On the interior geometry of a typical black hole microstate

Jan de Boer, Rik van Breukelen, Sagar F. Lokhande, Kyriakos Papadodimas and Erik Verlinde

Abstract: We argue that the region behind the horizon of a one-sided black hole can be probed by an analogue of the double-trace deformation protocol of Gao-Jafferis-Wall. This is achieved via a deformation of the CFT Hamiltonian by a term of the form $\mathcal{O}_e \mathcal{O}$, where $\mathcal{O}$ denote the state-dependent “mirror operators”. We argue that this deformation creates negative energy shockwaves in the bulk, which allow particles inside the horizon to escape and to get directly detected in the CFT. This provides evidence for the smoothness of the horizon of black holes dual to typical states. We argue that the mirror operators allow us to perform an analogue of the Hayden-Preskill decoding protocol. Our claims rely on a technical conjecture about the chaotic behavior of out-of-time-order correlators on typical pure states at scrambling time.

Keywords: AdS-CFT Correspondence, Black Holes in String Theory

ArXiv ePrint: 1804.10580
1 Introduction

The black hole information paradox is related to the question of smoothness of the black hole horizon \([1, 2]\). The latter question becomes particularly sharp for typical CFT states dual to a large black hole in AdS. It is challenging to reconcile the smoothness of their horizon to unitarity of the dual CFT \([3, 4]\), even though these black holes do not evaporate. In \([5–7]\) it was argued that these problems can be resolved by describing the space-time behind the horizon using state-dependent CFT operators, which are partly selected by their entanglement with fields in the exterior. A related proposal from a somewhat different perspective was described in \([8, 9]\). It remains a challenge to fully understand the geometry dual to a typical black hole microstate.

In \([10]\), it was realized how to probe the horizon of a two-sided eternal AdS black hole by using double-trace deformations of the CFT Hamiltonian. This protocol, reviewed in the next section, has provided evidence for the smoothness of the eternal black hole and the ER=EPR proposal \([11]\). It was further discussed in \([12–14]\) and applied to a class of a-typical pure states in \([15–17]\).

In this paper we develop a similar protocol for one-sided black holes dual to typical pure states in the CFT. This protocol relies on perturbing the Hamiltonian by state-dependent operators and allows us to connect the smoothness of the horizon of a typical pure state to properties of CFT correlators. Moreover, it provides an explicit CFT realization of an analogue of the Hayden-Preskill protocol \([18]\). More details will be provided in upcoming work \([19]\).
2 Two-sided black hole

The thermofield double state, which is holographically dual to an eternal two-sided AdS black hole \cite{20}, is an entangled state in the tensor product of two identical CFTs (called "left" and "right"),

$$|\text{TFD}\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_E e^{-\frac{\beta E}{T}} |E\rangle_L \otimes |E\rangle_R,$$  \hspace{1cm} (2.1)

where $\beta$ is the inverse temperature and we sum over energy eigenstates. The two CFTs are not interacting, therefore operators on the left and right commute $[O_L, O_R] = 0$ and no information can be transferred between the CFTs. Equivalently, in the bulk the Einstein-Rosen wormhole is not traversable.

In \cite{10} it was argued that the wormhole can become traversable if we couple the two CFTs by a double-trace interaction of the form $V = O_L(0)O_R(0)$ which then allows for geometric transfer of information, if the sign of the coupling is appropriately chosen. An example of a CFT correlator which can diagnose traversability is \cite{12}

$$C = \langle \text{TFD} | [\phi_L(-t), e^{-i\alpha V(t)} \phi_R(t)] | \text{TFD} \rangle.$$  \hspace{1cm} (2.2)

A probe is created on the left by $\phi_L(-t)$ and detected on the right by $\phi_R(t)$. Without the double trace interaction $V$, we would have $[\phi_L(-t), \phi_R(t)] = 0$. When including $V$ certain terms in (2.2) grow exponentially with $t$, as typical for out-of-time-order commutators in chaotic systems \cite{21}. Around scrambling time $t = \frac{\beta}{2\pi} \log S$, we see \cite{12} a signal in the correlator (2.2) representing the probe crossing the wormhole, thus demonstrating smoothness of the horizon of the two-sided eternal black hole.

3 One-sided black hole

We consider a typical state in a large $N$ holographic CFT, which can be thought of as a random superposition of energy eigenstates

$$|\Psi_0\rangle = \sum_{E_i \in (E_0, E_0 + \delta E)} c_i |E_i\rangle,$$  \hspace{1cm} (3.1)

from a narrow energy band of width $\delta E \sim O(N^0)$ and $c_i$ are randomly chosen with the uniform Haar measure. We take $E_0$ to be in the regime dominated by a large AdS black hole in the bulk.

