Black Holes in Cavities and Blinking Islands

Timofei Rusalev, Steklov Mathematical Institute

based on arXiv:2311.16244
with D. Ageev, I. Aref'eva

supported by the Russian Science Foundation, project 20–12–00200

Quarks-2024
21 May, 2024

Main idea and result

- The information paradox for black holes is one of the fundamental problems in the union of gravity and quantum theory, starting with seminal paper [Hawking '76] after the discovery of Hawking radiation [Hawking '74]
- Is the island approach [Almheiri et al. '19, '20, Penington '20] the solution to information paradox?

- Placing a black hole in a cavity (box) is known to be a natural way to study IR scales in gravity, the thermodynamic instability etc. [Hawking '76, Page '80, York '86, ...]
- We consider the dynamics of the entanglement entropy within island prescription for Hawking radiation in the generalization of the two-sided Schwarzschild black hole [Hashimoto-Iizuka-Matsuo '20] by introducing reflective boundaries ($BCFT_2$ techniques)
- We found a universal effect induced by the boundary presence, which we call *"blinking"* island – the disappearance of the island for a finite time interval, that for large cavity leads to non-unitary evolution

The information paradox – a (possible) violation of the unitary evolution of closed systems containing black holes, taking into account quantum effects, where unitary

$$\rho_{\text{tot}}(t_f) = U\rho_{\text{tot}}(t_0)U^\dagger, \qquad UU^\dagger = U^\dagger U = 1$$

Important: unitary evolution implies *pure state* → *pure state*

$$\text{if } \rho_{tot}(t_0) = \rho^2_{tot}(t_0), \text{ then } \rho_{tot}(t_f) = \rho^2_{tot}(t_f) \text{ for all } t_f > t_0$$

We consider a system "black hole + Hawking radiation", which

- at t_0: $\rho_{tot}(t_0) = \rho^2_{tot}(t_0)$ (pure state)
- if at $t_f > t_0$: $\rho_{tot}(t_f) \neq \rho^2_{tot}(t_f)$ (mixed state)

then there is an information paradox (non-unitary evolution)

Introduction: Entanglement Entropy

- Total system $A \cup B$ is described by ρ_{tot} and divided into A and B

$$\mathcal{H}_{\text{tot}} = \mathcal{H}_A \otimes \mathcal{H}_B , \;\; \rho_A = \operatorname*{Tr}_{\mathcal{H}_B} \rho_{tot} , \;\; \rho_B = \operatorname*{Tr}_{\mathcal{H}_A} \rho_{tot}$$

- The entanglement entropies of the subsystems are defined as

$$S(X) \equiv S_{vN}(\rho_X) = -\operatorname{Tr} \rho_X \log \rho_X , \;\; \text{where } X = A,\ B$$

Basic properties

1. $\rho_{tot} = \rho_{tot}^2$ (pure state) $\Leftrightarrow$ $S_{vN}(\rho_{tot}) = 0$ $\Rightarrow$ $S(A) = S(B)$
2. $|S(A) - S(B)| \neq 0$ $\Rightarrow$ $S_{vN}(\rho_{tot}) > 0$ $\Rightarrow$ $\rho_{tot} \neq \rho_{tot}^2$ (mixed state)
3. $S_{vN}(\rho_X) \leq S^{\text{therm}}(\rho_X)$ (thermod. entropy is always BIGGEST)

If $S(A) > S^{\text{therm}}(\rho_B) \Rightarrow |S(A) - S(B)| \neq 0 \Rightarrow \rho_{tot} \neq \rho_{tot}^2$ (mixed state)

