Canonical Purification and the Quantum Extremal Shock

Onkar Parrikar
Tata Institute of Fundamental Research, Mumbai

(Based on work in progress with Vivek Singh)

12th Jo'burg workshop on String Theory

December 05, 2022

Introduction

- Consider a general, full-rank, bi-partite state Ψ in the (for the moment, finite dimensional) tensor product Hilbert space $\mathcal{H}_L \otimes \mathcal{H}_R$.
- Any such state can always be written in the form:

$$|\Psi\rangle = \sum_n \sqrt{p_n}\, |\tilde{\chi}_n\rangle_L \otimes |\chi_n\rangle_R,$$

  where p_n are the eigenvalues of the reduced density matrices, χ_n are the eigenstates of the reduced density matrix on the right, and $\tilde{\chi}_n$ are the eigenstates of the reduced density matrix on the left.
- While the eigenvalues are common to both the parties, the eigenstates are not.

- On the other hand, suppose we were only given access to the left party.
- We could write down a purification, called the *canonical purification* of this reduced density matrix as:

$$|\Psi^\star\rangle := \sum_n \sqrt{p_n}\,|\tilde{\chi}_n\rangle_L \otimes |\tilde{\chi}_n^\star\rangle_{L^\star}.$$

  Here $|\tilde{\chi}_n^\star\rangle = \Theta|\tilde{\chi}_n\rangle$ where Θ is an anti-unitary operator on L (for instance, CPT).
- $\Psi^\star$ resembles the thermofield double state.
- We can think of $\Psi^\star$ as the "simplest" purification that one could build only given access to the left party.

- Since Ψ and $\Psi^\star$ are two different purifications of the same density matrix, they are related by a unitary transformation on the right factor:

$$|\Psi^\star\rangle := \mathcal{R}_\Psi|\Psi\rangle$$

$$\mathcal{R}_\Psi : \mathcal{H}_R \to \mathcal{H}_{L^\star}, \quad \mathcal{R}_\Psi := \sum_n |\tilde{\chi}^\star_n\rangle_{L^\star}\langle\chi_n|_R,$$

- The operator $\mathcal{R}_\Psi$ quantifies the complexity of reconstructing the original state Ψ from $\Psi^\star$.
- It goes beyond entanglement in quantifying properties of the state Ψ.
- We will call $\mathcal{R}_\Psi$ the *reflection operator* w.r.t the left party.

Motivation

- In holographic CFTs, it was shown by Harlow that for a bulk degree of freedom within the entanglement wedge of a boundary subregion A, the encoding map V into the dual CFT takes the general form [Harlow '16]:

$$|\psi_i\rangle_{\mathrm{CFT}} = V|i\rangle_{\mathrm{bulk}} = U_A|i\rangle_{A_1} \otimes |\chi\rangle_{A_2,\bar{A}},$$

  where $A = A_1 \otimes A_2 \oplus A_3$, with A_1 being the same dimension as that of the code subspace.

- Harlow's structure theorem is a general consequence of the Ryu-Takayanagi formula.

Motivation

- We can think of the unitary U_A appearing in Harlow's structure theorem as a reflection operator: introduce an auxiliary reference system ref which has the same dimension as that of the code subspace, and consider the maximally entangled state:

$$|\Psi\rangle = \frac{1}{\sqrt{d_{\text{code}}}} \sum_i |i\rangle_{\text{ref}} \otimes |\psi_i\rangle_{\text{CFT}}.$$

Then, the unitary U_A in Harlow's theorem is precisely the reflection operator with respect to $\text{ref} \cup \bar{A}$.

Motivation 2

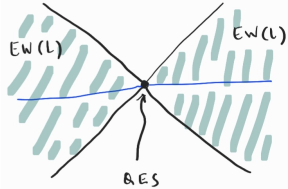
- A second motivation comes from the fact that the reflected entropy for a mixed two-party state ρ_{AB} is defined in terms of the canonical purification $\Psi^{\star}_{ABA^{\star}B^{\star}}$ as the entanglement entropy of $AA^{\star}$. [Dutta, Faulkner '19]
- The reflected entropy has a natural geometric interpretation in AdS/CFT in terms of the entanglement wedge cross section.