These are almost time-independent equilibrium states. The bulk dual contains at least the exterior of the black hole. It has been proposed \cite{5-7} that the interior can be described using the “mirror operators”, denoted as $\hat{O}$. These operators play a role similar to $O_L$ in the two-sided black hole and we will use them to perform an analogue of the experiment discussed in the previous section.

We will now review the mirror operator construction. First we define a “small algebra” $\mathcal{A}$ corresponding to simple observables in effective field theory. Then, given a typical black hole microstate $|\Psi_0\rangle$, we define the “small Hilbert space”, also called code-subspace, as

$$\mathcal{H}_{|\Psi_0\rangle} = \text{span}\{\mathcal{A}|\Psi_0\rangle\}.$$  \hspace{1cm} (3.2)

This subspace is the one relevant for describing effective field theory in the bulk.
Figure 1. Conjectured Penrose diagram of a typical black hole microstate, to leading order in $1/N$, with a probe created by mirror operators (blue) and two negative energy shockwaves (orange).

If $|\Psi_0\rangle$ is a black hole microstate, it follows [5–7] that the representation of the algebra $\mathcal{A}$ on the subspace $\mathcal{H}_{|\Psi_0\rangle}$ is reducible and the algebra has a non-trivial commutant $\mathcal{A}'$. The commutant $\mathcal{A}'$ can be concretely identified by an analogue of the Tomita-Takesaki construction\footnote{See [22] for a physics-motivated introduction, [5–7] for application to the black hole interior and [23] for a recent review of applications in QFT.} and it is natural to associate $\mathcal{A}'$ with the left region of the extended AdS-Schwarzschild solution.

Following this, we can define the mirror operators on the code subspace to act as

\[
\tilde{O}_\omega |\Psi_0\rangle = e^{-\frac{i}{\hbar}H} O_\omega e^{\frac{i}{\hbar}H} |\Psi_0\rangle, \\
\tilde{O}_\omega O_{\omega_1} \cdots O_{\omega_n} |\Psi_0\rangle = O_{\omega_1} \cdots O_{\omega_n} \tilde{O}_\omega |\Psi_0\rangle, \\
[H, \tilde{O}_\omega] O_{\omega_1} \cdots O_{\omega_n} |\Psi_0\rangle = \omega \tilde{O}_\omega O_{\omega_1} \cdots O_{\omega_n} |\Psi_0\rangle. \tag{3.3}
\]

Here $O_\omega$ denote the Fourier modes in time of single-trace operators. The extension of the operators on the rest of the Hilbert space is irrelevant for the following calculations. We notice that $[O, \tilde{O}] = 0$ only inside the code subspace and may be nonzero as an operator.

While the definiton of $\tilde{O}$ to subleading orders in $1/N$ is not unique, and in particular it will be related to the details of gravitational dressing of local bulk operators, for the purposes of this paper we extend these equations even when we include $1/N$ effects. Possible alternate extensions at the $1/N$ level may modify the interpretation of the bulk geometry at the subleading (i.e. $1/N$) order, but this would not change qualitatively the main point of the thought experiment which we consider, which is already visible in correlators of $O(1)$ size.

The equations (3.3) are defined only for modes with $|\omega| < \omega_*$ where $\omega_*$ is a large, but $N$-independent frequency. Because of this restriction, it is not meaningful to define the mirror operators for sharply localized operators $O(t)$. Moreover, the time argument of smeared mirror operators in position space is assigned so that $\langle \Psi_0 | O(t_1) \tilde{O}(t_2) | \Psi_0 \rangle$ depends on $t_1 + t_2$. This also determines the time-ordering of mirror operators. We emphasize that these operators are explicitly time-dependent, as will be discussed in more detail in [19].

Taking into account the mirror operators, it is natural to conjecture that the geometry dual to a typical state contains not only the exterior, but also the black and white hole.
interiors, as well as part of the left region. This was recently emphasized in [24]. However, we do not expect to be able to describe the full left asymptotic region due to the restriction in the frequencies $|\omega| < \omega_*$. This restriction introduces an effective cutoff of the left region, whose nature will be described more precisely in [19].