Introduction: Bipartition

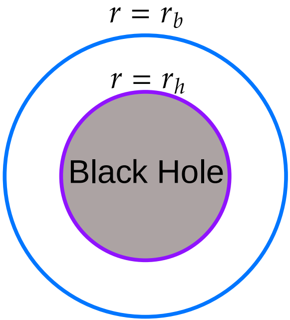
Total system "Hawking radiation (R) + black hole (BH)"

$$\mathcal{H}_{\text{tot}} = \mathcal{H}_{\text{R}} \otimes \mathcal{H}_{BH}$$

Entanglement entropies $S(R)$, $S(BH)$, thermodynamic entropy

$$S_{BH}^{\text{therm}} = \frac{\text{Area(horizon)}}{4G}$$

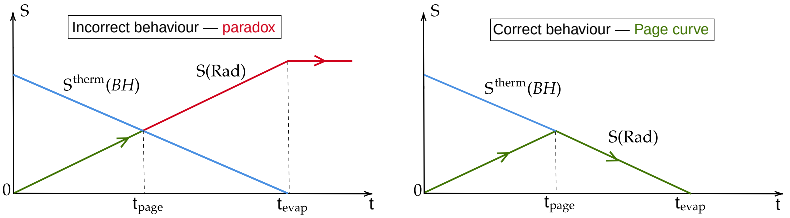
- $r = r_h$ is horizon
- $r = r_b > r_h$ is cutoff after which "gravity is weak"



- $r > r_b$: "Radiation" system
- $r < r_b$: "Black Hole" system

Recall that we assume $\rho_{tot}(t_0) = \rho^2_{tot}(t_0)$. If $\exists\, t_f > t_0$:

$$S_{\text{Rad}}(t_f) > S^{therm}_{BH}(t_f) \Rightarrow |S_{\text{Rad}} - S_{BH}|(t_f) \neq 0 \;\Rightarrow\; \rho_{tot}(t_f) \neq \rho^2_{tot}(t_f)$$



- "Hawking curve": monotonic increase in entropy, exceeding the upper limit at $t > t_{\text{page}}$ $\rightarrow$ unitarity violation
- "Page curve": at t_{page} the growth is replaced by a monotonic decrease to zero entanglement entropy, pure state at t_{evap} $\rightarrow$ unitary behavior

Introduction: island formula

The "island formula" for entanglement entropy of quantum field theory in gravitational systems for $R \subset \Sigma$, where $\mathcal{H}_{tot} = \mathcal{H}_R \otimes \mathcal{H}_{\bar{R}}$, $\bar{R} = \Sigma / R$

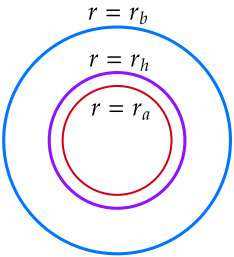
$$S(R) \simeq \min_{\mathcal{I}} \left\{ \operatorname{ext}_{\mathcal{I}} \left[\frac{Area(\partial \mathcal{I})}{4G} + S_{matter}(R \cup \mathcal{I}) \right] \right\}$$

[Almheiri et al. '19, '20, Penington '20, Penington et al. '22]

where

- Σ – spacelike Cauchy surface (e.g. constant time surface $t = \text{const}$)
- $\mathcal{I} \subset \bar{R}$ – "island", defined by extremization, $\partial \mathcal{I}$ – its boundary
- S_{matter} – entanglement entropy of QFT on the classical background

- $r = r_h$ is horizon
- $r = r_b > r_h$ is cutoff after which "gravity is weak"
- $r = r_a$ is boundary of island I



- $r > r_b$: "Radiation" system
- $r < r_a$: Island I

Schwarzschild black hole

- The metric of the four-dimensional Schwarzschild black hole (BH) is

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2, \qquad f(r) = 1 - \frac{r_h}{r}$$

  $r_h = 2GM$ is the horizon, M is BH mass, G is gravitational constant

- Introducing Kruskal coordinates

$$U = -\frac{1}{\kappa_h}\, e^{-\kappa_h(t - r_*(r))}, \quad V = \frac{1}{\kappa_h}\, e^{\kappa_h(t + r_*(r))}$$