Motivation 3

- Finally, it was argued in [Engelhardt, Folkestad '22] that for a black hole evaporating into a non-gravitational bath, the canonical purification of the total state with respect to the black hole side is dual to a connected wormhole, thus realzing the ER=EPR idea in the context of an evaporating black hole.
- While the original state of the radiation plus the evaporating black hole does not appear to have a wormhole in it, the state after the action of the corresponding $\mathcal{R}_\Psi$ does.
- In this way, the unitary $\mathcal{R}_\Psi$ in this case acts to geometrize the entanglement in the originally complex and non-geometric state.

Holographic Engelhardt-Wall prescription

- For holographic CFTs, it was proposed in [Engelhardt, Wall '18] that the classical, Lorentzian bulk geometry dual to the canonical purification w.r.t a boundary subregion L is obtained by taking the entanglement wedge of L and gluing it to its CPT image at the classical extremal surface.



- In gluing together solutions of Einstein equations, one must impose junction conditions at the gluing surfaces. When the co-dimension two surface we are gluing across is a classically extremal surface,

$$\theta_{\pm} = 0$$

these junction conditions are trivially satisfied.

Holographic Engelhardt-Wall prescription

- However, upon including quantum corrections, the gluing must be done across the quantum extremal surface (QES). [Bousso et al, '19]
- Due to quantum corrections, the QES is not generically classically extremal:

$$\theta_{\pm} + \frac{\delta S_{\text{bulk}}}{\delta x^{\pm}} = 0.$$

- In this case, the junction conditions imply that the state of the bulk matter dual to the canonical purification must have a stress tensor with a delta function "shock", in order for Einstein's equations to be satisfied:

$$\langle T^{\text{bulk}}_{++}(x^+, x^- = 0)\rangle_{\Psi^\star} = \frac{1}{\pi}\delta(x^+)\frac{\delta S_{\text{bulk}}}{\delta x^+}.$$

Our Goal

- Our primary goal will be to verify the above prediction of the Engelhardt-Wall construction in a perturbative setup.
- We will work to first order in perturbation theory around the TFD state/eternal black hole, and show the presence of this quantum extremal shock.
- Our arguments can also be generalized beyond perturbation theory with some mild assumptions.
- Since the EW prediction follows from Einstein's equation, we are seeing here the emergence of Einstein's equation from the boundary entanglement structure, in a context where bulk quantum corrections are important.

Flow equation for $\mathcal{R}_\Psi$

- Let us say that we have a general one-parameter family of states $\Psi_\lambda \in \mathcal{H}_L \otimes \mathcal{H}_R$ which are all full rank.
- At any value of λ, we can construct the reduced density matrices $\rho_L(\lambda)$ and $\rho_R(\lambda)$ corresponding to the left and right factors respectively.
- Accordingly, we have the one-parameter family of modular Hamiltonians $K_L(\lambda)$ and $K_R(\lambda)$, where the modular Hamiltonian for a density matrix ρ is defined as

$$K = -\log \rho.$$

Flow equation for $\mathcal{R}_\Psi$

- At any given value of λ, we define the modular eigenvalues and eigenstates as

$$K_R(\lambda)|\chi_n(\lambda)\rangle_R = E_n(\lambda)|\chi_n(\lambda)\rangle_R,$$

$$K_L(\lambda)|\tilde{\chi}_n(\lambda)\rangle_L = E_n(\lambda)|\tilde{\chi}_n(\lambda)\rangle_L,$$

  where note that the eigenvalues are common to both sides.
- Our first goal is to derive a differential equation for

$$\mathcal{R}_\lambda = \sum_n |\tilde{\chi}_n^\star\rangle_{L^\star}\langle\chi_n|_R$$

  along the flow parametrized by λ.

Flow equation for $\mathcal{R}_\Psi$

- Upon an infinitesimal deformation of the parameter λ, the change in the eigenstates of, say K_R, is given by (assuming non-degeneracy)

$$\frac{d}{d\lambda}|\chi_n\rangle_R = \sum_{m\neq n} \frac{\langle\chi_m|\frac{d}{d\lambda}K_R|\chi_n\rangle_R}{(E_n(\lambda)-E_m(\lambda))}|\chi_m\rangle_R.$$