We consider the mirror operators as gravitationally dressed with respect to the right. If we call $M, \tilde{M}$ the mass of the solution measured on the right, left respectively, the first law [25–29] applied to the two-sided Cauchy slice up to the left cutoff implies

$$
\delta M - \delta \tilde{M} = \delta K^\text{full}_{\text{bulk}},
$$

where $K^\text{full}_{\text{bulk}} = \int_{\Sigma} \xi T^\text{full}_{\text{bulk}}$ and $\xi$ is the Killing vector field. This can naturally be split into the right and left contributions $K^\text{full}_{\text{bulk}} = K - \tilde{K}$. Since the operators are right-dressed, we have $\delta \tilde{M} = 0$. This means that in the code subspace the CFT Hamiltonian acts as

$$
H = M = E_0 + K^\text{full}_{\text{bulk}},
$$

where $E_0$ is the energy of $|\Psi_0\rangle$.

We are now ready to set up the one-sided analogue of the double trace deformation protocol of [10], which will allow us to extract particles from behind the horizon.

A particle in the left region can be created in two ways. The first way is to actively perturb the CFT Hamiltonian at time $t$ by a “mirror-quench” $\tilde{\phi}(-t)$. The perturbation by $\tilde{\phi}(-t)$ creates a probe in the bulk indicated by the blue line in the figure 1. Without other perturbations the probe would end up in the singularity. The second way to create a particle in the left region is to consider a non-equilibrium state of the form

$$
U(\phi)|\Psi_0\rangle = e^{-\alpha_H} U(\phi)e^{\alpha_H} |\Psi_0\rangle.
$$

These states were extensively discussed in [24]. For definiteness, we will consider the first scenario.

After creating an excitation in the left region by perturbing the CFT with $\tilde{\phi}(-t)$, we perturb the CFT Hamiltonian by $e^{igV}$, where $V = \mathcal{O}(0)\tilde{\mathcal{O}}(0)$. With the appropriate choice of the sign of $g$, this creates two negative energy shockwaves as indicated in figure 1. When analyzing the trajectory of the probe in the region around $t = 0$, one should take into account the effect of the gravitational dressing of the $\tilde{\mathcal{O}}$ operators creating the shockwave [19]. Eventually, the probe particle intersects the right negative energy shockwave and thus undergoes a time-advance. This allows it to escape the horizon and to come out in the right region, where it can finally be detected by $\phi(t)$. This is captured by the correlator

$$
C' \equiv \langle \Psi_0 | [\tilde{\phi}(-t), e^{-igV}\phi(t)e^{igV}] |\Psi_0\rangle.
$$

The conjectured bulk geometry of figure 1 predicts that this correlator should show a sharp signal at $t \approx \frac{\beta}{2\pi} \log S$, similar to that of (2.2). The presence of a signal of the expected form in the CFT correlator (3.7) is thus a necessary (though not sufficient) condition that the conjectured bulk geometry is the one described above and that the horizon is smooth.

\footnote{Some care has to be taken about the interpretation of the bulk geometry to subleading order in $1/N$ [19].}
4 Comparison with two-sided case

We will now argue that the correlators $C$ and $C_0$ are the same in the large $N$ limit and that, therefore, the traversability of the two-sided black hole provides evidence for the smoothness of the horizon of the one-sided black hole. The argument is as follows. Operators in the TFD state obey the relations

\begin{align}
O_L;!j_{TFD}&= e^{\frac{\beta H}{2}} O_L;!, \\
O_L;!O_R;1 \cdots &O_R; n |_{TFD} = \omega O_{L};! O_{R};! 1 \cdots O_{R};! n |_{TFD}, \tag{4.1}
\end{align}

where $\hat{H} \equiv H_R - H_L$.

Now, we consider the correlator $C$ defined in (2.2) and we convert all left-CFT operators into right-CFT operators by repeatedly using the equations above. As a result, the correlator $C$ takes the form

\begin{equation}
C = \langle TFD|\chi(\phi_R, O_R)|TFD\rangle, \tag{4.2}
\end{equation}

where $\chi$ is some time-dependent expression involving only right-CFT operators. We can also write this as

\begin{equation}
C = \frac{1}{Z} \text{Tr}[e^{-\beta H}\chi(\phi, O)], \tag{4.3}
\end{equation}

where we have dropped the subscript $R$.

We now consider the correlator $C_0$ defined in (3.7) for the one-sided black hole. We follow a similar procedure by using equations (3.3) to convert all mirror operators into normal operators. The important point now is that by comparing equations (3.3) and (4.1) we will get exactly the same string $\chi$, i.e.

\begin{equation}
C_0 = \langle 0|\chi(\phi, O)|0\rangle. \tag{4.4}
\end{equation}

We have thus reduced the question about the smoothness of the horizon of a one-sided black hole, to a specific question about CFT expectation values of ordinary (non-mirror) CFT operators. In particular, smoothness requires the proximity of the expectation value of $\chi(\phi, O)$ in the thermal ensemble $e^{-\beta H}/Z$ and a typical pure state $|\Psi_0\rangle$. This is a well-defined CFT question which can in-principle be answered. It is important to notice that this condition needs to hold only for modes with $|\omega| < \omega_k$.