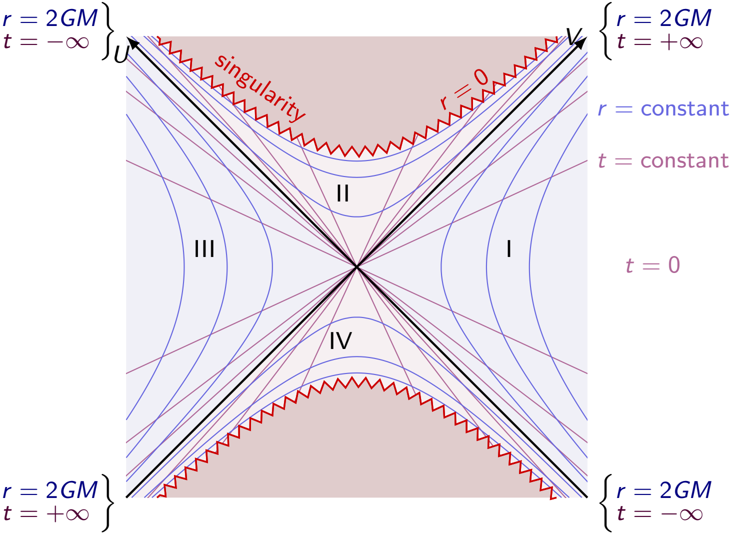
  with the tortoise coordinate $r_*(r) = r + r_h \log|r - r_h|/r_h$ and the surface gravity $\kappa_h = 1/2r_h$, we can rewrite the metric in the form

$$ds^2 = -e^{2\rho(r)} dUdV + r^2 d\Omega^2, \qquad e^{2\rho(r)} = \frac{r_h}{r} e^{-r/r_h}$$

- **Eternal (two-sided) black hole**: the maximal analytic extension of spactime to $-\infty < U, V < \infty$ with constraint $UV < 4r_h^2$

Schwarzschild black hole

Kruskal diagram of Lorentzian Schwarzschild spacetime is



ETERNAL Schwarzschild black hole

- The mass of a black hole is constant, $M = \text{const}$, thermodynamic equilibrium with radiation
- The loss of the black hole due to Hawking radiation is compensated by the flux from infinity (what is it???)
- Thermofield double state (TFD)

$$|\Psi\rangle = \sum_n e^{-\beta E_n/2} |n\rangle_L |n\rangle_R,$$

where $|n\rangle_{L,R}$ are the eigenstates with energy E_n of the matter theory Hamiltonian $H_{L,R}$ in the left/right wedges

- Time evolution is upward in both left and right wedges

$$H_{tot} = H_L + H_R.$$

TFD under H_{tot} is time-dependent!

s-wave approximation (Hashimoto et al. '20)

- The massless field $\Phi(x^\mu)$ on a spherically symmetric background of $d = 4$ spacetime (for instance, Schwarzschild) is decomposed into spherical harmonics: $\Phi(r, t, \theta, \varphi) = \sum_{l,m} Y_{lm}(\theta, \varphi) f_l(r, t)$

- There is a set of effective 2D theories with masses $m^2 \simeq l(l + 1)$

- The lowest harmonic with $l = 0$ (s-mode) corresponds to the effective *massless* 2D theory of matter and it is the largest part of the Hawking radiation away from the horizon

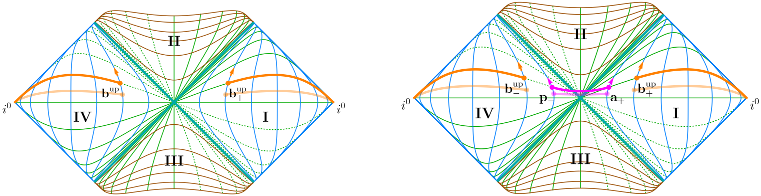
- Discarding $l > 0$ modes, we *assume* that s-mode corresponds to conformal theory (CFT) and entanglement entropy of this theory approximates the entropy of the original 4D problem

s-wave approximation [Hashimoto-Iizuka-Matsuo '20]

We consider CFT_2 on a two-dimensional part of Schwarzschild spacetime

$$ds^2 = -e^{2\rho(r)} dUdV, \qquad e^{2\rho(r)} = \frac{r_h}{r} e^{-r/r_h}$$

Schwarzschild without boundary (Hashimoto et al. '20)