We can rewrite this in the following way:

$$\begin{aligned} &= \sum_{m\neq n}\int_0^\infty idt\, e^{-\epsilon t}\left(\langle\chi_m|e^{itK_R(\lambda)}\frac{d}{d\lambda}K_R e^{-itK_R(\lambda)}|\chi_n\rangle_R\right)|\chi_m\rangle_R \\ &= \int_0^\infty idt\, e^{-\epsilon t} e^{itK_R(\lambda)}\frac{d}{d\lambda}K_R e^{-itK_R(\lambda)}|\chi_n\rangle_R - \frac{i}{\epsilon}\frac{d}{d\lambda}E_n(\lambda)\,|\chi_n\rangle_R. \end{aligned}$$

where we have introduced a regulator $\epsilon\to 0^+$

Flow equation for $\mathcal{R}_\Psi$

- So we have

$$\frac{d}{d\lambda}|\chi_n\rangle_R = i\mathcal{A}_R|\chi_n\rangle_R, \quad \frac{d}{d\lambda}|\tilde{\chi}_n\rangle_L = i\mathcal{A}_L|\tilde{\chi}_n\rangle_L,$$

where

$$\mathcal{A}_R(\lambda) = a_R(\lambda) + \int_0^\infty dte^{-\epsilon t}\, e^{itK_R^{(\lambda)}}\frac{d}{d\lambda}K_R^{(\lambda)}e^{-itK_R^{(\lambda)}},$$

$$\mathcal{A}_L(\lambda) = a_L(\lambda) + \int_0^\infty dte^{-\epsilon t}\, e^{itK_L^{(\lambda)}}\frac{d}{d\lambda}K_L^{(\lambda)}e^{-itK_L^{(\lambda)}}.$$

Modular Berry connections

- It is natural to interpret $\mathcal{A}$ as connection one-forms for a $\boldsymbol{U}(\dim \mathcal{H}_L) \times \boldsymbol{U}(\dim \mathcal{H}_R)$ bundle over parameter space.
- To see this more explicitly, imagine that we consider a modified state $\Psi' = U\Psi$, where U is a one-sided unitary transformation acting on R, but we can let U depend on the parameters λ.
- Then, it follows from a short calculation that the connections transform as

$$\mathcal{A}'_L = \mathcal{A}_L,$$

$$\mathcal{A}'_R = U\,\mathcal{A}_R\,U^{-1} - idU\,U^{-1},$$

  which is precisely the transformation property of a connection 1-form. The same formula is also true for the transformation of $\mathcal{A}_L$ under a one-sided unitary acting on L.

Flow equation for $\mathcal{R}_\Psi$

- Using this, we can work out the flow equation for $\mathcal{R}_\lambda$:

$$\frac{d}{d\lambda}\mathcal{R}_\lambda = \sum_n \left(\frac{d}{d\lambda} |\tilde{\chi}_n^\star\rangle_{L^\star} \langle \chi_n|_R + |\tilde{\chi}_n^\star\rangle_{L^\star} \frac{d}{d\lambda} \langle \chi_n|_R \right)$$

- In terms of the modular Berry connections, we get

$$i\frac{d}{d\lambda}\mathcal{R}_\lambda = \mathcal{A}^\star_{L^\star}(\lambda)\,\mathcal{R}_\lambda + \mathcal{R}_\lambda\,\mathcal{A}_R(\lambda).$$

where $\mathcal{A}^\star_{L^\star} = \Theta^{-1}\mathcal{A}_L\Theta$.

Solution to flow equation for $\mathcal{R}_\Psi$

- The general solution to this differential equation takes the form:

$$\mathcal{R}_\lambda = U_{L^\star}(\lambda) \cdot \mathcal{R}_0 \cdot U_R(\lambda),$$

$$U_{L^\star} = \mathcal{P} \exp\left\{-i \int_0^\lambda d\lambda' \mathcal{A}^\star_{L^\star}(\lambda')\right\},$$

$$U_R = \mathcal{P} \exp\left\{-i \int_0^\lambda d\lambda' \mathcal{A}_R(\lambda')\right\},$$

where $\mathcal{P}$ stands for path-ordering.

- The operators U_R and $U_{L^\star}$ are somewhat reminiscnet of Connes cocyles.