5 A conjecture

We conjecture that in the large $N$ limit, and for modes with $|\omega| < \omega_k$ we have

\begin{equation}
\lim_{N \to \infty} C' = \lim_{N \to \infty} C. \tag{5.1}
\end{equation}

As discussed above, this would provide evidence for the smoothness of the one-sided horizon.
The first step towards motivating (5.1) is to notice that based on general arguments, expectation values on typical pure states differ by $e^{-S}$ from those in the microcanonical ensemble $\rho_m$ of small energy spread \[30\]. Hence, we have

$$C' = \text{Tr}[\rho_m \mathcal{X}(\phi, \mathcal{O})] + O(e^{-S}).$$

(5.2)

This means that to establish (5.1), we need to compare the expectation value of $\mathcal{X}$ in the canonical and microcanonical ensembles

$$C = \frac{1}{Z} \text{Tr}[e^{-\beta H} \mathcal{X}(\phi, \mathcal{O})],$$

$$C'' \equiv \text{Tr}[\rho_m \mathcal{X}(\phi, \mathcal{O})].$$

(5.3)

These two correlators are sensitive only to the diagonal matrix elements of $\mathcal{X}$, since both ensembles in (5.3) are diagonal in the energy basis,

$$\langle E_i | \mathcal{X} | E_j \rangle = f(E_i) \delta_{ij} + R_{ij}.$$  

(5.4)

The Eigenstate Thermalization Hypothesis (ETH) \[31\] postulates that for “simple observables” the diagonal elements $f(E)$ are smooth functions of the energy and that they vary slowly with the energy, in particular $\frac{df}{dE} \sim O(1/S)$. This suggests that the expectation values $C, C''$ differ by $1/S$ corrections — which would imply our desired relation (5.1). However, the observable $\mathcal{X}$ consists of products of simple operators localized at very different times of the order of scrambling time $\frac{\beta}{2\pi} \log S$ and it is not obvious that the ETH will hold for such observables.

In particular, the interesting effect we want to see in the correlators $C, C''$, is coming from certain $1/S$ corrections, which get enhanced by exponential factors $e^{\frac{\beta}{2\pi} t}$ and which become of $O(1)$ at scrambling time. Hence, the non-trivial content of the conjecture (5.1) is that these “chaos-enhanced” $1/S$ corrections are the same in the two ensembles. Relatedly, it suggests that if some operators obey ETH, their product will also obey ETH, even when the operators are widely separated.

Some evidence for this conjecture follows from the observation that the ETH is expected to be robust under multiplication of operators \[4, 19\], at least for small separations in time between operators and for extremely large time separations, where the matrix elements become almost totally uncorrelated. It is natural that the same is true for intermediate times, which include times of the order of scrambling time.

Further evidence can be found by considering some simple models. Firstly, in large $c \leq 2d$ CFTs with sparse spectrum, it was argued in \[32\] that the commutator of two operators separated by times of the order of the scrambling time is dominated by certain time ordered terms. This can be used to transform the out-of-time-order correlator to a time-ordered correlator, for which it is generally assumed that factorization is still applicable and that would imply our conjecture. Secondly, if we assume that the correlators are dominated by the Virasoro identity block even at scrambling time, then there is evidence \[33\] that (5.1) follows.
Figure 2. Numerics in SYK model for \( N = 24 \). Left: \( \langle \{\psi^I(t),\psi^J(0)\}^2 \rangle \) in thermal (blue) and typical pure state (red). The scrambling time is designated by the vertical line. Right: diagonal matrix elements \( \langle E_i|\{\psi^I(t),\psi^J(0)\}^2|E_j\rangle \) for \( t \approx \) scrambling time. It shows slow variation with the energy, compatible with the ETH, within the dominant regions of the canonical ensemble (blue) and the microcanonical ensemble (red).

Finally, we have performed some numerical studies in the SYK model [34–36], a model of \( N \) Majorana fermions in 0+1 dimensions with the Hamiltonian

\[
H_{\text{SYK}} = \sum_{IJKL} J_{IJKL} \psi^I \psi^J \psi^K \psi^L
\]

The couplings \( J_{IJKL} \) are selected randomly from a Gaussian distribution. The analogue of the \( C, C' \) correlators can be represented in the SYK model in terms of correlators of the fermions \( \psi^I(t) \) at time separations of the order of scrambling time, evaluated on the canonical ensemble and on a typical pure state respectively. More details can be found in [19]. In the numerical comparison of the out-of-time-order correlators \( C, C' \) at scrambling time, we find good agreement as shown in figure 2. Unfortunately for the values of \( N \) that we were able to study numerically, there is no large separation between thermalization and scrambling times.