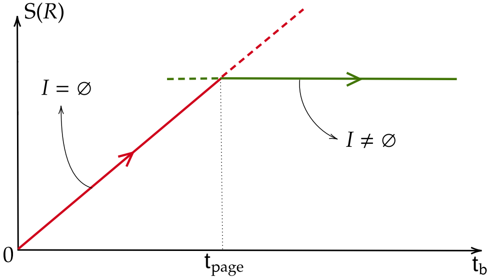


$$S(R) \simeq \min_{\mathcal{I}} \left\{ \operatorname*{ext}_{\mathcal{I}} \left[\frac{Area(\partial \mathcal{I})}{4G} + S_{matter}(R \cup \mathcal{I}) \right] \right\} = \min_{\mathcal{I}} \left\{ \operatorname*{ext}_{\mathcal{I}} S_{gen}(r_a, t_a) \right\}$$

- Entanglement region (orange) is located in both exteriors
  $R = R_- \cup R_+ = \left(i_L^0, b_-\right] \cup \left[b_+, i_R^0\right), \quad b_+ = \{r_b, t_b\}, \; b_- = \{r_b, -t_b\}$
- Island configuration (magenta) extends between both exteriors
  $\mathcal{I} = [a_-, a_+], \quad a_+ = \{r_a, t_a\}, \; a_- = \{r_a, -t_a\}$
- Evolution is upward in both exteriors with $H_{tot} = H_L + H_R$ for TFD
- Parameters (r_a, t_a) of the island $\mathcal{I}$ are determined by extremization

$$\operatorname*{ext}_{\partial \mathcal{I}} S_{gen} : \quad \begin{cases} \partial_{r_a} S_{gen}(r_a, t_a) = 0 \\ \partial_{t_a} S_{gen}(r_a, t_a) = 0 \end{cases}$$

Schwarzschild without boundary (Hashimoto et al. '20)



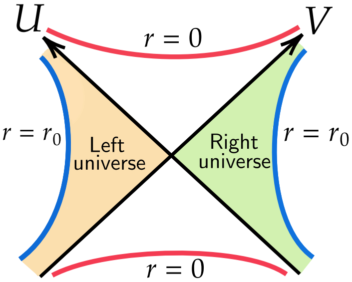
- $\mathcal{I} = \varnothing: \; S_{\mathcal{I}=\varnothing}(R) = \frac{c}{6}\log\left[\frac{16 r_h^2 (r_b - r_h)}{\varepsilon^2 r_b}\cosh^2\frac{t_b}{r_h}\right] \underset{t_b \gg r_h}{\simeq} \frac{c\, t_b}{6 r_h}$ (red)
- $\mathcal{I} \neq \varnothing: \; S_{\mathcal{I}}(R) \simeq \frac{2\pi r_h^2}{G} + \frac{c}{6}\left[\log\left(\frac{16 r_h^3 (r_b - r_h)^2}{\varepsilon^4 r_b}\right) + \frac{r_b - r_h}{r_h}\right]$ (green)

Island rule: at each moment $t = t_b$ we need to choose the *smallest* entropy

Note: $S_{\mathcal{I}}(R) \simeq S_{BH}^{\text{therm}} = 2\pi r_h^2/G$ ***only*** in the approximation $cG/r_h^2 \ll 1$ (neglect of backreaction), island provides consistent behaviour with unitarity

Introducing of boundaries

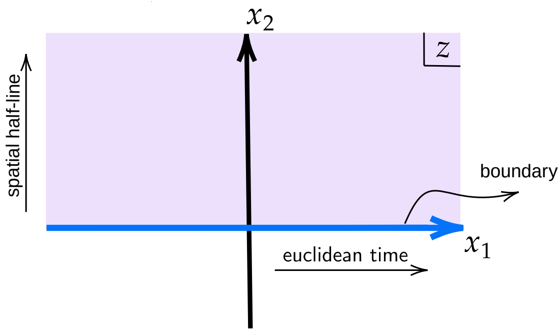
Geometry with a spherically symmetric boundaries:
double-boundary geometry with a boundary at $r = r_0 > r_h$ in the left and right exteriors



- We assume thermodynamic stable equilibrium at all times
- Loss due to Hawking radiation is compensated by reflection from the walls
- Quantum state - generalization of TFD
- Schwarzschild with walls - an analogue of AdS - Schwarzschild?