Perturbation theory

- Now we will specialize to perturbation theory around the TFD state:

$$|\Psi_0\rangle = \frac{1}{\sqrt{Z}} \sum_n e^{-\frac{\beta}{2} E_n(0)} |\chi_n(0)\rangle_L \otimes |\chi_n^\star(0)\rangle_R,$$

  where $E_n(0)$ and $\chi_n(0)$ are the eigenstates of some local Hamiltonian H.
- The TFD state can also be thought of as a Euclidean path integral over a Euclidean time segment of length $\beta/2$.
- The corresponding operator $\mathcal{R}_0$ is given by

$$\mathcal{R}_0 = \sum_n |\chi_n^\star(0)\rangle_{L^\star} \langle \chi_n^\star(0)|_R,$$

  and the corresponding canonical purification $\Psi_0^\star$ is essentially the same state, but with the right subsystem re-labelled as $L^\star$.

Perturbation theory

- We wish to consider a one-parameter deformation of the TFD state.
- A natural such family of states can be constructed by turning on a source $J(\tau)$ for some operator $O(\tau)$ in the Euclidean path integral.
- Concretely, we change the action inside the Euclidean path integral in the following way:

$$S_{\text{new}} = S_{\text{old}} + \lambda \int_{-\beta/2}^{0} d\tau J(\tau) O(\tau),$$

  where I have only written out the time integrals, leaving the space integrals implicit.
- This new path integral now constructs a new bi-partite state which we will call Ψ_λ. We wish to construct the operator $\mathcal{R}_\lambda$ for this family of states to first order in λ.

First order in Perturbation theory

- The derivative of the modular Hamiltonian under such a deformation is given by [Faulkner, Leigh, Parrikar, Wang '16, Balkrishnan, Parrikar '20]

$$\frac{dK_R}{d\lambda} = \int_0^{2\pi} d\tau\, J_R(\tau) \int_{-\infty}^{\infty} \frac{ds}{4\sinh^2(\frac{s+i\tau}{2})} e^{\frac{is}{2\pi}K_R(0)} O(0) e^{-\frac{is}{2\pi}K_R(0)}.$$

  Here $K_R(0)$ is the undeformed modular Hamiltonian for Ψ_0.
- $J_R(\tau)$ is a time-reflection symmetric version of $J(\tau)$:

$$J_R(t) = \begin{cases} J(\tau) & -\beta/2 < \tau < 0 \\ J^*(-\tau) & 0 < \tau < \beta/2. \end{cases}$$

- A similar formula can also be written for the left subsystem. The only difference is that the corresponding source J_L is related to J_R by a left-right reflection

$$J_L(\tau) = J_R(\beta/2 - \tau).$$

First order in Perturbation theory

- The previous formula essentially follows from the following:

$$\frac{d}{d\lambda}\log\rho = \int_{-\infty}^{\infty}\frac{ds}{4\sinh^2(\frac{s}{2})}\rho^{\frac{is}{2\pi}}\rho^{-1}\frac{d\rho}{d\lambda}\rho^{-\frac{is}{2\pi}},$$

which you can prove by re-summing the relevant terms in the BCH formula, together with the fact that

$$\rho^{-1}\frac{d\rho}{d\lambda} = \int d\tau J(\tau)O(\tau).$$

First order in Perturbation theory

- We can now obtain the first order change in U_R:

$$-i\frac{dU_R}{d\lambda}(0) = \mathcal{A}_R(0),$$

where recall

$$\mathcal{A}_R(0) = a_R(0) + \int_0^\infty dt e^{-\epsilon t}\, e_R^{itK^{(0)}} \frac{dK_R}{d\lambda} e^{-itK_R^{(0)}}.$$

- Substituting the expression for $\frac{dK_R}{d\lambda}$ and performing the t integral, and we get

$$\frac{dU_R}{d\lambda}(0) = a_R(0) + \frac{1}{2\pi i}\int d\tau J_R(\tau) \int_{-\infty}^{\infty} ds \frac{1}{\left(1 - e^{-(s+i\tau)}\right)} O(s),$$

$$O(s) = e^{isK_R^{(0)}} O e^{-isK_R^{(0)}}.$$

- We can also derive a "gravitational" formula for the modular Berry connection...

