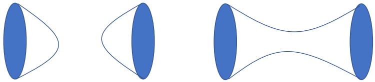
- The equilibrium approximation gives an explanation based on fluctuations around equilibrium. Valid for a theory with a single fixed Hamiltonian, no need for average over theories.

- Our derivation of the replica wormholes consisted of two steps:
  (1) Approximating Lorentzian expressions with a sum of Euclidean path integrals.
  (2) Using semiclassical gravity to compute the Euclidean path integrals.
  The problem with factorization arose at the first step.

- So in principle, the issue can be resolved within semiclassical gravity, by doing a Lorentzian calculation.

Comparison to wormholes in spectral form factors

- In Saad, Shenker, Stanford, trumpet contributions were found for the spectral form factor $\mathrm{Tr}[e^{iHt}]\mathrm{Tr}[e^{-iHt}]$ in Euclidean JT gravity,

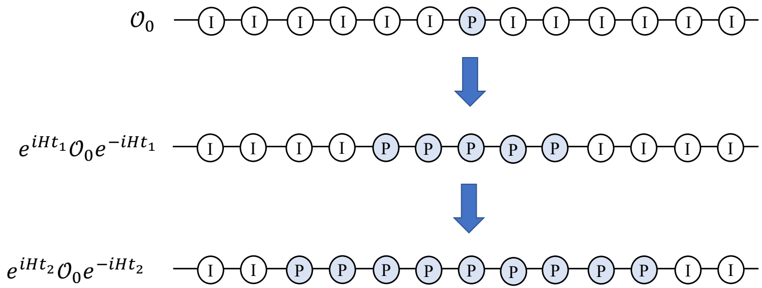


  which lead to the non-factorization of this quantity into the product of $\mathrm{Tr}[e^{iHt}]$ and $\mathrm{Tr}[e^{iHt}]$ computed by the same method.
- As discussed earlier, this quantity is beyond the scope of the equilibrium approximation.
- Note that here the Lorentzian path integral on 2 copies of the system is being evaluated with the help of a Euclidean path integral on 2 copies: this does not have the same structure as the equilibrium approximation.
- The "wormholes" appearing in these two scenarios are thus physically distinct.

Thermalization from an operator evolution perspective

Operator growth in chaotic systems

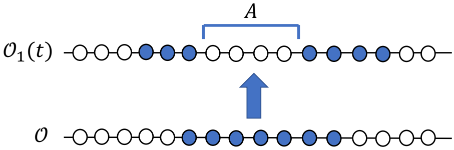
- Consider an operator which initially has support in a small subsystem of a chaotic system.
- The support of this operator increases with time (leads to decay of OTOCs). Susskind, Shenker, Roberts, Stanford



Void formation

Liu and SV, 1912.08918

- Within the growing support of the operator, it has some probability of being equal to the identity in any given region.



- At any time t, with respect to any region A, write $O(t)$ as

$$\mathcal{O}(t) = \mathcal{O}_1(t) + \mathcal{O}_2(t), \qquad \mathcal{O}_1(t) = \tilde{O}_{\bar{A}} \otimes \mathbf{1}_A, \qquad \mathrm{Tr}_A[\mathcal{O}_2(t)] = 0$$

- We refer to the presence of $\mathcal{O}_1(t)$ in $\mathcal{O}(t)$ as void formation from $\mathcal{O}$ in A.

Should expect this probability to be small at late times in chaotic systems, since the support of operators tends to grow.

However, we find that in any many-body system, void formation processes have important consequences:

- Ensuring unitarity of entanglement growth.
- Generation of multipartite entanglement between disjoint regions.
- Explaining differences in entanglement growth between integrable and chaotic systems.

Equilibration and the random void distribution

- Write the initial density matrix as

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{d}\mathbf{1} + \hat{\rho}, \quad \mathrm{Tr}[\hat{\rho}] = 0$$

- If we assumed that the probability of void formation of $\hat{\rho}$ in $\bar{A}$ is negligible at $t \gg t_s$, then we would find

$$S_n^{(A)} = \log d_A, \quad S_n^{(A)} \neq S_n^{(\bar{A})}$$

- Recall that from the equilibrium approximation at infinite temperature, if $d_A \ll d_{\bar{A}}$ or $d_{\bar{A}} \ll d_A$

$$S_n^{(A)} = S_n^{(\bar{A})} = \min(\log d_A, \log d_{\bar{A}})$$

- Thus, to include the contributions from $\tau \neq e$ that lead to a unitary result, from an operator growth perspective we need to take void formation processes into account.