BCFT$_2$ on upper half-plane

Simplest BCFT$_2$ geometry – Euclidean flat upper half-plane (UHP)



$$ds^2 = dx_1^2 + dx_2^2 = dzd\bar{z}, \quad z = x_1 + ix_2, \quad x_1 \in (-\infty, \infty), \quad x_2 \geq 0$$

Here x_1 is Euclidean time, x_2 is the spatial coordinate, $x_2 = 0$ is boundary

BCFT$_2$ on upper half-plane

- We consider regions R consisting of union of intervals

$$R = [z_{a_1}, z_{b_1}] \cup \ldots \cup [z_{a_m}, z_{b_m}]$$

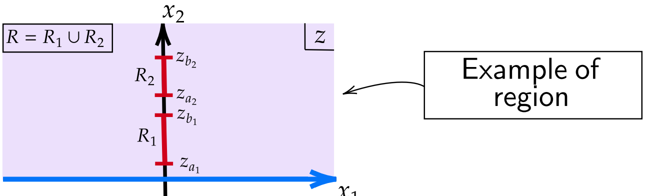
- Entanglement entropy in replica trick framework [Callan, Wilczek '94]

$$S(R) = -\mathrm{Tr}(\rho_R \log \rho_R) = -\lim_{n\to 1} \frac{1}{n-1} \log\left(\mathrm{Tr}\, \rho_R^n\right)$$

  where $\rho_R = \mathrm{Tr}_{\bar{R}}\, \rho_{\mathrm{tot}}$, and ρ_{tot} is vacuum state on UHP

- $$\mathrm{Tr}\, \rho_R^n = \langle \phi(z_{a_1}, \bar{z}_{a_1}) \widetilde{\phi}(z_{b_1}, \bar{z}_{b_1}) \ldots \phi(z_{a_m}, \bar{z}_{a_m}) \widetilde{\phi}(z_{b_m}, \bar{z}_{b_m}) \rangle_{\mathrm{UHP}}^{b.c.}$$

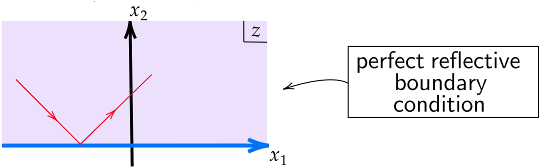
  The twist operators are primary with $h_n = \bar{h}_n = c/24(n - 1/n)$ for bulk and $h_n = \bar{h}_n = 0$ for boundary insertions [Calabrese, Cardy '04]



BCFT$_2$ of free Dirac fermions

We consider c copies of two-dimensional free massless Dirac fermions with perfectly reflecting boundary conditions

$$\psi = \begin{pmatrix} \psi_1(x_1, x_2) \\ \psi_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} \psi_1(z) \\ \psi_2(\bar{z}) \end{pmatrix}; \; \overbrace{\psi_1(x_1, 0) = \psi_2(x_1, 0)}^{\text{b.c.}}, \; x_1 \in (-\infty, \infty)$$



The EE of $R = [z_{a_1}, z_{b_1}] \cup \ldots \cup [z_{a_m}, z_{b_m}]$ [Kruthoff et. al '21, Rottoli et. al '23]

$$S(R) = \frac{c}{3} \sum_{i,j=1}^{m} \log |z_{a_i} - z_{b_j}| - \frac{c}{3} \sum_{i<j}^{m} \log |z_{a_i} - z_{a_j}||z_{b_i} - z_{b_j}| - m \log \varepsilon$$

$$+ \frac{c}{6} \sum_{i,j=1}^{m} \log |z_{a_i} - \bar{z}_{a_j}||z_{b_i} - \bar{z}_{b_j}| - \frac{c}{6} \sum_{i,j=1}^{m} \log |z_{a_i} - \bar{z}_{b_j}||z_{b_i} - \bar{z}_{a_j}|$$

BCFT$_2$ on a curved spacetime

What if the domain 1) is not UHP and 2) is not flat?