Gravity dual and the quantum extremal shock

- So far, we have worked out the first order change in the unitary $\mathcal{R}_\lambda$.
- From here, we can also work out the first order change in the canonical purification $\Psi_\lambda^\star = \mathcal{R}_\lambda \Psi_\lambda$.
- To see the quantum extremal shock, our goal is to compute the bulk stress tensor at first order in the state $\Psi_\lambda^\star$.
- We need to turn on an operator O in the Euclidean path-integral which sources the bulk stress tensor at $O(\lambda)$.
- For this purpose, we cannot take O to be a single-trace operator, as single trace operators only source the bulk stress tensor at $O(\lambda^2)$.
- Instead, we can consider a double-trace operator $O = :\phi\phi:$, for some single trace operator ϕ; although the details of what O we choose will not be relevant in the discussion below.

Gravity dual

- So in order to proceed, we wish to compute the bulk stress tensor in the canonically-purified state.
- We will compute the one-point function of a bulk operator Φ:

$$\langle \Phi \rangle_{\Psi_\lambda^\star} = \langle \Psi_\lambda | \mathcal{R}_\lambda^\dagger \, \Phi \, \mathcal{R}_\lambda | \Psi_\lambda \rangle,$$

  where we will take the bulk location x_B of this operator to be in the entanglement wedge of $L^\star$ in the geometry dual to $\Psi_\lambda^\star$. Eventually, we are interested in taking the limit where x_B approaches the QES.
- Note that the backreaction from turning on a double-trace operator is of $O(\lambda G_N)$ and we can ignore this effect for now. So, the classical bulk spacetime dual to the canonically purified state is the undeformed, eternal black hole spacetime. Here we are interested in computing the deformation of the bulk quantum state.

Gravity dual

- We now compute the first order change in $\langle\Phi\rangle_{\Psi_\lambda^\star}$:

$$\begin{aligned}\frac{d}{d\lambda}\langle\Phi\rangle_{\Psi_\lambda^\star} &= \langle\Psi_\lambda|\frac{d\mathcal{R}_\lambda^\dagger}{d\lambda}\,\Phi\,\mathcal{R}_\lambda|\Psi_\lambda\rangle + \langle\Psi_\lambda|\mathcal{R}_\lambda^\dagger\Phi\,\frac{d\mathcal{R}_\lambda}{d\lambda}|\Psi_\lambda\rangle \\ &+ \mathrm{Tr}_R\left(\frac{d\rho_R^{\Psi_\lambda}}{d\lambda}\mathcal{R}_\lambda^\dagger\,\Phi\,\mathcal{R}_\lambda\right).\end{aligned}$$

- Using the flow equation for $\mathcal{R}_\lambda$, we can rewrite this as

$$= i\langle\Psi_\lambda|\,[\mathcal{A}_R,\Phi_{\mathcal{R}}]\,|\Psi_\lambda\rangle + i\langle\Psi_\lambda^\star|\,[\mathcal{A}_{L^\star}^\star,\Phi]\,|\Psi_\lambda^\star\rangle + \delta_{J_R}\langle\Phi^\star\rangle,$$

where we have defined

$$\Phi_{\mathcal{R}} = \mathcal{R}_\lambda^\dagger\,\Phi\,\mathcal{R}_\lambda.$$

Gravity dual

- Now we evaluate this at $\lambda = 0$. Let us focus on one of the terms above, say, $\langle\Psi_0| [\mathcal{A}_R, \Phi] |\Psi_0\rangle$

$$
\begin{aligned}
&= \mathrm{Tr}_R \left(\rho_R^{(0)} [\mathcal{A}_R, \Phi(x_B)]\right) \\
&= \frac{1}{2\pi i} \int d\tau J_R(\tau) \int_{-\infty}^{\infty} \frac{ds}{\left(1 - e^{-(s+i\tau)}\right)} \mathrm{Tr}_R \left(\rho_R^{(0)} [O(s), \Phi]\right) \\
&= \frac{1}{2\pi i} \int d\tau J_R(\tau) \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \frac{ds}{\left(1 - e^{-(s+i\tau)}\right)} \mathrm{Tr}_R \left(\rho_R^{(0)} O(s)\Phi\right) \\
&- \frac{1}{2\pi i} \int d\tau J_R(\tau) \int_{-\infty-i(2\pi-\epsilon)}^{\infty-i(2\pi-\epsilon)} \frac{ds}{\left(1 - e^{-(s+i\tau)}\right)} \mathrm{Tr}_R \left(\rho_R^{(0)} O(s)\Phi\right)
\end{aligned}
$$

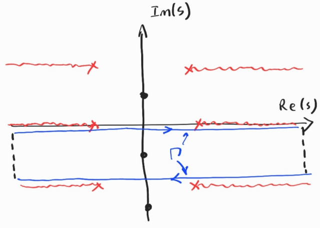
where note that the a_R term has dropped out because it commutes with $\rho^{(0)}$.