The SYK model does not have an Einstein gravity dual, however it captures some features of black hole physics, in particular it is maximally chaotic. The proximity of \( C, C' \) that we observe numerically in SYK, does not logically imply that theories with Einstein gravity duals (such as the \( N = 4 \) SYM at strong coupling) will also obey the conjecture (5.1). To address that further investigation would be needed. Results which appeared after the first version of this paper provide evidence in favor of (5.1) for a class of theories with Einstein duals [33]. These results, in combination with the numerical evidence that we find in SYK, suggest that the conjecture (5.1) may be true for a larger class of strongly coupled theories. It would be interesting to identify this class more precisely.

6 An analogue of Hayden-Preskill

We observe that the mirror operators \( \tilde{O} \) discussed in the previous sections realize an analogue of the Hayden-Preskill protocol, in the form described in [12]. We start with a black
Figure 3. A realization of the Hayden-Preskill protocol: the code subspace approximately factorizes into a tensor product corresponding to the algebras $\mathcal{A}, \mathcal{A}'$. These tensor factors are entangled and provide the reservoir of EPR pairs needed to perform the teleportation. Here $U, \tilde{U}$ is time evolution in the CFT and $V = \mathcal{O}\tilde{\mathcal{O}}$ denotes the perturbation of the CFT Hamiltonian.

hole in AdS dual to a microstate $|\Psi_0\rangle$. At some time $t_0 \approx -t_S$ (here $t_S$ is scrambling time), we throw a qubit from the boundary into the black hole. This qubit is created in the bulk by acting with the CFT operator $U_c = e^{i\phi(t_0)}$ (appropriately smeared). We wait until the particle has been absorbed, and then we ask what is the CFT operator we need to measure in order to extract the quantum information of the qubit.

One natural way to do this, is by perturbing the CFT Hamiltonian by an interaction of the form $V = \mathcal{O}(0)\tilde{\mathcal{O}}(0)$ with an appropriate coupling constant, which produces two negative energy shockwaves. The infalling particle collides with one of the shockwaves (the “mirror shock”) and undergoes a time-advance, pushing it into the left region. It can then be measured by the mirror operator $\tilde{\phi}(t_S)$. The result of this measurement is captured by a correlator similar to (3.7), with the roles of $\phi$ and $\tilde{\phi}$ reversed. The conjecture of the previous section implies that this correlator can extract the quantum information of the probe.

The Hayden-Preskill protocol can only be applied after the half-point of evaporation, when the black hole is maximally entangled with the early radiation. The analogue statement in our case is that in order to define the operators $\tilde{\phi}, \tilde{\mathcal{O}}$, one needs to have knowledge of the microstate, as the $\tilde{\phi}, \tilde{\mathcal{O}}$’s are state-dependent operators. We remind the reader that even in the original Hayden-Preskill protocol, the decoding operation is state-dependent.

7 Comments

We formulated a necessary condition for the smoothness of the horizon of a typical black hole microstate in terms of CFT correlators of local operator at scrambling time. We argued that smoothness of the horizon requires that these correlators are similar in the canonical and microcanonical ensembles. We provided some preliminary evidence in favor of this conjecture. These observations imply that for certain purposes it is meaningful to consider part of the left region of the extended AdS-Schwarzchild geometry.
Our arguments made use of state-dependent operators. Since the boundary observer has in principle unlimited resources, this fits within the conventional framework of quantum mechanics. Indirectly, this provides evidence for the relevance of the state-dependent operators for the infalling observer. Further details and open questions will be discussed in an upcoming longer article [19].

Acknowledgments

We would like to thank G. Gur-Ari, D. Harlow, D. Jafferis, E. Kiritsis, J. Maldacena, L. Motl, R. Myers, S. Raju, J. Sonner, H. Verlinde, S. Wadia for useful discussions. RvB would like to thank NCCR SwissMAP of the Swiss National Science Foundation for financial support. SFL thanks The Netherlands Organization for Scientific Research (NWO) for financial support and CERN for hospitality during the completion of this work. KP would like to thank KNAW and NWO for support and the University of Crete for hospitality.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


[34] A. Kitaev, *A simple model of quantum holography*,
http://online.kitp.ucsb.edu/online/entangled15/kitaev/.