Transform correlators using 1) conformal and 2) Weyl transformations

Recall within replica trick the EE is defined by correlator

$$S(R) = -\lim_{n\to 1}\frac{1}{n-1}\log\langle\phi(z_{a_1},\bar{z}_{a_1})\widetilde{\phi}(z_{b_1},\bar{z}_{b_1})\ldots\phi(z_{a_m},\bar{z}_{a_m})\widetilde{\phi}(z_{b_m},\bar{z}_{b_m})\rangle$$

- Conformal map, $z:\Omega\to\text{UHP}$, $\quad z=z(w), \quad \bar{z}=\bar{z}(\bar{w})$

$$\langle\phi(w_1,\bar{w}_1)\ldots\phi(w_m,\bar{w}_m)\rangle_\Omega = \prod_{j=1}^{m}\left(\frac{dz}{dw}\right)^{h_n}\Big|_{w=w_j}\left(\frac{d\bar{z}}{d\bar{w}}\right)^{\bar{h}_n}\Big|_{\bar{w}=\bar{w}_j} \times\langle\phi(z_1,\bar{z}_1)\ldots\phi(z_m,\bar{z}_m)\rangle_{\text{UHP}}.$$

- Weyl map (flat $\to$ curved), $ds^2=dw\,d\bar{w}\to ds^2=e^{2\rho(w,\bar{w})}dw\,d\bar{w}$

$$\langle\phi(w_1,\bar{w}_1)\ldots\phi(w_m,\bar{w}_m)\rangle_{e^{2\rho}g} = e^{-2h_n\rho(w_1,\bar{w}_1)}\ldots e^{-2h_n\rho(w_n,\bar{w}_n)} \times\langle\phi(w_1,\bar{w}_1)\ldots\phi(w_m,\bar{w}_m)\rangle_g$$

Note: it is entropy on a **fixed** curved background (the 2_{nd} term in the island formula)!

Euclidean Schwarzschild spacetime

Euclidean geometry is needed to apply the BCFT_2 EE technique $\Rightarrow$
$\Rightarrow$ let us consider Euclidean Schwarzschild spacetime with boundaries

- Define the (space-) time-like Kruskal coordinates $X, T = (V \mp U)/2$
- Wick rotate Kruskal $T = -i\mathcal{T}$ and Schwarzschild $t = -i\tau$ times
- Euclidean Schwarzschild does not contain an interior, i.e. $r \geq r_h$

We introduce a complex coordinates

$$w = X + i\mathcal{T}, \qquad \bar{w} = X - i\mathcal{T},$$

and the Euclidean version of the 2D part of Kruskal metric is

$$ds^2 = e^{2\rho(w,\bar{w})} dw d\bar{w}, \qquad \overbrace{e^{2\rho(w,\bar{w})}}^{\text{Weyl factor}} = \frac{W(e^{-1}\kappa_h^2 w\bar{w})}{\kappa_h^2 w\bar{w}\left[1 + W(e^{-1}\kappa_h^2 w\bar{w})\right]},$$