Gravity dual

- So we conclude that

$$\langle [\mathcal{A}_R, \Phi] \rangle_{\Psi_0} = \frac{1}{2\pi i} \int d\tau \, J_R \int_\Gamma \frac{ds}{\left(1 - e^{-(s+i\tau)}\right)} \mathrm{Tr}_R \left(\rho_R^{(0)} O(s) \Phi \right),$$

where the contour Γ is the union of the two horizontal contours at $\mathrm{Im}(s) = -\epsilon$ and $\mathrm{Im}(s) = -(2\pi - \epsilon)$.



- Using Cauchy's theorem, we can then rewrite this integral as the sum over three contributions: the pole at $s = -i\tau$, and the two "vertical" contours at $\mathbb{Re}(s) = \pm\Lambda$ (with $\Lambda \to \infty$).

Gravity dual

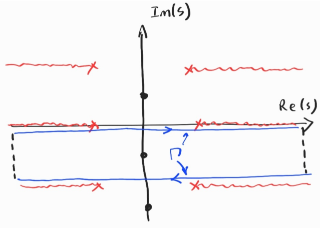
- The pole contribution is given by

$$\langle[\mathcal{A}_R, \Phi]\rangle_{\Psi_0}\Big|_{\text{pole}} = -\int d\tau J_R(\tau)\mathrm{Tr}_R\left(\rho_R^{(0)} O(\tau)\Phi\right) = -\delta_{J_R}\langle\Phi\rangle.$$

- This term cancels the third term above.
- The canonical purification does not know about the entanglement wedge of R. The third term comes from the change in the entanglement wedge of R; so this should cancel out and get replaced with the appropriate contribution coming from the change in the entanglement wedge of L.
- A similar calculation to the above with the $\mathcal{A}_{L^\star}$ term does precisely this!

Vertical contours

- However, the vertical contours at infinity cannot simply be dropped, especially when $\Phi = T^{(\text{bulk})}_{\pm\pm}$, and precisely when the operator approaches the QES.



- Consider, for instance, the vertical contour at $s = -\Lambda$:

$$
\begin{aligned}
I_{-} &= \frac{1}{2\pi}\int d\tau\, J_R \int_{\epsilon}^{2\pi-\epsilon} \frac{d\theta}{\left(1 - e^{(\Lambda - i\tau)} e^{i\theta}\right)} \\
&\times \quad \mathrm{Tr}_R \left(\rho_R^{(0)} O e^{i(\Lambda+i\theta)K_R^{(0)}}\, T^{\text{bulk}}_{++}(x^+, x^-) e^{-i(\Lambda+i\theta)K_R^{(0)}}\right)
\end{aligned}
$$

- Naively, in the $\Lambda \to \infty$ limit, it seems like this integral should vanish. But this is too quick!

Vertical contours

- Using the equality between boundary modular flow and bulk modular flow, we can write

$$e^{isK_R^{(0)}} T_{++}^{\text{bulk}}(x^+, x^-) e^{-isK_R^{(0)}} = e^{2s} T_{++}^{\text{bulk}}(x^+ e^s, x^- e^{-s}).$$

- When $x^+ > 0$, the operator gets boosted off to infinity, and so the correlator should vanish.
- However, when $x^+ = 0$, the e^{2s} factor makes the correlator diverge. In fact, this is a delta function divergence.
- It is easy to extract the coefficient of the delta function by integrating in x^+.
- This precisely produces the half-sided ANEC operator in the bulk, which then precisely gives the $\frac{\delta S_{\text{bulk}}}{\delta x^+}$ contribution.

Summary

- The operator $\mathcal{R}_\Psi$ goes beyond standard entanglement measures.
- For a one-parameter family of states Ψ_λ, we obtained a flow equation for $\mathcal{R}_\lambda$.
- We used this flow equation in a perturbative setting to derive the quantum extremal shock expected in the dual bulk geometry.
- Our arguments for the presence of this shock can also be extended to finite λ, up to some mild assumptions (such as locality of modular flow close to the entanglement cut).