Random void distribution

- In particular, if $\hat{\rho}_1(t)$ is the part of $\hat{\rho}(t)$ with a void in $\bar{A}$, to get

$$S_n^{(A)} = S_n^{(\bar{A})} = \min(\log d_A, \log d_{\bar{A}})$$

we need

$$P_{\hat{\rho},n}^{(\bar{A})} \equiv \frac{\mathrm{Tr}[(\hat{\rho}_1(t))^n]}{\mathrm{Tr}[\hat{\rho}^n]} = \frac{1}{d_{\bar{A}}^{2(n-1)}}, \quad \bar{A} \ll A, \quad n \geq 2$$

- From a calculation in chaotic local random unitary circuits, we find that $P_{O,n}^{(\bar{A})}$ takes this form for any traceless operator O.

- Thermalization to infinite temperature in unitary systems can be seen as a consequence of this general property of operator growth, which we call the “random void distribution.”

Further directions

- Applying the method to other kinds of observables, such as Renyi negativity and Renyi relative entropy. work in progress with J. Kudler-Flam and H. Liu

  Helps understand effects of finite temperature on transfer of information from the black hole to the radiation, and correlations within the radiation.

- Systematically understanding corrections to the approximation from the $\mathcal{Z}_Q$ term, and the time-scale for approaching equilibrium. work in progress with Z.D. Shi and H. Liu

- Can equilibration of pure states can be related to a universal property of operator growth in the finite temperature case?

Thank you!

Example: Infinite temperature

- Taking $\mathcal{I}_\alpha = \mathbf{1}$, and considering the limit where d_A and $d_{\bar{A}}$ are large and comparable,

$$\mathcal{Z}_n^{(A)} = \frac{1}{d^n}\sum_{k=1}^{n} N(n,k) d_A^{n+1-k} d_{\bar{A}}^{k}$$

$$= \frac{1}{d_{\bar{A}}^{n-1}} + \frac{1}{2}\frac{n(n-1)}{d_A d_{\bar{A}}^{n-2}} + \cdots + \frac{1}{2}\frac{n(n-1)}{d_A^{n-2} d_{\bar{A}}} + \frac{1}{d_A^{n-1}}$$

  $N(n,k)$: number of non-crossing partitions of n objects with k blocks/ Narayana numbers.

- By analytic continuation, we can recover Page's result

$$S_1^{(A)} = \begin{cases} \log d_A - \frac{1}{2}\frac{d_A}{d_{\bar{A}}} & d_A < d_{\bar{A}} \\ \log d_{\bar{A}} - \frac{1}{2}\frac{d_{\bar{A}}}{d_A} & d_{\bar{A}} < d_A \end{cases} .$$

- From equilibrium approximation: we can in principle systematically improve this result by including contributions from $\mathcal{Z}_Q$.

Canonical and microcanonical ensemble

- Taking $\mathcal{I}_\alpha = \mathcal{I}_E, \ \mathcal{I}_\beta$, we get universal results for pure states that equilibrate to these ensembles, applicable to any quantum many-body system.
- These results agree with averages over "ergodic bipartition" states Lu and Grover and "canonical thermal pure states" Nakagawa, Watanabe, Sugiura, Fujita.
- Unlike Page's calculation, these averages cannot be seen as averages over time-evolution operators.
- We find the same results as approximations to the time-evolved expressions, and also see how they come from a common structure.
- So far, we have assumed there is some time-scale t_s after which the full system thermalizes. Not true for uncompact systems with local interactions.

  From a simple modification for systems with a sharp light-cone:

$$S_n^{(A)}(t) = s_n^{\mathrm{eq}} \min(|A|, 2t)$$

Approximation for Renyi and von Neumann entropies

- We can further show that

$$\left[\left(\mathcal{Z}_n^{(A)}\right)^m\right]_{\text{eq app}} - \left(\mathcal{Z}_{n,P}^{(A)}\right)^m \ll \left(\mathcal{Z}_{n,P}^{(A)}\right)^m , \quad m = 2, 3, \ldots$$

- Through analytic continuation, we get the equilibrium approximation for the Renyi entropies,

$$\begin{aligned}
[S_n^{(A)}]_{\text{eq app}} &= -\frac{1}{n-1}\lim_{m\to 0}\frac{\partial[(\mathcal{Z}_n^{(A)})^m]_{\text{eq app}}}{\partial m} \\
&\approx -\frac{1}{n-1}\lim_{m\to 0}\frac{\partial(\mathcal{Z}_{n,P}^{(A)})^m}{\partial m} = -\frac{1}{n-1}\log(\mathcal{Z}_{n,P}^{(A)}) \ . \qquad (1)
\end{aligned}$$

- Similarly, for the von Neumann entropy:

$$[S_1^{(A)}]_{\text{eq app}} = -\lim_{n\to 1}\frac{\partial \mathcal{Z}_{n,P}^{(A)}}{\partial n} \ .$$