where $W(x)$ is Lambert W function

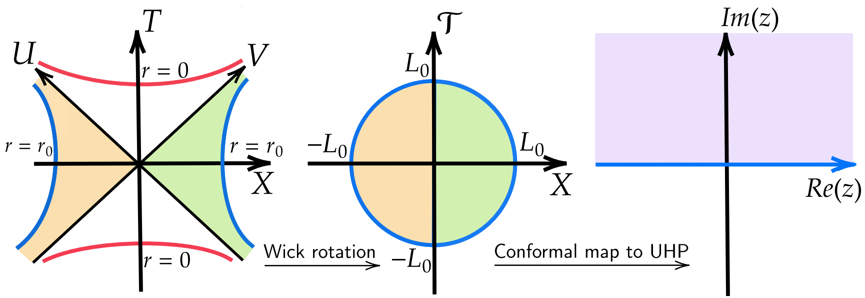
Euclidean double-boundary geometry

- Euclidean double-boundary geometry is the interior of the disc

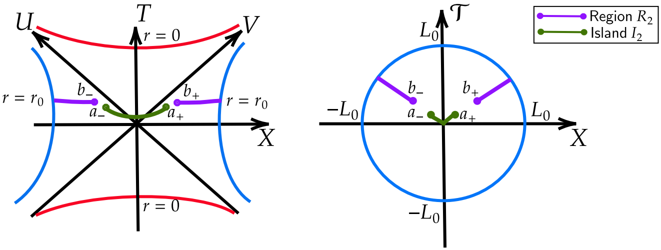
$$\{X^2 + \mathcal{T}^2 \leq L_0^2 \,|\, X, \mathcal{T} \in [-L_0, L_0]\}, \quad L_0 = \frac{e^{\kappa_h r_*(r_0)}}{\kappa_h}$$

- Conformal map from disc to UHP is

$$z = i\frac{L_0 + w}{L_0 - w}$$



Double-boundary entanglement entropy setup



- Entanglement region (purple) is located in both exteriors
  $R_2 = R_- \cup R_+ = [\mathbf{r}_0^-, b_-] \cup [b_+, \mathbf{r}_0^+], \quad b_+ = \{r_b, t_b\},\ b_- = \{r_b, -t_b\}$
- Island configuration (green) extends between both exteriors
  $\mathcal{I}_2 = [a_-, a_+], \quad a_+ = \{r_a, t_a\},\ a_- = \{r_a, -t_a\}$
- Evolution is upward in both exteriors with $H_{tot} = H_L + H_R$
- Parameters (r_a, t_a) of the island $\mathcal{I}$ are determined by extremization

$$\underset{\partial I}{\text{ext}}\, S_{gen} : \begin{cases} \partial_{r_a} S_{gen}(r_a, t_a) = 0 \\ \partial_{t_a} S_{gen}(r_a, t_a) = 0 \end{cases}$$

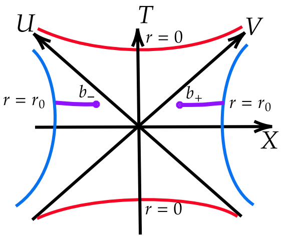
- Entanglement entropy with trivial island $\mathcal{I} = \varnothing$

$$S(R_2) = \frac{c}{6}\log\left(\frac{4f(r_b)\cosh^2\kappa_h t_b}{\kappa_h^2\varepsilon^2}\right) + \frac{c}{6}\log\left(\frac{2\sinh^2\kappa_h(r_*(r_0)-r_*(r_b))}{\cosh 2\kappa_h(r_*(r_0)-r_*(r_b)) + \cosh 2\kappa_h t_b}\right)$$

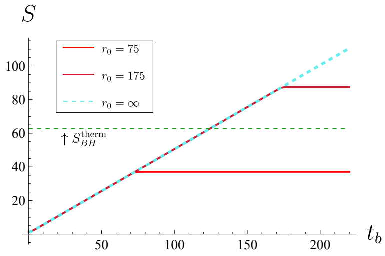
- Entropy saturates at $t_b \gg t_b^1,\ t_b^1 \equiv r_*(r_0) - r_*(r_b)$ at value

$$S^{\text{sat}}(R_2) = \frac{c}{6}\log\left(\frac{4f(r_b)\sinh^2\kappa_h(r_*(r_0)-r_*(r_b))}{\kappa_h^2\varepsilon^2}\right)$$

- There is such $\widetilde{r}_0(r_b, c)$ that
  - $\forall r_0 > \widetilde{r}_0$: $S^{\text{sat}}(R_2) > S_{BH}^{\text{therm}}$ (non-unitary evolution)
  - $\forall r_0 \in (r_h, \widetilde{r}_0)$: $S^{\text{sat}}(R_2) < S_{BH}^{\text{therm}}$ (consistent with unitarity)



$r_h = 1,\ r_b = 5,\ c = 3, G = 0.1$



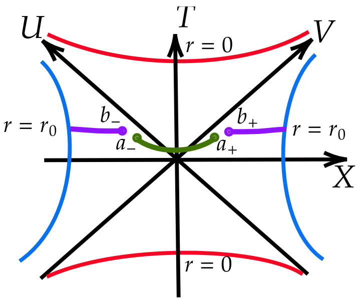
Double-boundary: blinking island

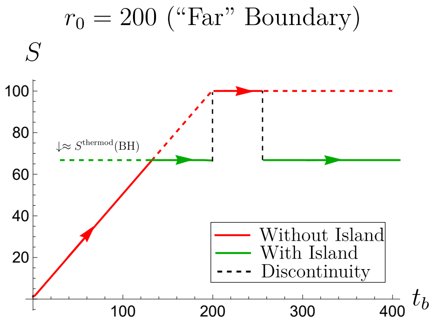
The general picture of the island evolution is

- First the island appears and at $t_b^1 = r_*(r_0) - r_*(r_b)$ disappears
- At $t_b^2 = r_*(r_0) + 2r_*(r_b) - 3/2\kappa_h \log\left(cG\kappa_h^2 e/3\pi\right)$ it appears again
- So, at $t_b \in (t_b^1, t_b^2)$ island solution near horizon *does not exist*

There is a **"blinking island"** in the approximation $cG\kappa_h^2 \ll 1$

$$t_{\text{blink}} \equiv t_b^2 - t_b^1 = 3r_*(r_b) - 3/2\kappa_h \log\left(cG\kappa_h^2 e/3\pi\right) > 0$$



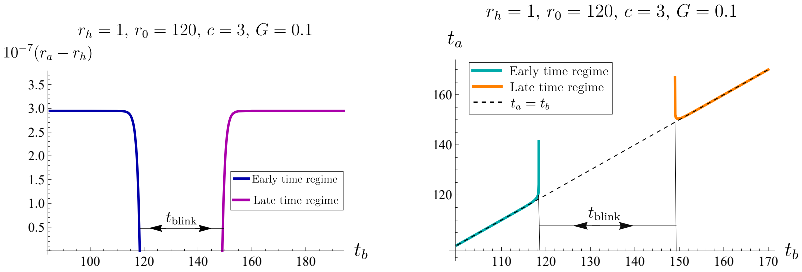


Properties of blinking island

- For $r_b/2GM$ = fixed and c = fixed, with $M \nearrow$ the time $t_{\text{blink}} \nearrow$

$$t_{\text{blink}} = 6GM\left(\frac{r_b}{2GM} - \log\left[\frac{ce}{48\pi GM^2(r_b/2GM-1)}\right]\right)$$

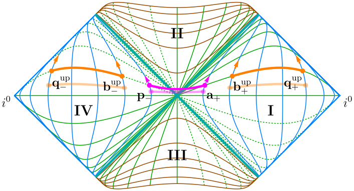
- Behavior of island's parameters (r_a, t_a) – loss of equilibrium?

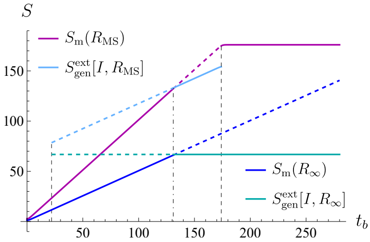


- Thermodynamic stability/instability depends on cavity size r_0 [Hawking '76, York '86]. Is there a connection with the "blinking effect"?

Finite-size effects

- EE of finite-size region [Ageev, Aref'eva, Belokon, Ermakov, Pushkarev, TR '23 PRD arXiv:2209.00036], for dS [Ageev, Aref'eva, Belokon, Pushkarev, TR arXiv:2304.12351]





- The island disappears ⇒ a jump in EE and non-unitary evolution

Hypothesis

Finite-size effects (IR regularization) lead to the problems with island?

<u>What is the reason</u>: massive graviton [Geng, Karch '20], thermodynamic instability, inapplicability of s-mode, ...? Direction of further research

Conclusions

1. We considered the dynamics of the entanglement entropy within island prescription for Hawking radiation in the generalization of the two-sided Schwarzschild black hole by introducing reflective boundaries
2. We showed a universal effect induced by the boundary presence, which we call *"blinking"* island – the disappearance of the island for a finite time interval, that for large cavity leads to non-unitary evolution
3. While for small sizes of cavity there is an evolution consistent with unitarity even without considering the island configuration

Thank you for attention!